

Einfluss von Personen- und Webseitenmerkmalen auf den Such- und Evaluations-
prozess beim Lösen eines Online-Informationsproblems

Dissertation
zur Erlangung des Grades eines Doktors
der wirtschaftlichen Staatswissenschaften
(Dr. rer. pol.)
des Fachbereichs Rechts- und Wirtschaftswissenschaften

der Johannes Gutenberg-Universität Mainz
vorgelegt von
M. Sc., Alice Laufer

in Mainz

im Jahre 2024

Erstgutachterin: Prof. Dr. Olga Zlatkin-Troitschanskaia

Zweitgutachter: Prof. Dr. Christian Schemer

Tag der mündlichen Prüfung: 12.06.2024

Danksagung

Ich bedanke mich von Herzen bei nachstehenden Personen:

Ich danke meinen beiden betreuenden PIs Prof. Dr. Olga Zlatkin-Troitschanskaia sowie Prof. Dr. Christian Schemer für die sehr gute Betreuung und die Hilfestellung bei u. a. der Ideenfindung, Konzeption oder Problemlösung.

Ich danke dem gesamten DIAPASON-Team für den fruchtbaren Austausch in jedem DIAPASON-Meeting, der durch die Interdisziplinarität und verschiedenen Blickwinkel wertvolle Ideen generiert hat.

Ich danke dem BRIDGE-Team erstens für die Datenerhebung und -aufbereitung, weiterhin aber auch für die wöchentlichen Meetings, die neue Ideen zutage gefördert oder so manche Probleme gelöst haben.

Ich danke Denis für die Zusammenarbeit hinsichtlich der mir damals noch unbekannten Auswertungsmethoden und für die Unterstützung in diesem Feld.

Ich danke Kevin dafür, dass er mir bei spontanen Ideen stets weiterhalf und sich die Zeit für mich und meine komplexen Anfragen nahm.

Ich danke Dominik dafür, dass ich ihn zu jeder Zeit mit Fragen nerven konnte und er mir insbesondere bei den schwierigeren Auswertungen zur Seite stand.

Ich danke Sebastian dafür, dass er mich überhaupt für das DIAPASON-Stipendium vorgeschlagen hat.

Ich danke Amina und Anika für die emotionale Unterstützung, die Pep Talks und das mir vermittelte Gefühl, dass wir diesen Weg gemeinsam gehen und schaffen.

Ich danke Julia für ihr offenes Ohr bei allen Fragen und Sorgen rund um den Ideenfindungs-, Schreib- und Veröffentlichungsprozess sowie ihr Verständnis bei Spontanabsagen.

Ich danke dir, Tom, weil du mein härtester Kritiker bist und ich dadurch das Beste aus mir und meiner Dissertation herausholen konnte. Ich danke dir für die allabendlichen, langen Diskussionen, für dein ständiges Fragen nach dem „Warum“, für deine Akzeptanz und Toleranz gegenüber meinem Selbstzweifel und deiner Energie mir in solchen Phasen wieder Mut zu machen.

Zusammenfassung

Die Fähigkeit des kritischen Online-Denkens, d. h. sich kritisch u. a. mit der Qualität und Glaubwürdigkeit von Online-Informationen auseinanderzusetzen, ist seit der Entwicklung KI-gestützter Sprachmodelle notwendiger denn je. Zwei Faktoren können den Prozess der Informationsverarbeitung und das kritische Online-Denken beeinflussen: Faktoren der beurteilenden Person sowie der zu beurteilenden Webseiten. Diese Einflussfaktoren zu ermitteln und zu verstehen, liefert Ansatzpunkte für die Förderung des kritischen Online-Denkens. Die Arbeit verwendet einen Mixed-Methods-Ansatz mit dem Ziel, den Einfluss von Personen- und Webseitenmerkmale auf das Such- und Evaluationsverhalten beim Lösen eines (generischen) Online-Informationsproblems zu ermitteln. 64 Universitätsabsolventen recherchieren mithilfe des Internets zur Fragestellung, ob E-Bikes gesundheitsfördernd sind. Ihre personellen Merkmale, wie das Kognitionsbedürfnis und die Informationsüberlastung, wurden mittels eines Fragebogens erhoben. Das Such- und Evaluationsverhalten sowie die besuchten Webseiten und deren Merkmale wurden anhand getrackter Logdaten ermittelt. Die Regressionsanalysen zeigen, dass sowohl Personen- als auch die aggregierten Webseitenmerkmale einen Einfluss auf zwei Schritte im Suchprozess (Nachrecherche und Linkverfolgung) haben. Die ganzheitliche Betrachtung der Personen- und Webseitenmerkmale liefert somit wichtige Erkenntnisse, um den Online-Informationsproblemlöseprozess besser verstehen zu können. Informationssuchende sollten für den Einfluss der eigenen Dispositionen sowie der Webseitenmerkmalen beim Lösen von Informationsproblemen sensibilisiert und geschult werden.

The ability of critical online reasoning, i.e., critically assessing the quality and credibility of online information, is especially necessary since the development of AI-based language models. Two factors can influence the process of information processing and critical online reasoning: factors of the evaluating person and of the websites being evaluated. Identifying and understanding these factors provides starting points for promoting critical online reasoning. This study uses a mixed-methods approach to determine the influence of personal and website characteristics on search and evaluation behaviors when solving a (generic) online information problem. Sixty-four university graduates conduct online research on whether e-bikes are health-promoting. Their personal characteristics, such as need for cognition and information overload, were collected using a questionnaire. The search and evaluation behavior and the visited websites and their characteristics were determined based on tracked log data. The regression analyses show that personal and aggregated website characteristics influence two steps in the search process (further research and link following). The holistic consideration of personal and website characteristics provides important insights to better understand the online information problem-solving process. Information seekers should be made aware of and trained on the influence of their own dispositions and website characteristics when solving information problems.

Inhaltsverzeichnis

Inhaltsverzeichnis	V
Abbildungsverzeichnis	IX
Tabellenverzeichnis	X
Abkürzungsverzeichnis	XII
1. Relevanz und Ziel der Arbeit.....	1
1.1. Schwerpunktsetzung und Forschungsfrage.....	3
1.2. Aufbau der Arbeit	5
2. Theoretischer Rahmen des Forschungsvorhabens.....	6
2.1. Begriffliche Grundlagen: Glaubwürdigkeitsdefinition, innere Differenzierung und abzugrenzende Konstrukte	7
2.2. Informationsverarbeitungs- und Selektionsmodelle	13
2.2.1. Grundlagen der Selektion und Informationsverarbeitung.....	13
2.2.2. Das Heuristisch-Systematische Modell als Vertreter der Zwei-Prozess-Theorien der klassischen Informationsverarbeitungsmodelle	19
2.2.2.1. Ausgewählte Entscheidungsregeln bei der Informationsverarbeitung.....	21
2.3. Digitale Informationsverarbeitungs- und Selektionsmodelle.....	25
2.3.1. Einführender Überblick zum 3-S-Modell des Vertrauens	26
2.3.2. Das IPS-I-Modell	27
2.3.3. Das Konstrukt des Critical Online Reasoning	30
2.4. Transfer und Anwendung der theoretischen Grundlagen in den Forschungskontext	32
2.5. Theoretische Herleitung der Prozess- und Performanzindikatoren als abhängige Variablen	37
2.5.1. Prozessindikatoren aus dem Suchprozess.....	37
2.5.2. Performanzindikatoren aus dem Evaluationsprozess	38
2.6. Merkmale der Akteure im Informationsverarbeitungsprozess	41
2.6.1. Für Informationsverarbeitung relevante Merkmale im Informationsangebot.....	41
2.6.1.1. Aktualität	45
2.6.1.2. Verständlichkeit.....	46
2.6.1.3. Transparenz.....	47
2.6.1.4. Quellenherkunft.....	49
2.6.1.5. Angabe von Quellen.....	49
2.6.1.6. Struktur der Information bzw. Webseite.....	51
2.6.1.7. Faktische Korrektheit	52
2.6.1.8. Statistiken	53
2.6.1.9. Vollständigkeit.....	54

2.6.1.10.	Sprachstil	54
2.6.1.11.	Wissenschaftlichkeit	55
2.6.1.12.	Argumentqualität	56
2.6.1.13.	Narrative	57
2.6.1.14.	Spezialisierungsgrad	58
2.6.1.15.	Design	58
2.6.1.16.	Genre	59
2.6.1.17.	Kommerzielles Interesse	60
2.6.1.18.	Bekanntheit	62
2.6.1.19.	Reputation	63
2.6.1.20.	Angabe des Autors	64
2.6.1.21.	Foto eines Autors	65
2.6.2.	Theoretische Zusammenfassung der Webseitenmerkmale	66
2.6.3.	Individuelle Merkmale	68
2.6.3.1.	Need for Cognition	69
2.6.3.2.	Need for Cognitive Closure	70
2.6.3.3.	Information Overload	72
2.7.	Forschungsdesiderat	74
3.	Hypothesenherleitung	76
3.1.	Direkte Effekte des Suchprozesses	78
3.1.1.	Need for Cognition (NfC) und Suchprozess	78
3.1.2.	Need for Cognitive Closure (NfCC) und Suchprozess	79
3.1.3.	Information Overload (IO) und Suchprozess	80
3.1.4.	Einfluss der wahrgenommenen Glaubwürdigkeit besuchter Webseiten	81
3.2.	Interaktionseffekte in der Suchstrategie	82
3.3.	Direkte Effekte im Evaluationsprozesses	85
3.3.1.	Need for Cognition (NfC) und Evaluationsprozess	86
3.3.2.	Need for Cognitive Closure (NfCC) und Evaluationsprozess	87
3.3.3.	Information Overload (IO) und Evaluationsprozess	88
3.3.4.	Need for Cognition (NfC) und wahrgenommene Glaubwürdigkeit selektierter Webseiten	89
3.3.5.	Need for Cognitive Closure (NfCC) und wahrgenommene Glaubwürdigkeit selektierter Webseiten	89
3.3.6.	Information Overload (IO) und wahrgenommene Glaubwürdigkeit selektierter Webseiten	90
3.3.7.	Wahrgenommene Glaubwürdigkeit selektierter Webseiten und Performanzindikatoren	90
3.4.	Interaktionseffekte im Evaluationsprozess: Einfluss der wahrgenommenen Glaubwürdigkeit selektierter Webseiten	91

4. Methodik	94
4.1. Projektbeschreibung BRIDGE	95
4.1.1. Erhebungsablauf und -design.....	96
4.1.2. Beschreibung und Zusammensetzung der Stichprobe	98
4.2. Datenaufbereitung und -optimierung	98
4.2.1. Vorgehen zur Operationalisierung der Variablen	101
4.2.2. Operationalisierung der Personenmerkmale als unabhängige Variable.....	102
4.2.3. Problematische Webseitenmerkmale	103
4.2.4. Operationalisierung der Webseitenmerkmale als Zwischenergebnis der vorbereitenden Analyse.....	108
4.2.4.1. Aktualität	109
4.2.4.2. Verständlichkeit.....	110
4.2.4.3. Transparenz.....	111
4.2.4.4. Quellenherkunft (primär, sekundär).....	112
4.2.4.5. Angabe von Quellen wird zu Nachweisqualität.....	112
4.2.4.6. Kommerzielles Interesse	113
4.2.4.7. Genre.....	114
4.2.4.8. Bekanntheit.....	115
4.2.4.9. Angabe des Autors.....	115
4.2.4.10. Blockweise praxisbezogene Zusammenfassung der Webseitenmerkmale	115
4.2.4.11. Zwischenfazit zu den verbliebenen Webseitenmerkmalen	117
4.2.4.12. Personenbezogene Glaubwürdigkeitsindices auf Basis der Webseitenmerkmale	118
4.2.5. Generierung der Prozess- und Performanzindikatoren als abhängige Variablen	121
4.3. Datenanalyse	129
5. Ergebnisse.....	131
5.1. Deskriptive Ergebnisse.....	131
5.1.1. Beurteilung der Variablengüte der Personen- und Webseitenmerkmale.....	131
5.1.2. Beschreibung der Häufigkeiten der abhängigen Variablen	135
5.2. Korrelative Ergebnisse	137
5.3. Ergebnisse der Inferenzstatistik	142
5.3.1. Prämissenüberprüfung	142
5.3.2. Einfluss der Personen- und Webseitenmerkmale auf die abhängigen Variablen der Suchstrategie	150
5.3.2.1. Modell I A-B mit Spezifizität der Suchbegriffe.....	150
5.3.2.2. Modell II A-C mit Nachrecherche.....	155
5.3.2.3. Modell III A-C mit Weiterverlinkung anklicken.....	169

5.3.3. Einfluss der Personen- und Webseitenmerkmale auf die abhängigen Variablen der Evaluationsstrategie	178
5.3.3.1. Modell IV A-B mit Korrektheit der Quelleneinschätzung	178
5.3.3.2. Modell V A-B mit Systematische Begründung	183
5.3.3.3. Modell VI A-C mit Argumentationsstruktur	186
6. Diskussion und Interpretation der Befunde	190
6.1. Non-linearer Zusammenhang und Glaubwürdigkeitsbeurteilung	191
6.2. Defizite bezüglich COR	203
7. Limitationen	210
8. Implikationen für Forschung und Praxis	214
9. Ausblick und Fazit	221
Literaturverzeichnis	226
Anhang	266

Abbildungsverzeichnis

Abbildung 1: Prozessuale Darstellung der subjektiven Glaubwürdigkeitsbeurteilung.	9
Abbildung 2: Abgrenzung der Glaubwürdigkeit zu Vertrauen und Persuasion.	13
Abbildung 3: Makro- und Mikroperspektive der Informationsverarbeitung und Selektion.	14
Abbildung 4: 3-S-Modell, eigene Abbildung nach Lucassen et al. (2013).	27
Abbildung 5: Das IPS-I-Modell, eigene Abbildung in Anlehnung an Brand-Gruwel et al. (2005, 2009) sowie Eisenberg und Berkowitz (1990).	28
Abbildung 6: COR-Konstrukt, eigene Abbildung in Anlehnung an Molerov et al. (2020).	32
Abbildung 7: Transfer der theoretischen Grundlagen in Forschungskontext.	35
Abbildung 8: Einteilung der Webseitenmerkmale in drei Kategorien.	68
Abbildung 9: Exemplarische Darstellung der Hypothese H4a.	82
Abbildung 10: Exemplarische Abbildung für Nachrecherche und NfC.	83
Abbildung 11: Exemplarische Abbildung für Nachrecherche und NfCC.	84
Abbildung 12: Exemplarische Abbildung für Nachrecherche und IO.	85
Abbildung 13: Darstellung der Hypothesen aus Modell VI-A.	92
Abbildung 14: Darstellung der Hypothesen aus Modell VI-C.	93
Abbildung 15: Darstellung der Hypothesen aus Modell VI-C.	94
Abbildung 16: Mixed-Methods-Ansatz im Forschungsvorhaben.	95
Abbildung 17: Finale Einteilung der Webseitenmerkmale bezüglich Kategorien.	117
Abbildung 18: Beispielhafter Boxplot.	146
Abbildung 19: Effekte aus Modell II-A.	159
Abbildung 20: Effekte aus Modell II-B.	163
Abbildung 21: Effekte aus Modell II-C.	167
Abbildung 22: Effekte aus Modell III-B.	174
Abbildung 23: Allgemeines Mediationsmodell.	186
Abbildung 24: Mediationsmodell VI-A NfC und Argumentationsstruktur.	188
Abbildung 25: Mediationsmodell VI-B IO und Argumentationsstruktur.	189
Abbildung 26: Mediationsmodell VI-C NfCC und Argumentationsstruktur.	190
Abbildung 27: Gesamtbetrachtung der Ergebnisse aus Modell III.	193

Tabellenverzeichnis

Tabelle 1: Anwendung des 3-S-Modell für den Forschungskontext.	39
Tabelle 2: Kategorisierung der Merkmale nach unterschiedlichen Erhebungen in Untersuchungen.	42
Tabelle 3: Übersicht Modelle und Variablen.	78
Tabelle 4: Meist aufgerufene unique Webseiten.	101
Tabelle 5: Übersicht der verwendeten Input-, Prozess- und Performanzfaktoren.	102
Tabelle 6: Berechnungsbeispiel des personenbezogenen Glaubwürdigkeitsindices.	119
Tabelle 7: Ergebnisse der qualitativen Analyse der Probandenantwort zur Glaubwürdigkeit.	125
Tabelle 8: Häufigkeitsverteilung der Codes	126
Tabelle 9: Beispielhafte Ausschnitte aus den Stellungnahmen der Probanden.	128
Tabelle 10: Übersicht aller Modelle.	130
Tabelle 11: Deskriptive Statistiken der Personenmerkmale.	131
Tabelle 12: Deskriptive Statistiken der Soziodemografie.	133
Tabelle 13: Deskriptive Statistiken der Merkmale der besuchten und selektierten Webseiten.	134
Tabelle 14: Häufigkeitsverteilung der dichotomen abhängigen Variablen.	135
Tabelle 15: Aufgerufene Webseitenverteilung und Weiterverlinkung.	136
Tabelle 16: Deskriptive Statistiken der metrischen abhängigen Variablen.	137
Tabelle 17: Korrelationen.	141
Tabelle 18: Prämissenübersicht der Modelle	143
Tabelle 19: Modellübergreifende Ausreißer-Analyse	146
Tabelle 20: Modellgüte des Modells I-A NfC und IO und Spezifität des Suchbegriffs.	151
Tabelle 21: Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell I-A NfC, IO und Spezifität der Suchbegriffseingabe.	154
Tabelle 22: Modellgüte des Modells II-A NfC und Nachrecherche	156
Tabelle 23: Ausschnitt der Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-A NfC und Nachrecherche.	157
Tabelle 24: Hypothesenentscheidung Modell II-A.	160
Tabelle 25: Modellgüte des Modells II-B IO und Nachrecherche	161
Tabelle 26: Ausschnitt der Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-B IO und Nachrecherche.	162
Tabelle 27: Hypothesenentscheidung Modell II-B.	164
Tabelle 28: Modellgüte von Modell II-C NfCC und Nachrecherche.	165
Tabelle 29: Ausschnitt der Ergebnisse der Berechnung der logistischen Regression von Modell II-C NfCC und Nachrecherche.	166

Tabelle 30: Hypothesenentscheidung Modell II-C.....	168
Tabelle 31: Modellgüte von Modell III-A NfC und Weiterverlinkung anklicken.....	169
Tabelle 32: Ausschnitt der Ergebnisse der schrittweisen Berechnung von Modell III-A NfC und Weiterverlinkung.....	170
Tabelle 33: Hypothesenentscheidung aus Modell III-A.	171
Tabelle 34: Modellgüte des Modells III-B IO und Weiterverlinkung anklicken.	172
Tabelle 35: Ausschnitt der Ergebnisse der schrittweisen Berechnung von Modell III-B IO und Weiterverlinkung.....	173
Tabelle 36: Hypothesenentscheidung aus Modell III-B.	176
Tabelle 37: Modellgüte des Modells III-C NfCC und Weiterverlinkung anklicken.	177
Tabelle 38: Ausschnitt der Ergebnisse der schrittweisen Modellberechnung III-C NfCC und Weiterverlinkung anklicken.....	178
Tabelle 39: Modellgüte von Modell IV-A.	180
Tabelle 40: Ergebnisse der linearen Regression vom finalen Modell IV-A NfC und IO auf Korrektheit der Quelleneinschätzung.....	181
Tabelle 41: Modellgüte des Modells IV-B.....	182
Tabelle 42: Modellgüte des Modells V-A NfC, IO und systematische Begründung.	183
Tabelle 43: Ausschnitt der Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-A NfC, IO und systematische Begründung.....	184
Tabelle 44: Modellgüte des Modells V-B NfCC und systematische Begründung.....	185
Tabelle 45: Ausschnitt der Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-B NfCC und systematische Begründung.....	185

Abkürzungsverzeichnis

ASL	Average Sentence Length (durchschnittliche Satzlänge)
ASW	Average Number of Syllables per Word (durchschnittliche Wortlänge)
AUC	Area Under Curve
AV	Abhängige Variable(n)
BRIDGE.....	Berufspraktische Bildungsprozesse im Recht- und Lehramtsreferendariat sowie der Medizin unter Nutzung digitaler Medien
CIE	Critical Information Evaluation
COR	Critical Online Reasoning
df.....	Freiheitsgrade
DOM-COR.....	domain-specific Critical Online Reasoning
ELM.....	Elaboration-Likelihood-Modell
F.....	Wert der F-Verteilung
GEN-COR	generic Critical Online Reasoning
H.....	Hypothese
HSM	Heuristic-Systematic Model
i. S. v.	im Sinne von
IO	Information Overload
IPS-I-Modell	internetbezogenes Informationsproblem-Löse-Modell
IQA.....	Interquartilsabstand
KI.....	Konfidenzintervall
LLQ	Likelihood-Quotienten-Tests
LLR	Likelihood-Ratio-Statistik
M.....	Mittelwert
N bzw. n	Gesamtstichprobengröße bzw. Teilstichprobengröße
NfC.....	Need for Cognition
NfCC	Need for Cognitive Closure
OIA.....	Online Information Acquisition
OR	Odds Ratio
REAS	Reasoning with Evidence, Argumentation and Synthesis
ROC	Receiver Operating Characteristics
SD	Standardabweichung
SResid	studentisierte Residuen
std	standardisiert
t.....	Wert der T-Verteilung
URL.....	Uniform Resource Locator
UV	Unabhängige Variable(n)

VIF	Varianz-Inflations-Faktor
zentr	zentriert
β	Standardisierter Regressionskoeffizient
η	Eta
χ^2	Chi ²

In dieser Arbeit wird ein Punkt als Dezimaltrennzeichen verwendet und es wird auf drei Nachkommastellen gerundet.

Die Zitation erfolgt nach APA 7.

1. Relevanz und Ziel der Arbeit

„Wagen Sie, selbst zu denken“ (Voltaire, 1786, S. 417)

KI-basierte Textgenerierungsprogramme, wie etwa das populärste und hier exemplarisch genannte ChatGPT (*ChatGPT*, o. J.), können das Berufs- oder Alltagsleben durch z. B. Internetrecherchen erleichtern. ChatGPT verfügt über Such-, Selektions- und Zusammenfassungsfunktionen (B. D. Lund & Wang, 2023) und kann dem Nutzer¹ diese Tätigkeiten und somit einen großen Teil des Rechercheprozesses abnehmen (Rasul et al., 2023). Die Kehrseite dieser Entwicklung könnte jedoch sein, dass Nutzer insgesamt nicht mehr selbst suchen, selektieren oder denken. Sie verlassen sich jetzt schon auf das Ergebnis von KI-Textgeneratoren, ohne die Informationen oder ihre Glaubwürdigkeit kritisch zu hinterfragen (Fuchs, 2023). Auch wenn die Ausgaben von solchen KI-basierten Systemen das gesamte Internet als Quelle nutzen, können sie dennoch fehlerhaft sein (Qadir, 2023). Insbesondere in Zeiten einer „Infodemie“ – einer Überflutung mit Informationen und Fehlinformationen zum Zeitpunkt eines Krankheitsausbruchs (WHO, 2021) – wie Covid-19, wird das Ausmaß vom Lesen von fehlerhaften oder unglaubwürdigen Informationen, ohne kritisches Einschätzen, deutlich. Menschen konnten nur schlecht glaubwürdige von unglaubwürdigen Informationen unterscheiden (Paus & Börsch-Supan, 2020, S. 6)² und der wenig versierte bzw. unkritische Umgang mit (Fehl-)Information hatte teilweise tödliche Folgen: So tranken Menschen beispielsweise aufgrund einer Fehlinformation und ohne diese zu überprüfen Methanol, und starben (Coleman, 2020).

Die Ursachen für diese noch immer andauernde Entwicklung liegt in der Vernetzung von mehreren früheren Entwicklungen: Zum einen ermöglicht die Digitalisierung die schnelle Verbreitung von sehr vielen Informationen im Online-Kontext (Otto et al., 2020). Diese Informationen werden aber nicht, wie bei traditionellen Medien, durch so genannte Gatekeeper überprüft (Koltay, 2017; Neuberger, 2007; Rains & Karmikel, 2009). Gatekeeper wählen die wichtigsten Informationen für das Publikum aus, sind beispielsweise Reviewer oder Journalisten, und haben in der Regel gewisse Qualitätssicherungsstandards für ihre Berichterstattung (*Medienpolitik: Gatekeeper*, o. J.). Im Online-Kontext kann hingegen jeder alles verbreiten, ohne dass es geprüft werden muss. Dies hat u. a. zur Folge, dass viele falsche Informationen im Internet kursieren (Maurer et al., 2018). Gleichzeitig führt die zunehmende Informationsmenge dazu, dass Personen schneller überlastet werden, weil sie an ihre kognitiven Grenzen stoßen (Röetzel, 2019). Es befinden sich folglich zu viele, teilweise fehlerhafte oder unglaubwürdige

¹ Zur besseren Lesbarkeit wird in dieser Arbeit das generische Maskulinum verwendet. Die in dieser Arbeit verwendeten Personenbezeichnungen beziehen sich – sofern nicht anders kenntlich gemacht – jedoch auf alle Geschlechter.

² 64 % der Befragten fällt es schwerer bei Informationen zu Covid-19 im Vergleich zu Informationen aus anderen Bereichen glaubwürdige von unglaubwürdigen Nachrichten zu unterscheiden; $N = 2.064$, 14 - 24 Jahre.

Informationen im Umlauf, zugleich werden die Informationen aufgrund der Informationsüberlastung von Menschen aber immer weniger überprüft.

Wie können Menschen in Anbetracht dieser Entwicklung besser mit Online-Informationen umgehen? Wie können sie trotz der Informationsmenge „gute“ Informationen suchen, selektieren und verwenden? Ein in diesem Vorhaben genutzter Lösungsansatz für einen effizienteren Umgang mit Online-Informationen ist es, für das kritische Denken im Online-Kontext zu sensibilisieren und dieses zu fördern, was in einem bereits etablierten Konstrukt (Molero et al., 2020; Nagel et al., 2020, 2022; Schmidt, 2020) beschrieben wird: das kritische Online-Denken (Englisch: Critical Online Reasoning, im Folgenden abgekürzt mit COR).

Im Zuge der Verbesserung von COR muss allerdings zunächst verstanden werden, inwieweit die Fähigkeit kritisch zu denken und gut mit Informationen umzugehen bei Personen vorhanden ist (Ist-Zustand) und welche Faktoren den Prozess im kritischen Umgang mit Informationen beeinflussen.

Für die Erfassung des Ist-Zustands wird daher der Hochschulbereich gewählt. Es kann zumindest von Masterabsolventen erwartet werden, dass sie gemäß ihrem Bildungsniveau ausreichend gut kritisch denken bzw. Online-Informationen einschätzen können (OECD, 2022). Die Schwerpunktsetzung im darauffolgenden Kapitel wird dies näher spezifizieren.

Auf den Prozess zum Umgang mit Online-Informationen haben zwei wesentliche Faktoren Einfluss: Die Person, die beurteilt, und die Webseite, die beurteilt wird (Wirth et al., 2007). Zum einen sind also individuelle Prädispositionen, so genannte Traits (Dorsch Lexikon der Psychologie, o. J.) daran beteiligt, wie Informationen gesucht, selektiert oder evaluiert werden (Al-Samarraie et al., 2017; Uslu & Durak, 2022). Folgende Forschungsarbeit wird sich hierbei auf das Kognitionsbedürfnis (Cacioppo & Petty, 1982), das Bedürfnis nach Abgeschlossenheit (Webster & Kruglanski, 1994) und die Informationsüberlastung (Roetzel, 2019) fokussieren. Denn diese individuellen Dispositionen können u. a. die Informationsverarbeitung beeinflussen (Wirth et al., 2007). Zudem können sie jederzeit im Rahmen des Studiums erhoben werden, sodass zudem mögliche Handlungsempfehlungen im Zusammenhang damit abgeleitet werden können.

Zum anderen spielen Merkmale des zu beurteilenden Objekts (Wirth et al., 2007), also der Webseite, eine Rolle. Die Merkmale der Webseiten sind situations- und aufgabenspezifisch. Die Wahrnehmung dieser Merkmale durch das Individuum ist folglich ebenfalls situativ und erfasst einen Zustand, einen „State“ (Schmitt, 2022).

Im Sinne der Literatur zu Trait-State-Theorien (Spielberger et al., 1971; Steyer et al., 1999) können Trait und State nicht separat betrachtet werden, da die Messung eines Traits einer

Person nie in einem situativen Vakuum geschieht (Steyer et al., 1999). Obwohl für vorliegende Forschungsarbeit somit zwei unterschiedliche Betrachtungsebenen zu Grunde liegen (Personen und Webseiten), wird sich dennoch bewusst für diese Ebenen entschieden, weil sie im Zusammenspiel zu betrachten sind. Denn eine mit einem zeitlich stabilen Trait ausgestattete Person wird in einer Situation mit Webseiten und ihren Merkmalen konfrontiert, welche situationsspezifisch einen Zustand (z. B. Wahrnehmung dieser) und weiterhin damit verbundenes Verhalten hervorrufen kann. Gemäß der angenommenen Wechselwirkung zwischen Trait und State in der Trait-State-Theorie (Steyer et al., 2015) können Personen- und Webseitenmerkmale in Wechselwirkung treten und einen gemeinsamen Einfluss auf z. B. Verhalten aufweisen. Zudem liegen bislang, wie u. a. auch die spätere theoretische Aufarbeitung des Forschungsstandes herausstellen wird, keine Forschungen vor, die beide Einflussfaktoren im Zusammenwirken und derart detailliert untersuchen (Zlatkin-Troitschanskaia, Hartig, et al., 2021). Insofern rechtfertigt das die Betrachtung beider möglicher Einflussfaktoren.

Um die Unterschiede der Personen im Umgang mit Online-Informationen zu erkennen, werden Performanz- und Prozessschritte vom COR-Konstrukt abgeleitet (Molero et al., 2020).

Die vorliegende Forschungsarbeit hat daher einerseits das Ziel, den Einfluss der Personen- und Webseitenmerkmale auf Prozess- und Performanzschritte beim Lösen eines Online-Informationsproblems zu ermitteln, die wiederum als Indikatoren für die COR-Fähigkeiten herangezogen werden können. Hierzu soll ein holistisches Modell abgebildet werden, in dem sowohl die Personen- als auch Webseitenmerkmale enthalten sind. Ein holistisches Modell ist deshalb dienlich, damit es auf andere Situationen und Kontexte anwendbar ist und um folglich einen besseren Umgang mit Online-Informationen zu gewährleisten.

Zudem soll ein Bewusstsein für den Einfluss und das Zusammenspiel der Personen- und Webseitenmerkmale geschaffen werden, welches eine Verhaltensänderung in Richtung effizientere und weniger fehleranfällige Informationsverarbeitungsstrategie ermöglichen kann. Untrennbar damit verbunden ist das perspektivische Ziel, die Erkenntnisse rund um das kritische Online-Denken und die von den Personen- und Webseitenmerkmalen abhängige Fehleranfälligkeit in die Curricula zu integrieren.

1.1. Schwerpunktsetzung und Forschungsfrage

Die Daten für das vorliegende Forschungsvorhaben wurden aus dem durch Drittmittel geförderten Projekt BRIDGE (*Berufspraktische Bildungsprozesse im Recht- und Lehramtsreferendariat sowie der Medizin unter Nutzung digitaler Medien*)³ (Zlatkin-Troitschanskaia, Brückner, et al., 2021) gewonnen. Aus der Drittmittelförderung sowie der Teilnahme weiterer

³ Gefördert vom Bundesministerium für Bildung und Forschung; Förderkennzeichen: 01JD1906.

Projektbeteiligter an den Daten ergeben sich deshalb Einschränkungen, auf die die Forschende keinen Einfluss mehr hat, ihr jedoch bewusst sind. Mit diesem Bewusstsein soll dennoch versucht werden, die oben genannten Ziele zu erreichen. Hierzu werden kürzlich graduierte Masterabsolventen – Berufsanfänger – herangezogen.

Aktuelle Forschung zeigt, dass Studierende in ihrem Studium vermehrt Online-Medien im Vergleich zu den klassischen Lehrbüchern nutzen, um sich z. B. auf Prüfungen vorzubereiten (Maurer et al., 2020). Das zeigt, dass das Internet als (Lern-)Medium in der Hochschulbildung bereits Einzug gehalten hat, obwohl die klassischen Lehrbücher oder wissenschaftlichen Veröffentlichungen durch die Gatekeeper eigentlich ausreichend glaubwürdige Informationen bereitstellen könnten, ohne dass eine zusätzliche Glaubwürdigkeitsprüfung seitens der Studierenden notwendig ist. Durch die zunehmende Nutzung des Internets als Medium drängt sich die Notwendigkeit der Glaubwürdigkeitsbeurteilung und des kritischen Umgangs mit Informationen verstärkt in den Vordergrund (s. z. B. folgende Reviews: López-González et al., 2023; Zlatkin-Troitschanskaia, Hartig, et al., 2021). Es ist daher notwendig, dass im Rahmen des Studiums ein geschulter, kritischer Umgang mit Online-Informationen gelehrt werden sollte.

Berufsanfänger sollten mit Erreichen ihres Master-Abschlusses ein gewisses Niveau an kritischem Denken erlangt haben (OECD, 2022). Außerdem beginnt damit ein Lebensabschnitt, ab welchem die Hochschulbildung ihre Einflussmöglichkeit auf die Fähigkeit der Probanden kritisch zu denken bzw. Online-Informationen zu hinterfragen, verliert. Fünf Jahre Studium haben im Idealfall die kognitiven Fähigkeiten, welche möglicherweise sowieso schon durch die schulische Bildung vorhaben waren, weiter ausgebaut und Personen zu handlungsfähigen Erwachsenen ausgebildet (OECD, 2022). Das Wissen um das kritische Denken und die weitere Förderung dessen könnte somit auch das Resultat der universitären Lehre sein und erfasst einen Ist-Zustand. Mit dem kürzlichen Abschluss und noch vor Eintritt in die Berufswelt sind Berufsanfänger genau die Zielgruppe, die einen Aufschluss darüber geben kann, inwieweit COR-Fähigkeiten (noch) vorhanden sind.

Für die vorliegende Arbeit werden Berufsanfänger fokussiert, welche jedoch innerhalb des Projektkontextes BRIDGE in drei Domänen (Jura bzw. Rechtswissenschaften, Lehramt, Medizin) erhoben wurden. Ihr Abschluss an der Universität endet noch nicht mit einem „richtigen“ Einstieg in die Berufswelt i. S. v., dass sie selbst lehren oder praktizieren können. Sie beginnen ihre Karrieren zunächst als Rechtsreferendar, Lehramtsreferendar oder auszubildender Arzt im praktischen Jahr. An ihre universitäre Ausbildung schließt erneut eine praktische, betreute Ausbildung an, ehe sie nach weiteren ein bis zwei Jahren als Lehrer, Jurist oder Arzt selbstständig arbeiten können.

Die Daten aus BRIDGE untersuchen deshalb die drei Domänen, da insbesondere diese Berufsgruppen die Fähigkeit kritisch zu denken besitzen sollten (Bäsler, 2021; Kuhn et al., 2020; Zlatkin-Troitschanskaia, Brückner, et al., 2021). Alle drei Berufsgruppen sind in besonderem Maße von der Digitalisierung betroffen (Zlatkin-Troitschanskaia, Brückner, et al., 2021), weisen aber nur unzureichend Fähigkeiten im Umgang mit Online-Informationen auf: Obwohl beispielsweise mit dem Aufkommen der evidenzbasierten Medizin den Suchfähigkeiten und -strategien beim Lokalisieren relevanter Informationen eine immer größere Rolle zugeschrieben wurde (Abdekhoda et al., 2016), verfügen Medizinstudierende noch nicht über die notwendigen Fähigkeiten für ihr späteres Berufsleben (Bazrafkan et al., 2017). Jedoch auch nicht nur für sich selbst, sondern ebenfalls für ihre Bezugsgruppen (Patienten, Klienten, Schüler) ist es relevant, dass diese Berufsgruppen ausreichend kritisch denken können (Zlatkin-Troitschanskaia, Brückner, et al., 2021). Denn wenn Patienten, Klienten oder Schüler selbst nach Informationen recherchieren, können sie diese aufgrund ihres laienhaften Wissens möglicherweise nicht richtig einschätzen (Zlatkin-Troitschanskaia, Brückner, et al., 2021). Wenn angehende Ärzte, Juristen oder Lehrkräfte damit konfrontiert werden, müssen sie die gefundenen Informationen einschätzen, mit dem scheinbar erworbenen Internetwissen der Laien umgehen und ihnen kompetent antworten können (Zlatkin-Troitschanskaia, Brückner, et al., 2021).

Obgleich im Projekt BRIDGE aus den oben genannten wichtigen Gründen die drei unterschiedlichen Domänen untersucht wurden, fokussiert sich vorliegende Arbeit jedoch ausschließlich auf personenübergreifende Merkmale der Berufsanfänger, ohne explizit die Domänen zu untersuchen.

Aus der soeben und im vorherigen Unterkapitel dargestellten Synthese eingängiger aktueller Statistiken und neuster Erkenntnisse aus der Literatur soll mithilfe dieser Arbeit folgende Forschungsfrage beantwortet werden:

In welchem Maße beeinflussen kognitive Personenmerkmale und Webseitenmerkmale den Such- und Evaluationsprozess von Berufsanfängern bei der Lösung eines allgemeinen Online-Informationsproblem?

1.2. Aufbau der Arbeit

Um die zuvor genannte Forschungsfrage zu beantworten und die gesetzten Ziele der Arbeit zu erreichen, werden Begriffe definiert (2.1) und eine ausgiebige Betrachtung der theoretischen Grundlagen geschaffen (2.2, 2.3), ehe diese Grundlagen im weiteren Verlauf in den Forschungskontext transferiert werden (2.4). Darauf folgt die Herleitung der abhängigen Variablen (im Folgenden abgekürzt mit AV bzw. AVs) (2.5) sowie die Beschreibung der Webseiten- (2.6.1) und Personenmerkmale (2.6.3) als unabhängige Variablen (im Folgenden abgekürzt mit UV bzw. UVs). Kapitel 2 endet mit der Herausarbeitung der Forschungslücke und der sich

daraus ergebenden Relevanz zur Untersuchung dieser Fragestellung (2.7). In Kapitel 3 werden die Hypothesen (im Folgenden abgekürzt mit H) unterteilt nach Prozessphase hergeleitet. Innerhalb des vierten Kapitels wird zunächst das Projekt BRIDGE (4.1) vorgestellt, in welchem die Datenerhebung und Forschungsfrage eingebettet sind. Wie die Daten aufbereitet und optimiert werden, wird in 4.2 beschrieben. Kapitel 4.3 deckt die eigentliche Auswertung der zuvor herausgearbeiteten Variablen in den Modellen ab, was mit den Softwares SPSS (Version 27) und JASP (Version 0.17.3 vom 17.03.2023) durchgeführt wird. Das Ergebniskapitel beginnt mit der Beschreibung der deskriptiven Statistiken und Häufigkeiten und der damit verbundenen Beurteilung der Variablungüte (5.1). Es thematisiert weiterhin die korrelativen Ergebnisse (5.2), ehe in 5.3 die Regressionsberechnungen je nach Phase und Modell vorgestellt werden. In Kapitel 6 werden die Ergebnisse modellübergreifend zusammengefasst und diskutiert. Das Kapitel 7 thematisiert die mit der Arbeit verbundenen Limitationen. In Kapitel 8 werden die theoretischen und praktischen Implikationen aus den Erkenntnissen dieser Forschung abgeleitet, ehe das letzte Kapitel 9 ein schlussfolgerndes Fazit präsentiert.

2. Theoretischer Rahmen des Forschungsvorhabens

Das nachfolgende Kapitel führt zunächst in die Begrifflichkeiten der Glaubwürdigkeit sowie weiteren abzugrenzenden Konstrukten ein, um eine Verständnisbasis des Begriffes sowie der Abgrenzungen zu schaffen. Daran anknüpfend werden theoretische Grundlagen dargestellt, die sich thematisch in die Bereiche allgemeine Informationsverarbeitung und Selektion niederschlagen, um die Verarbeitungsprozesse auf individueller Ebene zu veranschaulichen. Anschließend wird mit dem Heuristisch-Systematischen Modell (Original: Heuristic Systematic Model; im Folgenden abgekürzt mit HSM) ein klassisches Informationsverarbeitungsmodell angeführt, welches grundlegend für die Forschung und das Verständnis der kognitiven Prozesse wichtig ist und zudem für die Hypothesenherleitung verwendet wird. Für den Übergang zur digitalen Informationswelt und für die Verknüpfung der beiden Einflussfaktoren – Personen und Webseiten – werden kurz das 3-S-Modell des Vertrauens (Original: 3 S Model of Trust), sowie das internetbezogene Informationsproblem-Löse-Modell (Original: Model of Information Problem Solving regarding Internet usage, im Folgenden abgekürzt mit IPS-I-Modell) mit dem anknüpfenden Konstrukt des kritischen Online-Denkens (COR) dargestellt. Diese geschaffenen Grundlagen werden genutzt, um das Forschungsproblem und den Kontext zu rahmen. Die überblicksartige Darstellung aller beteiligten theoretischen Grundlagen ermöglicht eine prozessuale Abfolge, aus der in Kapitel 2.5 die abhängigen Variablen als Prozess- und Performanzindikatoren abgeleitet werden. Anschließend werden die Personen- und Webseitenmerkmale als erklärende Variablen thematisiert. Zuletzt werden die Erkenntnisse aus den aktuellen Forschungsständen zu den abhängigen und unabhängigen Variablen zusammengetragen, um daraus das Forschungsdesiderat und die Hypothesen abzuleiten.

2.1. Begriffliche Grundlagen: Glaubwürdigkeitsdefinition, innere Differenzierung und abzugrenzende Konstrukte

Es ist essenziell im Zusammenhang mit der rapiden Verbreitung von (Fehl-)Informationen und ihrer unterschiedlichen Qualitäten insbesondere Glaubwürdigkeit als ein Qualitätsmerkmal zu verstehen, und im Sinne einer ausgeprägten Informationskompetenz (Original: information literacy⁴) Informationen hinsichtlich der Glaubwürdigkeit beurteilen zu können (Hobbs, 2006). Individuen stehen jedoch vor der Herausforderung, Glaubwürdigkeit zu definieren und glaubwürdige Informationen zu erkennen, denn es gibt keine klare und einheitliche Definition hiervon: „The literature on credibility is plentiful, contradictory, and confused⁵“ (Self & Roberts, 2019, S. 435). Glaubwürdigkeit ist eine subjektiv von einem Individuum wahrgenommene Qualität von einer Person oder einem Gegenstand (Fogg & Tseng, 1999). Nichts ist per se glaubwürdig, sondern ein Individuum, das etwas oder jemanden wahrnimmt, misst dem Wahrgenommenen einen gewissen Wert an (Un-)Glaubwürdigkeit bei (Wirth, 1999). Wenn in vorliegender Forschungsarbeit von Glaubwürdigkeit oder der Beurteilung dieser die Rede ist, wird damit die anzunehmende, subjektiv wahrgenommene Glaubwürdigkeit gemeint und nicht die tatsächliche, objektive. Es geht nicht darum, ob oder dass ein Merkmal auf einer Webseite tatsächlich glaubwürdig ist. Es gibt vielmehr darum, dass es als glaubwürdig wahrgenommen werden könnte, weil bisherige Forschungen gezeigt haben, dass gewissen Merkmalen aus unterschiedlichen Gründen (Heuristiken, s. Kapitel 2.2.2.1) subjektiv mehr Glaubwürdigkeit zugeschrieben wird als anderen. Zudem ist anzumerken, dass die vorliegende Arbeit zur Vereinfachung die Unterscheidung zwischen glaubwürdig und unglaubwürdig vornimmt. Dennoch ist Glaubwürdigkeit als ein Kontinuum zu verstehen. Das bedeutet, dass es auch zwischen Glaubwürdigkeit und Unglaubwürdigkeit weitere Abstufungen geben kann. Da Glaubwürdigkeit an sich schon subjektiv (Fogg & Tseng, 1999) und schwer zu definieren (Self & Roberts, 2019) ist, sind Aussagen über entsprechende Zwischenstufen ebenfalls subjektiv und schwierig. Die Arbeit geht daher von einer Unterscheidung zwischen Glaubwürdigkeit und Unglaubwürdigkeit aus, ohne dabei explizit auf mögliche Zwischenstufen hinzuweisen.

Die folgenden Ausführungen zeigen in Anlehnung an Fogg (2003) auf, welche zwei notwendigen Bedingungen laut seiner Prominenz-Interpretations-Theorie für die Glaubwürdigkeitsbeurteilung essenziell sind: Zum einen muss das Merkmal mit einer gewissen (hohen) Wahrscheinlichkeit wahrgenommen werden, sprich prominent sein (S. 722 f.). Zum anderen muss die Person das wahrgenommene Merkmal entsprechend interpretieren und beurteilen (S.

⁴ Je nach Kontext kann statt Informationskompetenz auch die Medienkompetenz [für Definition, s. Aufderheide (1997, S. 79)] und/oder Digitalkompetenz [für Definition, s. Peng und Yu (2022, S. 2)] begrifflich verwendet werden, denn wie ein jüngster Literaturreview zeigt, sind die Konstrukte nicht trennscharf voneinander abzugrenzen (Wuyckens et al., 2022, S. 177).

⁵ Übersetzung ins Deutsche: "Die Literatur zur Glaubwürdigkeit ist zahlreich, widersprüchlich und verworren." (Self & Roberts, 2019, S. 435).

723). Nur wenn beide Bedingungen laut der Theorie erfüllt sind, führt dies zur Beurteilung der Glaubwürdigkeit (S. 723).

Vorliegende Forschungsarbeit wird allerdings diese beiden Bedingungen lediglich theoretisch annehmen bzw. voraussetzen, ohne sie direkt im Untersuchungsdesign zu integrieren oder zu messen. Grund dafür ist, dass der Fokus der Arbeit auf den Webseiten- und Personenmerkmalen und die Unterschiede in den Performanz- und Prozessindikatoren liegt. Eine Analyse der detaillierten Prozess- und Kognitionsschritte würde zu weit führen. Diese Einschränkungen bzw. Annahmen werden in den später folgenden Limitationen entsprechend kritisch gewürdigt.

Folgender Prozess verdeutlicht den Zusammenhang der subjektiven Glaubwürdigkeitsbeurteilung für den Online-Kontext (s. auch Abbildung 1): Webseiten stellen Hinweise bzw. Merkmale für die Glaubwürdigkeit oder Unglaubwürdigkeit bereit. Die Merkmale, die als subjektive Glaubwürdigkeitsbeurteilung heranziehbar sind, basieren auf bisherigen Forschungen. Es wird im Sinne der Prominenz-Interpretations-Theorie (Fogg, 2003) angenommen, dass Individuen diese Merkmale wahrnehmen und sie als glaubwürdig oder unglaubwürdig interpretieren. Es wird weiterhin angenommen, dass die Probanden auf Basis ihrer Einschätzung ein bestimmtes Verhalten im Such-, Selektions- und Evaluationsprozess aufweisen, welches anschließend gemessen wird. Eine Webseite stellt folglich bestimmte Merkmale bereit, die zu einer subjektiven positiven oder negativen Glaubwürdigkeitsbeurteilung beitragen (können). Auf der Output-Seite gibt es ein bestimmtes Verhalten, gemessen als abhängige Variablen. Folgende Forschung möchte mithilfe von Zusammenhangsanalysen herausfinden, ob die unterschiedlichen Merkmale auf einer Webseite, die subjektiv als (un-)glaubwürdig empfunden werden könnten, sowie Personenmerkmale Unterschiede im Such-, Selektions- und Evaluationsverhalten erklären können. Ob Webseitenmerkmale tatsächlich wahrgenommen und als solches als glaubwürdig erachtet werden, wird hierbei nicht thematisiert werden.

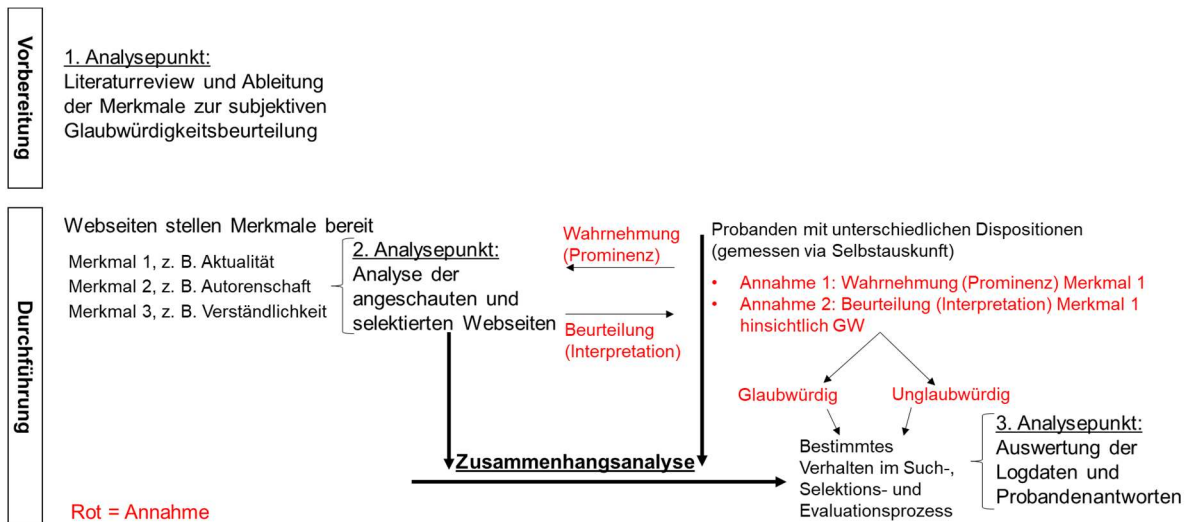


Abbildung 1: Prozessuale Darstellung der subjektiven Glaubwürdigkeitsbeurteilung.

Viele verschiedene Komponenten können bei der Bewertung der Glaubwürdigkeit eine Rolle spielen, z. B. u. a. Dynamik oder wahrgenommene Ähnlichkeit (Nawratil, 1999; Wirth, 1999), dennoch einigt sich das Gros der Forschenden auf zwei, die maßgeblich zur Glaubwürdigkeitsbeurteilung beitragen: Expertise und Vertrauenswürdigkeit respektive Zuverlässigkeit (Original: Trustworthiness) (Fogg & Tseng, 1999; Stiff & Mongeau, 2003). Hovland und Kollegen (1953, S. 21) beschreiben die Vertrauenswürdigkeit als ein Maß an Vertrauen in die Absicht eines Kommunikators, die von ihm als besonders stichhaltig erachteten Behauptungen zu kommunizieren. Vertrauenswürdigkeit ist weiterhin beschrieben als die wahrgenommene Güte und Moralvorstellung eines Kommunikators (Fogg & Tseng, 1999, S. 80) und, dass man sich auf jemanden oder etwas verlassen kann (Oxford Learner's Dictionaries; o.J. - c).

Expertise definieren Hovland et al. (1953, S. 21) als der Umfang, in dem ein Gesprächspartner als eine Quelle valider bzw. zuverlässiger Behauptungen wahrgenommen wird. Expertise beinhaltet die wahrgenommenen Fähigkeiten und das Wissen der Quelle bzw. des Gesprächspartners (Fogg & Tseng, 1999, S. 80). Für vorliegende Forschung wird daher Glaubwürdigkeit als subjektiv wahrgenommene Qualität von jemandem oder etwas definiert. Diese Definition löst die Notwendigkeit aus, eine weitere Abgrenzung innerhalb des Glaubwürdigkeitsbegriffs vorzunehmen. Das, was wahrgenommen und auf Glaubwürdigkeit beurteilt wird, kann unterschiedlicher Art sein.

Innere Differenzierung. Das, was wahrgenommen und als glaubwürdig beurteilt wird, meint im weiteren Sinne Objekte und Personen; alles, was von einem Beurteiler über die Sinne wahrgenommen und zur Glaubwürdigkeitsbeurteilung herangezogen werden kann. Somit resultieren hieraus auch unterschiedliche Subkategorien der Glaubwürdigkeit, wenn diese auf den digitalen Kontext transferiert werden: Nachrichtenglaubwürdigkeit (Appelman & Sundar, 2016), Autorenglaubwürdigkeit und die Glaubwürdigkeit der Quelle respektive des Mediums

(Flanagin & Metzger, 2007; Metzger, Flanagin, Eyal, et al., 2003; Saleh, 2016; Sundar & Nass, 2001). Wird eine Person mit einer Information im Internet konfrontiert, kann sie zunächst beurteilen, ob das Medium, auf welchem die Informationen präsentiert werden, als glaubwürdig erachtet wird. Im weiteren Verlauf könnte die Person feststellen, ob der Verfasser des Textes, die in den meisten Fällen ungleich des Mediums ist – sofern überhaupt ersichtlich – als glaubwürdig angesehen wird. Zuletzt kann die Information selbst, sprich die Nachricht, auf Glaubwürdigkeit beurteilt werden. Für vorliegende Arbeit werden an der entsprechenden Stelle in *Kapitel 2.6.1* die zu untersuchenden Merkmale, die in Zusammenhang mit der Glaubwürdigkeitsbeurteilung stehen könnten, korrespondierend den drei Subkategorien zugeordnet. Wenn im Folgenden von Glaubwürdigkeit gesprochen wird, wird aus Gründen der Einfachheit jedoch zunächst nicht explizit eine Unterkategorie, sondern der gesamte Begriff gemeint. Werden die expliziten Unterkategorien angesprochen, wie bei der Erklärung der Ergebnisse, wird darauf hingewiesen.

Nachdem der Begriff der Glaubwürdigkeit eingeführt, die Bestandteile erläutert und weiterhin innerlich differenziert wurde, müssen weitere, verwandte Konzepte von Glaubwürdigkeit abgegrenzt werden. Das Vertrauen (Tseng & Fogg, 1999; Wells et al., 2021) und die Persuasion (Wathen & Burkell, 2002) stehen in Zusammenhang mit Glaubwürdigkeit, können aber, wie nachstehend eingängig beschrieben, davon unterschieden werden.

Vertrauen. Eng verwoben mit dem Konzept der Glaubwürdigkeit ist auch das Vertrauen (Hovland et al., 1953; Wells et al., 2021). Vertrauen wird oftmals synonym zu Glaubwürdigkeit verwendet, dennoch gibt es Unterschiede (Lucassen & Schraagen, 2011; Tseng & Fogg, 1999). Daher ist es für die vorliegende Forschung notwendig, die beiden Konzepte voneinander abzugrenzen.

Vertrauen in einem Satz punktgenau zu definieren, ist unmöglich. Es gibt erstens unzählige Definitionen, zweitens sind diese kontextabhängig und zuletzt sind sie je nach Autor geprägt von den Paradigmen seiner Disziplin (Aljazzaf et al., 2010). Es gibt also keine allumfassende universal gültige und auf alle Situationen anwendbare Definition von Vertrauen. Dennoch kann angenommen werden, dass das Konzept immer dann Anwendung findet, wenn mindestens zwei Akteure involviert sind, bei denen Vertrauen entgegenbringendes Verhalten die Folge aus einer Vertrauenszuschreibung werden kann. Dabei spielen u. a. Interdependenzen und Unsicherheiten (McEvily et al., 2002), Zuverlässigkeit und Kontrollverlust (Aljazzaf et al., 2010) sowie Risiko, Verletzbarkeit und positive Erwartungen (Rousseau et al., 1998) eine Rolle. Nachstehende Definition versucht zu erfassen, wie sich einige der genannten Begrifflichkeiten in einem möglichen Vertrauensverhältnis und zur Abgrenzung von Glaubwürdigkeit niederschlagen können: „Vertrauen ist die Bereitschaft des Vertrauensgebers, sich darauf zu

verlassen [Zuverlässigkeit], dass ein Vertrauensträger das tut, was er in einem bestimmten Zusammenhang versprochen hat [positive Erwartungen], und zwar unabhängig von der Möglichkeit, den Vertrauensträger zu überwachen oder zu kontrollieren [Kontrollverlust; Unsicherheit], und auch wenn negative Folgen auftreten können [Risiko]⁶ (Aljazzaf et al., 2010, S. 168).

Wird dies mit dem Wörterbuch-Eintrag zu Glaubwürdigkeit kombiniert „the quality that somebody/something has that makes people believe or trust them⁷“ (Oxford Learner's Dictionaries; o.J. - b) wird dadurch auch die Unterscheidung zu Glaubwürdigkeit ersichtlich. Vertrauen bezieht sich folglich u. a. auf (erhoffte) Zuverlässigkeit, aber auch Unsicherheit, gegenseitige Abhängigkeit, die gleichzeitig mit einer Akzeptanz von Risiko und Verletzlichkeit einhergehen. Glaubwürdigkeit hingegen bezieht sich auf die wahrgenommene Qualität von jemandem oder etwas, das zu Vertrauen und folglich Vertrauen entgegenbringenden Verhalten führen kann, aber nicht zwingend muss (Lucassen & Schraagen, 2011; Rieh & Danielson, 2007; Winkel, 1999).

Dem Verständnis zufolge bedeutet dies, dass eine Beurteilung der Glaubwürdigkeit (respektive der beiden Komponenten Expertise oder Vertrauenswürdigkeit/Zuverlässigkeit) gegenüber einer Person oder eines Objektes (z. B. Information, Medien) erfolgt, was zunächst nur innerhalb von einer Person und ihrem subjektiven Verständnis geschieht. Je nachdem, wie die subjektive Beurteilung der Person ausfällt, resultiert daraus Vertrauen, das sich ebenfalls nur innerhalb der Person und ihrem subjektiven Verständnis befindet. Die Glaubwürdigkeit ist daher eine Vorstufe zu Vertrauen (Lucassen & Schraagen, 2011). Fällt die Beurteilung der Glaubwürdigkeit und des Vertrauens positiv aus, so schlägt sich das in einem Vertrauen entgegenbringenden Verhalten nieder, was somit nicht mehr innerhalb der Person selbst, sondern im Umgang mit dem beurteilten Objekt oder der Person stattfindet. Ein Beispiel soll dies illustrieren:

Eine Studierende möchte im Rahmen einer schriftlichen Hausarbeit Informationen zu Ihrem Thema, zu welchem sie bereits eine Meinung hat, online recherchieren. Sie findet eine Nachrichten-Webseite, die sie aufgrund wahrgenommener Hinweisreize (z. B. faktische Korrektheit des Inhalts, Aktualität, Angabe von Quellen) als glaubwürdig beurteilt. Sie beschließt, der Quelle zu vertrauen und sich darauf zu verlassen, dass die Informationen scheinbar glaubwürdig sind, ohne dass sie es abschließend kontrollieren und nachvollziehen kann, da ihr beispielsweise die Möglichkeit fehlt auf die in der Webseite referierte Studie zuzugreifen. Aus

⁶ Original: "Trust is the willingness of the trustor to rely on a trustee to do what is promised in a given context, irrespective of the ability to monitor or control the trustee, and even though negative consequences may occur" (Aljazzaf et al., 2010, S. 168).

⁷ Übersetzung ins Deutsche: „die Eigenschaft von jemandem/etwas, die dazu führt, dass man ihm/ihr glaubt oder vertraut“ (Oxford Learner's Dictionaries; o.J. - b).

dem Vertrauen resultiert das Verhalten, dass sie die Informationen auf der Webseite durcharbeitet und letztlich für ihre Hausarbeit verwendet.

Nachdem an dieser Stelle die Abgrenzung zwischen Vertrauen und Glaubwürdigkeit erfolgt ist, muss die Persuasion von der Glaubwürdigkeit abgegrenzt werden, weil diese ebenfalls mit der wahrgenommenen Glaubwürdigkeit verbunden ist (Wathen & Burkell, 2002).

Persuasion beschreibt einen Kommunikationsakt, als Folge dessen ein Individuum bewusst bzw. intentional von der Meinung eines anderen Individuums überzeugt wurde (*lateinisch: persuadere = überzeugen, überreden*) (M. D. Miller & Levine, 2019). Die Persuasion hat damit das intendierte Ziel eine kognitive, affektive oder verhaltenstechnische Modifikation einer anderen Person zu erreichen (M. D. Miller & Levine, 2019). Die Beurteilung der Glaubwürdigkeit eines Objektes oder einer Person hingegen kann dazu beitragen, ob die beurteilende Person überzeugt wurde und eine entsprechende Verhaltensänderung erfolgen kann (M. D. Miller & Levine, 2019). Die Glaubwürdigkeitsbeurteilung kann daher als vorgelagerter Schritt in den Persuasionsakt eingeordnet werden (Wathen & Burkell, 2002). Das praxisnahe Beispiel wird an dieser Stelle nochmal aufgegriffen, um den Unterschied zu verdeutlichen:

Die Studierende hat beschlossen, den Informationen auf der Webseite zu vertrauen und diese somit zu verwenden. Sie hat bereits eine Meinung zu ihrem Thema, lässt sich aber durch die genaue Bearbeitung der Informationen auf der Webseite und die für sie als gut dargestellten Argumente überzeugen, dass ihre Meinung nicht korrekt ist. Sie adaptiert damit die Meinung des Journalisten der Nachrichten-Webseite. Möglicherweise hat sie auch die Intention des Journalisten hinterfragt, jedoch kann davon nicht immer ausgegangen werden. Dennoch resultiert daraus eine Meinungsänderung, sodass sie ihre Hausarbeit zu dem Thema mit der veränderten Meinung verfasst.

Diese Arbeit beschäftigt sich mit der wahrgenommenen Glaubwürdigkeit und soll aufzeigen, welche Glaubwürdigkeitsbeurteilungen unterschiedlichen Informationen und deren Darstellungsweise im Internet zugeschrieben werden und welche Verhaltensweisen respektive Strategien hieraus folgen können. Da sich Vertrauen und Persuasion beide als nachgelagerte Schritte darstellen, ist es unmöglich, diese völlig aus diesem Prozess auszuschließen, wie in Abbildung 2 dargestellt. Es sei anzumerken, dass das Bild eine Vereinfachung darstellt und nicht auf alle Verbindungen ausführlich eingegangen wird. Bedeutsam zu verstehen ist jedoch, dass die Glaubwürdigkeitsbeurteilung zuerst stattfindet, und erst dann Vertrauen und Persuasion folgen, weshalb die beiden letzteren nur an ausgewählten Stellen Beachtung finden werden. Insbesondere bei der Darstellung des Modells von Lucassen und Schraagen (2011) und der Hypothesenherleitung wird das Vertrauen beispielsweise erneut aufgegriffen. Auch die Persuasion findet sich thematisch im HSM wieder. Dennoch liegt der Fokus der

Forschungsarbeit auf der wahrgenommenen Glaubwürdigkeit, sodass Vertrauen und Persuasion zwar begrifflich erwähnt werden, jedoch zur Beantwortung des Forschungszwecks nur soweit notwendig herangezogen werden.

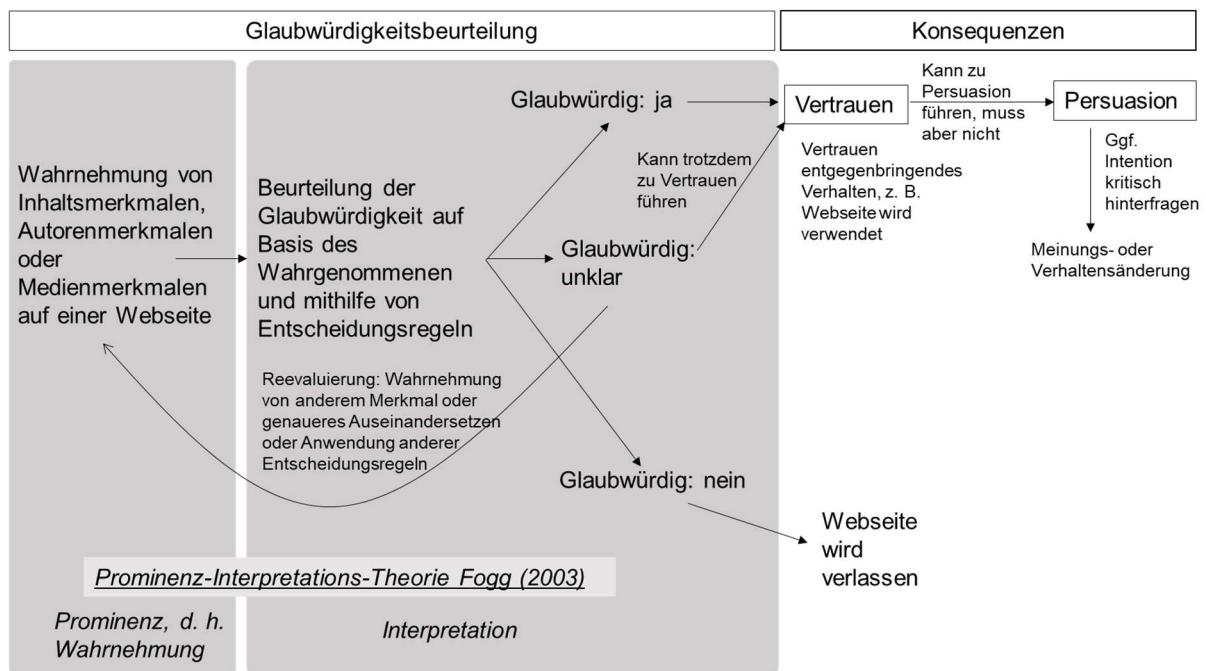


Abbildung 2: Abgrenzung der Glaubwürdigkeit zu Vertrauen und Persuasion.

2.2. Informationsverarbeitungs- und Selektionsmodelle

Das folgende Kapitel thematisiert die allgemeine Informationsverarbeitung und Selektion. Hierzu werden diese beiden Theorien zusammen anhand eines Beispiels dargestellt. Anschließend wird ein Zwei-Prozess-Modell als Vertreter der allgemeinen Informationsverarbeitung präsentiert, bevor hierunter eine beispielhafte Aufführung von Heuristiken und Entscheidungsregeln das Unterkapitel abschließen.

2.2.1. Grundlagen der Selektion und Informationsverarbeitung

Das Gehirn als Teil des Nervensystems mit den für die Informationsübertragung notwendigen Nervenzellen hat eine enorme Verarbeitungskapazität, vergleichbar mit mehreren Milliarden miteinander vernetzten Computern (Anderson, 2013). Es wäre daher naheliegend anzunehmen, dass aufgrund der großen Verarbeitungskapazität sehr viele Informationen verarbeitet und folglich gespeichert werden könnten. Die Realität zeigt allerdings, dass die Informationsspeicherfähigkeit des Gehirns begrenzt ist, was beispielsweise durch das Modell der limitierten Kapazität von Lang (2000) modelliert ist. Der Kognitionspsychologe George A. Miller (1956) postulierte bereits Mitte der 50er Jahre, dass nur in etwa sieben Informationseinheiten verarbeitet und im Kurzzeitgedächtnis gespeichert werden können.

Welche Informationen gespeichert werden (können) und was ursächlich dafür ist, dass es überhaupt unterschiedliche Informationsverarbeitungswege gibt, wird nachfolgend mithilfe der einzelnen Prozesse, die an der Informationsverarbeitung beteiligt sind, erklärt.

Die Informationsverarbeitung ist ein Oberbegriff, der viele verschiedene Subprozesse unter diesem vereint. Für vorliegende Forschung wird zunächst simplifiziert angenommen, dass der Informationsverarbeitungsprozess gemäß der Wortbedeutung her grundsätzlich beschreibt, wie Informationen von einer Person verarbeitet werden. Diesem Schritt der Verarbeitung sind jedoch weitere, detailliertere Prozesse vor- und nachgelagert, nämlich die Informationswahrnehmung, Informationsverarbeitung, Selektion, Speicherung sowie der Informationsabruf im Sinne von Lernen und Erinnern. In Anlehnung an das Selektionsmodell von Donsbach (1991), dem Clickstream-Modell (Wirth & Brecht, 1998) und der Informationsverarbeitung beschrieben in Brünken (2019) und Anderson (2013) werden die Prozesse nachstehend erklärt, sodass eine ausreichende Grundlage zum Verständnis des Forschungsvorhabens geschaffen wird. Zum Zwecke der Vollständigkeit werden alle Begrifflichkeiten erwähnt und in Abbildung 3 kombiniert dargestellt. Für die Forschungsfrage sind jedoch hauptsächlich die Prozesse der Selektion und die anschließende Verarbeitung im Fokus.

Die Informationsverarbeitung und Selektion sind keine untrennbaren Prozesse. Zum einen ist eine kognitive Selektion innerhalb des Informationsverarbeitungsprozesses (Mikroperspektive) an dem Gesamtprozess beteiligt. Zum anderen spielt die Selektion zur Handlung (Makroperspektive) ebenso eine große Rolle.

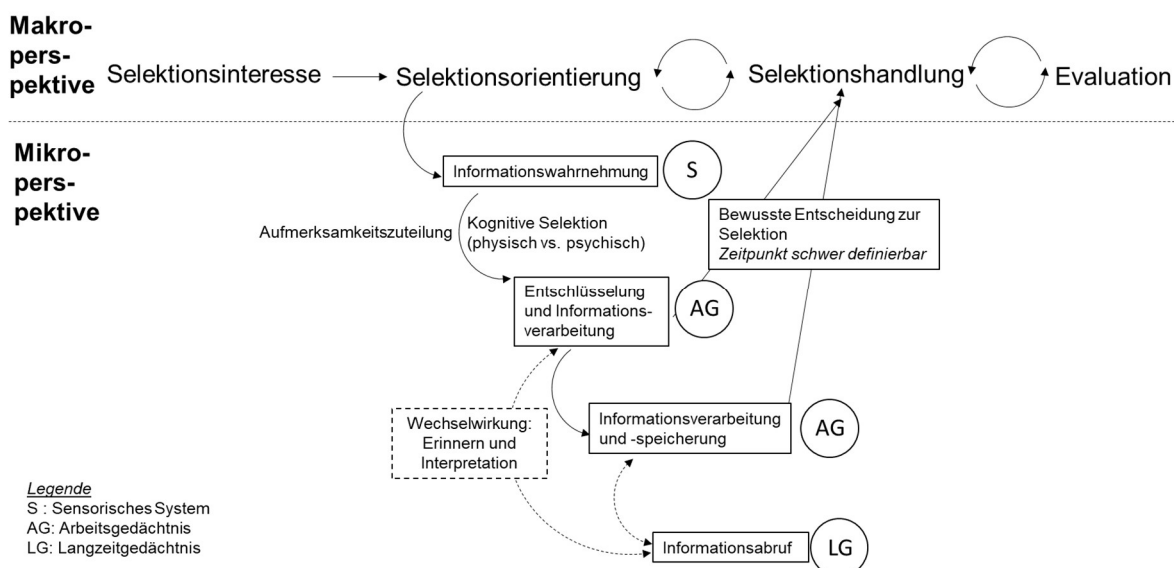


Abbildung 3: Makro- und Mikroperspektive der Informationsverarbeitung und Selektion.

Der Gesamtprozess beginnt auf der Makroperspektive und somit mit der *Selektion* an sich, die am Ende zu einer Selektionshandlung führt.

Die Selektion wird allgemein definiert als ein Prozess „in dem Individuen aus den in ihrer Umwelt potenziell zur Verfügung stehenden Signale mit Bedeutungsgehalt aufgrund von deren physischen und inhaltlichen Merkmalen bestimmte Signale bewusst oder unbewusst auswählen oder vermeiden“ (Donsbach, 1991, S. 28). Diese Definition beschreibt die Selektionshandlung auf der Makroperspektive.

Der Selektionsprozess beinhaltet die Elemente Selektionsinteresse, Selektionsorientierung, Selektions- bzw. Navigationshandlung und Selektionsevaluation (Wirth & Brecht, 1998). Dieser Definition und Phasen zufolge haben Individuen zunächst eine Ausgangslage bzw. ein Ausgangsproblem (Wirth & Brecht, 1998), zu dessen Lösung es einer Informationsgrundlage bedarf. Sie haben also ein *Selektionsinteresse* und verfolgen ein Ziel, wie z. B. die Lösung einer Aufgabenstellung (Wirth & Brecht, 1998). Um dieses Ziel zu erreichen, tätigen sie beispielsweise eine erste Informationssuche, indem sie Suchaufträge über eine Suchmaschine im Internet verfassen.

Das führt zur nächsten Phase, der *Selektionsorientierung*. In dieser orientieren sich die Nutzer an wahrnehmbaren Reizen aus den für sie inhaltlich oder gestalterisch auffälligsten Merkmalen ihrer Umwelt (Wirth & Brecht, 1998), wie z. B. der Schriftgröße oder -art als visueller Reiz des Merkmals „Informationstext“. Mit der Wahrnehmung von Informationen beginnt der eigentliche Informationsverarbeitungsprozess, weshalb folgend nun die Informationsverarbeitungsschritte auf der Mikroprozess-Ebene genauer betrachtet werden, ehe das Selektionsmodell mit den letzten beiden Phasen (Selektionshandlung und -evaluation) fortgeführt wird.

Informationswahrnehmung. Wie erwähnt, ist unser Nervensystem hauptsächlich für die Informationsverarbeitung zuständig. Weitere Beteiligte im menschlichen Nervensystem sind Nervenzellen, die sich nicht nur im Gehirn, sondern auch in unseren Sinnesorganen befinden (Anderson, 2013). So werden Reize aus unserer Umwelt visuell, auditiv oder olfaktorisch mithilfe der entsprechenden sensorischen Systeme wahrgenommen (Anderson, 2013) und im jeweiligen sensorischen Gedächtnis zwischengespeichert (Brünken et al., 2019), wie beispielsweise beim Anschauen eines Textes von Suchergebnissen in einer Online-Recherche mithilfe unseres Auges.

Auf dem Weg zur nächsten Phase, der Informationsverarbeitung, spielen die *Zuteilung von Aufmerksamkeit* und eine erste *kognitive Selektion* eine entscheidende Rolle. Denn ist ein Mensch parallel mehreren Reizen gleichzeitig ausgesetzt, stoßen die sensorischen Systeme an ihre Grenzen (Anderson, 2013; Lang, 2000) und es muss entschieden werden, welcher Reiz priorisiert Aufmerksamkeit erhält. „Die Steuerung der Aufmerksamkeit ist bei fast allen Informationsverarbeitungsprozessen und Handlungsentscheidungen notwendig. Würden wir nicht zwischen wichtigen und unwichtigen Informationen unterscheiden, würden wir alles

unterschiedslos als relevant erachten, dann wäre unsere Informationsverarbeitung sehr schnell hoffnungslos überladen“ (Brünken et al., 2019, S. 96). Welcher Reiz mehr Aufmerksamkeit erhält und den Verarbeitungsprozess fortführt, hängt mit der Art des Reizes zusammen, wobei zwischen physischem und psychischem Reiz differenziert werden kann (Donsbach, 1991). Wird anhand von z. B. der beigemessenen Relevanz (Brünken et al., 2019), der Nützlichkeit oder Interesse (Donsbach, 1991) entschieden, und der Reiz folglich selektiert, ist die Reizart und die anschließende bewusste Selektion psychischer Natur. Sie unterliegt dem Willen des Individuums und ist prädispositionell verankert (Donsbach, 1991). Dem gegenüber steht die Reizart der physischen Reize (Farben, Laute) und dessen Selektion. Physische Reize lösen fest implementierte Mechanismen aus, die aufgrund der Evolution im menschlichen Gehirn verankert, zugleich unbewusst sind, und bei allen Menschen i. d. R. ähnliche Reaktionen hervorrufen (Donsbach, 1991). So wird beispielsweise die Farbe „rot“ die Aufmerksamkeit erregen, da sie aus evolutionärer Sicht mit Gefahr assoziiert wurde (Elliot & Maier, 2007). Auch heute noch ist diese Verlinkung nachweisbar (Pravossoudovitch et al., 2014).

Reize, die der physischen Art angehören, werden von Individuen also unbewusst ohne aktives Zutun mit höherer Aufmerksamkeit bemessen, wohingegen bei Reizen der psychischen Art eine bewusste Aufmerksamkeitssteuerung geschehen kann. Es gibt verschiedene Ansätze aus der Kognitionspsychologie, die Aufmerksamkeitszuteilung zu erklären, z. B. die Filtertheorie von Broadbent (1958), die Dämpfungstheorie von Treisman (1964) oder die Theorie der späten Auswahl von Deutsch und Deutsch (1963). Diese seien an dieser Stelle der Vollständigkeit wegen erwähnt, ohne jedoch im weiteren Verlauf näher darauf einzugehen. Grund für die Aussparung dieser Thematik ist, dass die Aufmerksamkeitszuteilung nicht im Fokus der Arbeit stehen wird.

Wird also anhand der soeben beschriebenen Mechanismen und Kriterien einem Reiz mehr Wert beigemessen als einem anderen, wird ihm mehr Aufmerksamkeit zuteil. Die bewusste oder unbewusste Aufmerksamkeitszuteilung stellt daher eine erste Selektion dar, welche hier als kognitive Selektion von der später aufkommenden (Selektionshandlung) differenziert werden muss. Reize, die es über dieses Stadium der kognitiven Selektion und Aufmerksamkeitszuteilung hinausgeschafft haben, werden anschließenden in das Kurzzeit- bzw. Arbeitsgedächtnis weitergeleitet (Anderson, 2013).

Entschlüsselung und Informationsverarbeitung. Ehe die aus der Umwelt wahrgenommenen Reize im Arbeitsgedächtnis (zu Repräsentationen) verarbeitet werden können, müssen sie dort entschlüsselt werden (Lang, 2000), d. h. in einen kategorialen Code umgewandelt werden (Brünken et al., 2019). Kategorien sind beispielsweise phonologische Einheiten (Laute einer Sprache) oder semantische Konzepte (Bedeutung einer Lauteinheit) (Brünken et al., 2019).

Jeder Reiz wird unterschiedlichen Kategorien zugeordnet, wobei Inhalte des Langzeitgedächtnisses aktiviert bzw. erinnert werden, sodass der zu entschlüsselnde Reiz mit bestehenden Kategorien abgeglichen werden kann (Brünken et al., 2019). Da diese Inhalte aus dem Langzeitgedächtnis abgerufen werden müssen, spielt das Vorwissen des Individuums bei diesem Prozess eine entscheidende Rolle: die aufgenommenen Daten werden erst dann zu einer Information mit Bedeutungsgehalt, wenn sie auf Basis des bereits gespeicherten Vorwissens aus dem Langzeitgedächtnis interpretiert werden (Renkl, 2020). Es herrscht also eine stetige Wechselwirkung zwischen dem Arbeits- und Langzeitgedächtnis. Resultierend daraus entstehen für den wahrgenommenen Reiz mentale Repräsentationen (Lang, 2000), welche individuell basierend auf Erfahrungen aus dem Langzeitgedächtnis (Brünken et al., 2019) erschaffen werden. Die mentalen Repräsentationen bilden folglich niemals eine exakte Darstellung der ursprünglich wahrgenommenen Informationen ab (Lang, 2000), weil sie vom (Vor-) Wissen, der Umgebung oder den Zielen (Lang, 2000) des Individuums abhängen und lediglich eine Interpretation sind.

Informationsspeicherung und -verarbeitung: Ist eine Encodierung erfolgt und sind daraus eine oder mehrere mentale Repräsentationen der ursprünglichen Informationen entstanden, sind sie so lange Bestandteil des Arbeitsgedächtnisses, wie sie aktiviert sind (Brünken et al., 2019). Es findet daher neben der Verarbeitung auch eine erste Speicherung statt (Brünken et al., 2019). Hierbei gilt, dass wenn eine Person über die gerade erworbene Information nachdenkt, zunehmend Verlinkungen bzw. Assoziationen geschaffen werden, die diese Information mit älteren Informationen aus dem Langzeitgedächtnis verbindet (Lang, 2000). Je mehr Verlinkungen zwischen den neuen und alten Informationen bestehen, desto besser wird die neue Information gespeichert (Lang, 2000) und gelernt (Renkl, 2020). Es findet somit erneut eine Wechselwirkung zwischen dem Arbeits- und Langzeitgedächtnis statt. Jedoch obliegt auch das Arbeitsgedächtnis einer restringierten Kapazität (Brünken et al., 2019; Renkl, 2020), weshalb die Aufmerksamkeit auf bestimmte Informationen gerichtet werden muss, sodass diese ausgewählten Informationen in das Langzeitgedächtnis überführt werden können.

Es ist an dieser Stelle zu erwähnen, dass sich Personen hinsichtlich ihrer Arbeitsgedächtniskapazität unterscheiden können und die Unterschiede mit kognitiven Leistungen, wie z. B. schlussfolgerndem Denken, in Zusammenhang stehen (Brünken et al., 2019).

Wird nun wieder die Makroperspektive und das Selektionsmodell betrachtet, folgt anschließend die *Selektionshandlung*, was anhand des Beispiels verdeutlicht werden soll: Bei einer ersten (kognitiven) Selektion werden einigen Reizen mehr Aufmerksamkeit zuteil als anderen, weil sie von höherer persönlicher Relevanz sind. Beispielsweise wird ein erster Satz eines Textes einer Webseite gelesen, weil das Wort „Gesundheit“ heraussticht und der Suchende

sich über Gesundheitsaspekte von E-Bikes informieren möchte. Die Reize, also Informationen im Text, werden durch Abgleich von Inhalten aus dem Langzeitgedächtnis, z. B. Vorwissen über E-Bikes und ihre Funktionsweise, entschlüsselt und in mentale Repräsentationen umgewandelt, d. h. interpretiert. Die bewusste Entscheidung zur Selektion der Reize findet dann mit der weiteren Informationsverarbeitung statt, wobei der exakte Zeitpunkt schwierig zu bestimmen ist. Denn die Prozesse agieren stetig miteinander und verlaufen nicht sequenziell (Trepte et al., 2021). Hat der Abgleich mit dem Langzeitgedächtnis ergeben, dass die Informationen aus dem Text z. B. gut an das Vorwissen anknüpfen, und auch neue Informationen enthalten, wird der Text wahrscheinlich weitergelesen und eindringlicher bearbeitet. Es folgt die Selektionshandlung.

Die *Selektionshandlung* ist das Ergebnis der bewussten Selektionsentscheidung und äußert sich in einer beobachtbaren Handlung, wie z. B. das Anklicken eines Links (Wirth & Brecht, 1998) oder, wie hier im Beispiel, das Auswählen des Textes.

Schließlich endet der Selektionsprozess mit der *Evaluation*, d. h. es wird ein Abgleich mit den Erwartungen getätigt (Wirth & Brecht, 1998). Dies könnte beispielsweise der Abgleich der gefundenen Informationen auf der angeklickten Webseite zur Erfüllung der Aufgabenstellung mit den entsprechenden Erwartungen zur Erfüllung der Aufgabenstellung sein.

Die Prozesse der Selektion und Verarbeitung können daher auch nicht-linear ablaufen, wenn beispielsweise die suchende Person erkennt, dass es eine Erwartungsdiskrepanz gibt. Dies könnte zum Anlass genommen werden, die Phase der Selektionsorientierung erneut durchzuführen, um möglicherweise neue Informationen heranziehen zu können. Die beiden Prozesse stehen in einer iterativ-wechselseitigen Beziehung und können in unterschiedlichen Schleifen oder Längen erfolgen.

Zuletzt wird die Perspektive des Wissenserwerbs i. S. v. Informationsabruf dargestellt, wenn die suchende Person beispielsweise die für die Aufgabenstellung gefundenen Informationen beabsichtigt zu lernen.

Informationsabruf. Sind Informationen erfolgreich aus dem Arbeitsgedächtnis in das Langzeitgedächtnis überführt worden, sind sie für zukünftige Situationen und neue Informationen abrufbar. Die Abrufbarkeit variiert, je nachdem wie viele Verlinkungen mit den Informationen im Vorfeld eingegangen wurde (Lang, 2000). Tritt eine neue Information auf, so werden mentale Repräsentationen und damit verknüpfte Elemente aktiviert, die in den aktuell laufenden Informationsverarbeitungsprozess integriert werden (Brünken et al., 2019). Das eigentliche Abrufen von diesen Informationen stellt daher nichts anderes dar als das Abrufen von Wissen, welches zuvor gelernt wurde. Lernen bedeutet wiederum „(...) organisierte und abrufbare Repräsentationen im Langzeitgedächtnis zu speichern“ (Brünken et al., 2019, S. 105) oder

anders ausgedrückt: „Lernen ist Andocken neuer Information an das Vorwissen“ (Renkl, 2020, S. 9).

Wie intensiv und oft die einzelnen Schritte der Informationsverarbeitung getätigt werden, hängt von vielen Faktoren an. Zum einen spielt die kognitive Kapazität des Individuums eine Rolle (Brünken et al., 2019), das Vorwissen (Renkl, 2020) weiterhin jedoch auch die Anzahl, Komplexität (Eppler & Mengis, 2004; Schneider, 1987) sowie Präsentationsart (insbesondere verbal vs. visuell) (Anderson, 2013) der zu verarbeitenden Informationen.

Es zeigt sich, dass beide Perspektiven, die des Informationsverarbeiters und des Informationsanbieters, betrachtet werden müssen, um den Prozess gänzlich verstehen zu können. Der Notwendigkeit zufolge zeigt sich die Rechtfertigung für die hier zu untersuchende Betrachtung eines ganzheitlichen Modells, in der die Probanden als Informationsverarbeiter und die Webseiten als Informationsanbieter zusammenwirken. Im nächsten Unterkapitel werden auf der soeben dargestellten Theorie weitere Grundlagen aufgebaut, die zum Verständnis der theoretischen Überlegungen und anschließenden Untersuchung beitragen sollen.

2.2.2. Das Heuristisch-Systematische Modell als Vertreter der Zwei-Prozess-Theorien der klassischen Informationsverarbeitungsmodelle

Bis zu diesem Zeitpunkt wurde dargestellt, wie Informationen im menschlichen Gehirn verarbeitet werden und dass die Schritte der Verarbeitung individuell unterschiedlich sein können. Petty und Cacioppo (1986) oder Chen und Chaiken (1999) präzisieren nochmals die Verarbeitung hinsichtlich ihrer Schritte und reihen sich, neben anderen (für einen Review siehe Chaiken & Trope, 1999), als Vertreter der so genannten Zwei-Prozess-Modelle ein (Metzger, 2007). Die Zwei-Prozess-Modelle postulieren eine Zweiteilung des Weges zur Informationsverarbeitung (Gawronski & Creighton, 2013). Konkrete Beispiele hierfür sind die Unterteilung von System 1 und 2 (Kahneman, 2012), das Elaboration-Likelihood-Modell (ELM) von Petty und Cacioppo (1986) mit der peripheren und zentralen Route oder das hier näher ausgeführte Heuristisch-Systematische Modell (Chaiken, 1980; S. Chen & Chaiken, 1999). Alle genannten Modelle ähneln sich in ihrer grundsätzlichen Zweier-Einteilung (z. B. System 1 = peripher = heuristisch), jedoch konzentriert sich vorliegende Forschung auf das HSM, was nach der Gegenüberstellung mit und Abgrenzung vom ELM begründet werden wird.

Chen und Chaiken (1999) postulieren, dass es für die Informationsverarbeitung zwei Routen gibt, die die Individuen einschlagen können. Bei der *systematischen* Route wird eine analytische Betrachtung der Informationen erfolgen (S. Chen & Chaiken, 1999) und Menschen werden einen hohen kognitiven Beitrag aufwenden, um die Informationen zu beurteilen (K. Z. K. Zhang et al., 2014). Die auf diese Weise getätigten Beurteilungen können laut ihrer Annahme auf den tatsächlichen Inhalt der Informationen zurückzuführen sein (S. Chen & Chaiken, 1999).

Weiterhin gilt in Anlehnung an Chen und Chaiken (1999): Bei der *heuristischen* Route verlassen sich Menschen auf einfache Daumen- bzw. Beurteilungsregeln (Heuristiken) oder erkennbare Hinweisreize, um schnell zu einem Urteil oder einer Entscheidungsfindung zu gelangen (S. 74). Diese Beurteilungsregeln werden laut der Annahme der Forschenden zu einem früheren Zeitpunkt erlernt und als Erinnerung gespeichert (S. 74). Durch die Speicherung können sie in neuen Situationen abgerufen und angewandt werden (S. 74). Drei Bedingungen sind für die heuristische Informationsverarbeitung laut den Autoren (S. Chen & Chaiken, 1999, S. 82 f.) notwendig: Erstens, dass die Beurteilungsregeln überhaupt gelernt und gespeichert wurden, also dass sie *verfügbar*, zweitens aus der Erinnerung *abrufbar* sind und zuletzt müssen die Regeln für die entsprechende Entscheidungssituation auch *anwendbar* sein. Die auf dieser Route getätigten Urteile berücksichtigen weniger den Inhalt, sondern vielmehr die leicht zu verarbeitenden Hinweise (S. 74). Beispiele für die Heuristiken im Online-Kontext werden aufgrund der Ausführlichkeit als gesondertes Unterkapitel in 2.2.2.1 erläutert.

Beide Routen haben als eine erste Grundannahme gemein, dass eine gewisse kognitive Fähigkeit vorhanden sein muss, auch wenn dies bei der heuristischen Informationsverarbeitung verhältnismäßig weniger die Anforderung ist (Chaiken, 2014; S. Chen & Chaiken, 1999). Als zweite Grundannahme formulieren die Autoren, dass Motivation (Deci & Ryan, 1993) existent sein muss (Chaiken, 2014). Die vorherrschende Meinung lautet, dass Menschen mit größerer Motivation (Sillence et al., 2007) und größeren (kognitiven) Fähigkeiten (Flanagin & Metzger, 2000) auch sehr wahrscheinlich Informationen systematischer verarbeiten, beispielsweise größere Aufmerksamkeit auf die Informationsqualität als Hinweismerkmal richten und gründlicher Informationen auswerten (Metzger, 2007; Metzger et al., 2010) als Menschen mit vergleichsweise niedrig ausgeprägter Motivation oder Fähigkeit. Der Annahme zufolge wird auch hier erneut ersichtlich, dass die Rolle der kognitiven Fähigkeit bei der Informationsverarbeitung von wesentlicher Bedeutung sein kann und die Betrachtung der (kognitiven) Personenmerkmale rechtfertigt.

Wie bereits erwähnt, ähneln sich das ELM und HSM. Abgesehen von der auffälligsten Ähnlichkeit, dass beide zwei unterschiedliche Wege der Informationsverarbeitung postulieren, eint die Modelle auch die soeben ausgeführte Annahme, dass Motivationen und Kognitionen Antezedenzen zur Informations- oder Nachrichtenverarbeitung sind (S. Chen & Chaiken, 1999).

Unterschiedlich zum ELM und als ein Rechtfertigungsgrund zur Wahl des HSM kann jedoch aufgeführt werden, dass beim HSM kein exklusives Kontinuum („heuristisch vs. systematisch“) formuliert wird (S. Chen & Chaiken, 1999, S. 81), wohingegen aus der Beschreibung zum ELM implizit angenommen werden kann, dass entweder „nur peripher“ oder „nur elaboriert“ verarbeitet wird (Meijnders et al., 2001, S. 759; Petty & Cacioppo, 1986, S. 187; Stiff & Mongeau,

2003, S. 223). Das bedeutet, dass die Autoren des HSM eine gewisse Freiheit modellieren, die es erlaubt anzunehmen, dass systematische und heuristische Informationsverarbeitung auch parallel oder gleichzeitig stattfinden, und so gemeinsam auftreten können (S. Chen & Chaiken, 1999; Maheswaran & Chaiken, 1991; K. Z. K. Zhang et al., 2014). In Betrachtung des geplanten Forschungsrahmens ist ein Verarbeiten auf diese Weise hier ebenso denkbar: Wenn Probanden online nach Informationen suchen, könnten sie beispielsweise aufgrund der heuristischen Verarbeitung den ersten Link einer Suchergebnisseite auswählen, ehe sie dann den Inhalt der Webseite ausgiebig analysieren und somit systematisch verarbeiten (Wirth et al., 2007, S. 795). Insofern könnte ein Hinweisreiz im angedachten Forschungsrahmen sowohl heuristisch als auch systematisch verarbeitet werden.

Ein zweiter Grund, der für das HSM und gegen das ELM spricht, ist, dass das HSM für einen allgemeineren Kontext anwendbar ist, wohingegen sich das ELM auf persuasive Kommunikation limitiert (S. Chen & Chaiken, 1999, S. 80; Sussman & Zhang, 2008, S. 75). Im entfernten Sinne können die Informationen, die Probanden im hier angedachten Forschungsrahmen sehen oder lesen werden, als „persuasive Kommunikation“ bezeichnet werden. Webseiteninhalte haben das Ziel respektive die Intention, den Leser zu einem Thema, wie der Gesundheitsförderung von E-Bikes, zunächst zu informieren und dadurch möglicherweise vom Inhalt zu überzeugen. Dennoch wird Persuasion als ein Kommunikationsakt gesehen, der überwiegend das Ziel hat, jemanden zu überzeugen und eine Modifikation des Verhaltens hervorzurufen (M. D. Miller & Levine, 2019). Für vorliegende Forschung wird davon ausgegangen, dass der persuasive Charakter in der Intention der Autoren oder Webseitenbetreiber weniger präsent ist, weshalb das HSM mit dem größeren und breiter angelegten thematischen Spielraum dem ELM vorzuziehen ist. Im Einklang mit der Erörterung der beiden Unterschiede wird für vorliegende Forschung somit das HSM verwendet.

2.2.2.1. Ausgewählte Entscheidungsregeln bei der Informationsverarbeitung

Menschen können sich bei der Informationsverarbeitung so genannter Daumenregeln bzw. Heuristiken bedienen. Diese erleichtern die Entscheidungsfindung, weil sie den kognitiven und zeitlichen Aufwand verringern (Metzger et al., 2010, S. 417, 434). Sie können als mentale Abkürzungswege verstanden werden (Metzger et al., 2010, S. 417). Sie ermöglichen, dass Situationen oder Informationen nicht nochmal von Neuem beurteilt werden müssen, sondern dass abgespeicherte Denkmuster, wie Entscheidungsregeln, abzurufen und auf die Situation anzuwenden sind (S. Chen & Chaiken, 1999, S. 74).

Die Beurteilung der Glaubwürdigkeit aufgrund von Heuristiken hat insbesondere in der digitalen Welt eine große Relevanz: Konnten vor einigen Jahren analoge Medien noch relativ einfach auf Glaubwürdigkeit oder vergleichbaren Qualitätskriterien beurteilt werden, weil es eine

überschaubare Anzahl an traditionellen Medien gab, ist dies für Internetnutzer aufgrund der zahlreichen Informationen und Medien schier unmöglich geworden (Sundar, 2008). Das rückt gerade im digitalen Kontext zunehmend die Glaubwürdigkeitsbeurteilung anhand von bereits abgespeicherten mentalen Denkmustern oder Abkürzungen in den Fokus.

Es ist naheliegend anzunehmen, dass die heuristische Verarbeitung keine ausreichende Verarbeitung darstellen und zu einem schlechteren Resultat führen könnte, beispielsweise weil aufgrund der schnellen Entscheidung und dem geringen kognitiven Aufwand möglicherweise leicht Fehler übersehen werden. Die Literatur zur Nachrichtenverarbeitung liefert Argumente für die Annahme und zeigt auf, dass eine oberflächliche und somit heuristische Informationsverarbeitung z. B. zu einer geringeren Risikoeinschätzung (Trumbo, 1999) führt oder das Lernen aus Nachrichten (Eveland, 2002) behindert. Eine weitere Forschung zeigt jedoch, dass bei der Online-Suche die Verwendung von Heuristiken im Vergleich zu einer detaillierten und systematischen Auseinandersetzung keineswegs zu einem schlechteren Ergebnis, i. S. v. Suchperformanz führt (Wirth et al., 2007, S. 795). Insofern ist es sinnvoll Entscheidungsregeln und Heuristiken zu begutachten, welche bei der Beurteilung der Glaubwürdigkeit von Online-Informationen angeführt werden und somit einen wesentlichen Teil zum Lösen eines Informationsproblems darstellen kann. Grundsätzlich kann jeder Hinweisreiz bzw. jedes Merkmal heuristisch verarbeitet werden und folglich eine Heuristik darstellen. Es werden nachstehend jedoch nur eine Auswahl an Heuristiken präsentiert, welche für den Forschungskontext relevant sind und zum Verständnis der Verbindung zwischen Glaubwürdigkeitsbeurteilung und Webseitenmerkmal beitragen. Es sei noch zu erwähnen, dass eine Abgrenzung aller möglicher Heuristiken und Entscheidungsregeln nur schwer ist (Gigerenzer et al., 1999), denn sie lassen sich beispielsweise kombinieren oder bilden die Basis für eine Weiterentwicklung (Metzger et al., 2010).

Bekanntheitsheuristik. Auf Basis der Heuristik beurteilen Individuen Informationen oder andere Individuen als glaubwürdiger, wenn sie (ihnen) bekannter sind (z. B. Hilligoss & Rieh, 2008, S. 1476). Eine vorangegangene Erfahrung ist hierzu eine vorgelagerte Bedingung, dass die Bekanntheitsheuristik greifen kann. Wenn bei einer Entscheidung zwei Alternativen zur Wahl stehen, wird die Alternative bevorzugt bzw. erhält einen höheren Wert für ein bestimmtes Kriterium (z. B. Glaubwürdigkeit), welche bereits bekannt ist und mit der Erfahrungen gemacht wurden (Gigerenzer & Gaissmaier, 2011). Wenn eine Person in der Vergangenheit bereits Erfahrungen mit einem anderen Menschen gemacht hat, so wird dies abgespeichert. In einer neuen Situation wird diese Erfahrung bzw. Erinnerung aktiviert und dies wird bei der Entscheidung dazu führen, dass die bekannte Alternative als glaubwürdiger beurteilt wird. Dieses Muster kann dann weiterhin auf ähnliche Situationen analog angewandt werden, obwohl nur die Vermutung besteht, jemanden zu kennen.

Wiedererkennungsheuristik. Unter die Bekanntheitsheuristik kann weiterhin auch die Wiedererkennungsheuristik eingeordnet werden. Diese besagt, dass etwas oder jemand als qualitativ besser (z. B. Menschen als glaubwürdiger) bewertet wird (Gigerenzer et al., 1999, S. 40 f.; Metzger et al., 2010, S. 426), wenn es wiedererkannt wird im Vergleich zu etwas nicht Wiedererkanntem bzw. Unbekanntem. Anhand dieser Aussage wird deutlich, dass die Bekanntheitsheuristik nicht gänzlich von der Wiedererkennungsheuristik zu trennen ist. Denn wenn etwas wiedererkannt wird, muss in der Vergangenheit schon einmal zwingend Erfahrung damit gemacht worden und somit bekannt sein.

Reputationsheuristik. Eine weitere sich überschneidende Heuristik ist die Reputationsheuristik, welche dafür verantwortlich ist, dass Menschen, die eine höhere Reputation genießen, als glaubwürdiger eingeschätzt werden (Hilligoss & Rieh, 2008). Die Definition des Begriffs Reputation sei an dieser Stelle notwendig, um die anschließend folgende Autoritätsheuristik zu verstehen und von dieser differenzieren zu können. Eine Reputation oder ein Ruf beschreibt das Ansehen, das eine Person oder eine Institution „(...)für ihren spezifischen Beitrag zur Realisierung kollektiv geteilter Ziele und Werte in der Öffentlichkeit genießt“ (Eisenegger, 2015, S. 450). Eine positive Reputation kann dadurch entstehen, dass die Erwartungen wichtiger Bezugsgruppen dauerhaft erfüllt werden (Eisenegger, 2015). Abgeleitet aus diesem Verständnis ist eine notwendige Bedingung für das Entstehen einer (guten/schlechten) Reputation somit, dass entsprechende Erfahrungen mit der Person oder beispielsweise einer Institution gemacht wurden. Auch im Bereich der Online-Medien⁸ tätigen Metzger et al. (2010, S. 426) bspw. genau diese Überlegungen und greifen zudem auf, dass Bekanntheit und Wiedererkennung obligatorische Voraussetzungen für das Zutreffen der Reputationsheuristik sind. Denn sie führen an, dass die Reputationsheuristik ihren Ursprung wahrscheinlich in der Bekanntheitsheuristik hat mit dem Ziel, kognitiven Aufwand bei Entscheidungen gering zu halten (Metzger et al., 2010, S. 426).

Auch die *Expertenheuristik* könnte, je nach Kontext, mit der Reputationsheuristik gleichgesetzt werden. Die Expertenheuristik verursacht, dass Menschen, die von z. B. Rezipienten als Experten wahrgenommen werden eine höhere Glaubwürdigkeit zugeschrieben werden als solche, die nicht als Experten wahrgenommene werden (Sundar, 2008). Für die Wahrnehmung als Experte müssen allerdings Hinweise, wie z. B. ein Dokortitel im Fachgebiet der Medizin, wenn es um medizinische Entscheidungen geht, erkennbar sein. Die Zuschreibung der Glaubwürdigkeit resultiert letztlich daraus, dass ein Rezipient zu einem vorherigen Zeitpunkt bereits Erfahrungen mit diesem Experten gemacht hatte (Sundar, 2008). Alle zukünftigen Entscheidungen werden dahingehend erleichtert, in dem Sinne, dass bei Erkennung von einem

⁸ Die Definition lässt sich auch in ähnlicher Weise auf analoge Medien transferieren. Da das Forschungshaben allerdings Medien im Online-Kontext untersucht, wurde sich für die entsprechende Definition entschieden.

Experten-Hinweis diese Heuristik unbewusst aktiviert wird und zu einer positiv ausfallenden Glaubwürdigkeitsbeurteilung führt. Auch konnte dieser Zusammenhang bereits experimentell nachgewiesen werden: Forschende (Meinert & Krämer, 2022) zeigten auf, dass Informationsquellen mit einem Expertenhinweis als glaubwürdiger erachtet wurden im Vergleich zu Informationsquellen ohne solchen Hinweis.

Autoritätsheuristik. Eine weitere Heuristik ist die Autoritätsheuristik. Einer Quelle oder einer Person wird eine hohe Glaubwürdigkeit zugeschrieben, wenn sie offizieller Art ist und dadurch eine Autorität suggeriert (Sundar, 2008, S. 84). Die Autoritätsheuristik ist hinsichtlich ihrer Entstehung von der Reputationsheuristik abzugrenzen, obwohl dies herausfordernd ist, denn die Reputationsheuristik könnte unter die Autoritätsheuristik subsumiert werden (Metzger et al., 2010): Eine Reputation kann sich über eigene Erfahrungen aufbauen, wenn z. B. eine Lehramtsstudierende online immer wieder bei der gleichen Webseite Informationen recherchiert, mithilfe derer sie sich gut auf Klausuren vorbereiten und sie anschließend bestehen kann. Die Erwartung der Studierenden, dass die Webseite für die Klausur hilfreiche Informationen bereitstellt, wird somit dauerhaft erfüllt, weshalb sie bei ihr mit einer guten Reputation abgespeichert ist. Wird die Studierende zu einem späteren Zeitpunkt mit einem anderen Problem konfrontiert, wird sie die positive Erinnerung an die Reputation der Webseite abrufen und diese erneut nutzen. Bei der Autoritätsheuristik hingegen wird eine Person oder Webseite beispielsweise aufgrund ihrer offiziellen Betitelung bzw. Position als autoritär und folglich glaubwürdiger wahrgenommen, ungeachtet dessen, ob bereits eigene Erfahrungen zu dieser Person respektive Quelle vorliegen. Die Lehramtsstudierende aus vorliegendem Beispiel möchte sich auf ihre Klausur vorbereiten und mithilfe einer offiziellen Webseite, beispielsweise dem Bundesministerium für Bildung und Forschung, nach klausurrelevanten Informationen suchen. Obwohl sie keine bisherigen Erfahrungen mit der Quelle gemacht hat, wird sie die Informationen als glaubwürdig erachten, da sie offizieller Art ist. - i. S. v. betitelt durch die aktuelle Bundesregierung. An dieser Stelle könnte allerdings auch *die Expertenheuristik ansetzen*, da das Bundesministerium für Bildung und Forschung von Seiten der Lehramtsstudierenden möglicherweise auch als Experte für das klausurrelevante Thema angesehen werden kann. Anhand der Ausführungen zeigt sich, dass sich die Heuristiken der Bekanntheit, Wiedererkennung und Reputation sowie der Reputation, Autorität und Experten nur schwierig trennen lassen.

Leichtigkeit der Verarbeitung. Die Heuristik der Verarbeitungsleichtigkeit (Original: Processing Fluency) (Schwarz, 2015) verursacht, dass Menschen erstens Informationen besser abspeichern können und zweitens positiver bewerten (Graf et al., 2017), weil sie die Informationen leichter verarbeiten können (Schooler & Hertwig, 2005). Aufgrund einer erleichterten Verarbeitung fallen beispielsweise ästhetische Urteile (Landwehr et al., 2011, 2013), Wahrheitsurteile (Hansen et al., 2008) oder Vertrauensurteile (Reber & Schwarz, 1999) positiver aus. Bezogen

auf die wahrgenommene Glaubwürdigkeitsbeurteilung bedeutet dies, dass Menschen Informationen als vertrauenswürdiger und somit glaubwürdiger erachten, weil sie diese leichter verarbeiten (z. B. durch gleiche Kommunikationssprache, einfache Sprache etc.) und weiterhin, weil sie sie besser abspeichern und zu einem späteren Zeitpunkt wieder erinnern können.

Reaktanz. Bei anderen Autoren als „Heuristik der Überredungsabsicht“ (Metzger et al., 2010, S. 432) oder „Heuristik der Aufdringlichkeit“ (Sundar, 2008, S. 82) bezeichnet, wird die Reaktanz (Brehm, 1966) hier als Prozess verstanden. Sie wird dennoch in diesem Kapitel der Heuristiken angeführt, da sie ebenfalls den Zusammenhang zwischen Merkmalen und Glaubwürdigkeitsbeurteilung beschreiben kann. Der Begriff drückt eine unangenehme, motivationale Erregung aus, die auftritt, wenn Menschen die Bedrohung oder den Verlust ihrer Verhaltensfreiheit wahrnehmen (Steindl et al., 2015). Resultierend aus dieser Erregung empfinden Menschen u. a. Aggression (Brehm, 1966; Steindl et al., 2015) und möchten ihre bedrohte bzw. verlorene Freiheit schützen oder wiedererhalten, in dem sie z. B. die aufgezwungene Option ablehnen (Steindl et al., 2015). Beispielhaft im Online-Kontext anzuwenden ist diese Art der Heuristik bei der Begegnung mit einer bewussten Manipulation durch Werbung. Die Präsenz einer klaren Beeinflussung in Form einer Werbeanzeige führt dazu, dass sich ein Mensch in seiner freien Entscheidung über die Bewertung eingeschränkt fühlt (Metzger & Flanagin, 2013) und so den Schutzmechanismus zur Ablehnung der aufgezwungenen Option aktiviert. Für eine potenzielle Glaubwürdigkeitsbeurteilung bedeutet dies beispielsweise, dass eine negative Bewertung (Webseite oder Information ist unglaubwürdig) oder gar Ablehnung (Webseite oder Information wird verworfen) lediglich deshalb erfolgt, weil eine mögliche Beeinflussung bei der Person erkennbar ist.

Darüber hinaus seien an dieser Stelle noch weitere Heuristiken erwähnt, wie z. B. die Befürwortungsheuristik (Metzger et al., 2010) respektive Bandwagon-Effekt (Sundar, 2008), bei dem Probanden eine Information oder Webseite deshalb als glaubwürdig erachten, weil ihnen durch andere (z. B. Empfehlung) signalisiert wird, dass sie glaubwürdig sei. Bei der Konsistenzheuristik (Metzger et al., 2010) wird eine Webseite oder Information als glaubwürdig erachtet, wenn unterschiedliche Webseiten die gleichen und somit konsistenten Informationen wiedergeben, im Sinne einer Bestätigung. Letzte zu nennende ist die Erwartungsverletzung respektive Selbstbestätigung: Die Probanden haben bereits vor Kenntnisnahme der Webseite oder Information eine Erwartung oder eine Meinung, wobei die Webseite als unglaubwürdig erachtet wird, wenn die Erwartung nicht erfüllt bzw. die Meinung nicht bestätigt wird (Metzger et al., 2010).

2.3. Digitale Informationsverarbeitungs- und Selektionsmodelle

Das vorherige Kapitel diene zur Schaffung der Grundlagen für die allgemeine Informationsverarbeitung. Im Folgenden werden zwei Modelle und ein Konstrukt dargestellt, die die Modelle in den Online-Kontext transferieren und im weiteren Verlauf zur Hypothesenherleitung herangezogen werden. Das 3-S-Modell von Vertrauen liefert Anhaltspunkte zur Unterteilung von Webseitenmerkmalen. Das IPS-I-Modell sowie das COR-Konstrukt thematisieren die (idealen) Prozessschritte beim Lösen eines Informationsproblem online.

2.3.1. Einführender Überblick zum 3-S-Modell des Vertrauens

Die Wahl für das Modell wird damit begründet, dass dieses den wie hier ebenfalls benötigten Online-Kontext thematisiert. Weiterhin werden die Hinweise bzw. Merkmale, die zur Beurteilung herangezogen werden können, im 3-S-Modell in drei Kategorien (Semantik, Oberfläche, Quelle) unterteilt. Die Vorüberlegungen der Forscherin zu den Merkmalen sind sich zu den Kategorien des 3-S-Modells inhaltlich ähnlich, was somit ebenfalls für die Wahl des Modells spricht. Zudem sind die drei Kategorien des 3-S-Modells angelehnt an die zwei Verarbeitungswege (heuristisch/systematisch), was für den Zweck dieses Forschungsvorhabens im Hinblick auf die später folgende Hypothesenherleitung dienlich ist.

Vertrauen und Glaubwürdigkeit sind, wie bereits in Kapitel 2.1 beschrieben, eng miteinander verwoben. Laut eigener Differenzierung muss die Glaubwürdigkeitsbeurteilung zuerst erfolgen, ehe ein Vertrauen entgegenbringendes Verhalten in einem zweiten Schritt erfolgen kann. Die zur Beurteilung der Glaubwürdigkeit wahrgenommenen Merkmale (von z. B. Informationen) sind diesen beiden Schritten (Glaubwürdigkeitsbeurteilung und Vertrauensbildung) noch vorgelagert und finden sich in dem Modell von Lucassen und Schraagen (2011) wieder. Es wird angenommen, dass die Beurteilung der Glaubwürdigkeit als Zwischenschritt durchgeführt wird, der in diesem Modell allerdings nicht explizit dargestellt wird, weil die Autoren den Fokus auf Vertrauen legen. Da sich für diese Forschung allerdings Glaubwürdigkeit und Vertrauen nicht explizit trennen lassen, könnten die gleichen Merkmale, die zur Vertrauensbildung beitragen, auch zur Beurteilung der Glaubwürdigkeit angewandt werden. Insofern kann das Modell (Abbildung 4) für die Einteilung der Charakteristiken in die unterschiedlichen Kategorien verwendet und für das vorliegende Forschungsproblem angewandt werden⁹. Diese Ausführung liefert abschließend einen letzten Grund für die Wahl des Modells.

⁹ Der Fokus des Modells liegt auf unterschiedlichen Strategien, die von den Autoren im Zusammenhang mit den Nutzermerkmalen gebracht werden. Diese nähere Betrachtungsweise der Strategien ist für die vorliegende Forschung nicht notwendig, da ausschließlich die Kategorien der Merkmale zur Herleitung der Hypothesen und Bildung der abhängigen Variablen verwendet werden.

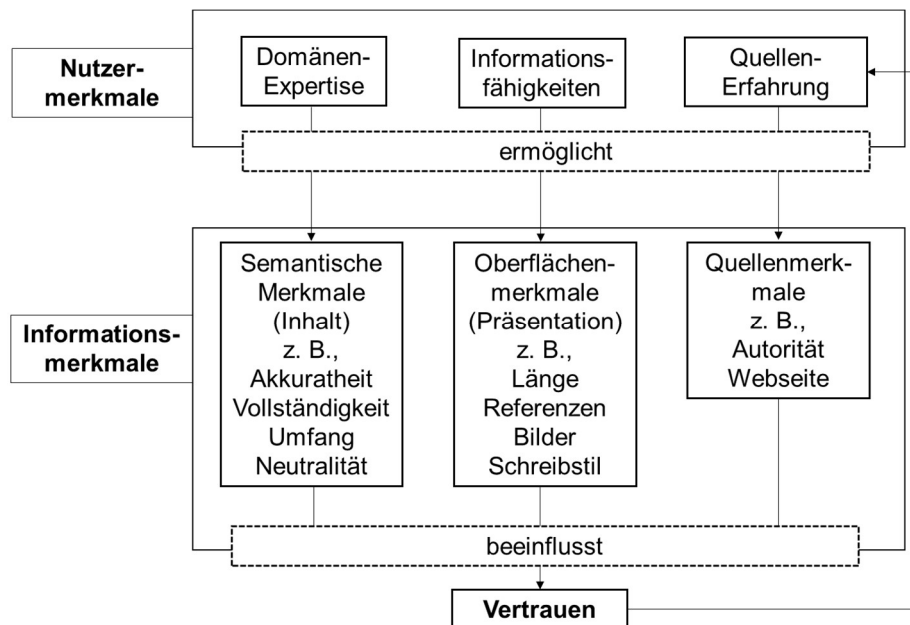


Abbildung 4: 3-S-Modell, eigene Abbildung nach Lucassen et al. (2013).

Die Autoren beschreiben das Zusammenspiel der Informationsmerkmale (semantische Merkmale sowie Oberflächen- und Quellenmerkmale) mit den Nutzermerkmalen (Informationsfähigkeit sowie Domänenexpertise und Quellenerfahrung), welches zur Vertrauensbildung beitragen kann (Lucassen & Schraagen, 2011). Die Expertise oder das Fachwissen in der Domäne ermöglicht dabei das Erkennen und Beurteilen von inhaltlichen bzw. semantischen Merkmalen (wie z. B. faktische Korrektheit, Vollständigkeit oder Neutralität) (Lucassen & Schraagen, 2011, S. 1234 f.). Die Informationsfähigkeit ermöglicht ein Erkennen von Oberflächenmerkmalen, wie Bilder, Schreibstil oder Angabe von Referenzen (Lucassen & Schraagen, 2011, S. 1235 f.). Zuletzt führt vorangegangene Erfahrung mit einer Quelle dazu, dass Quellenmerkmale, wie Autorität oder die Webseite, beurteilt werden können (Lucassen & Schraagen, 2011, S. 1234, 1236).

Die Autoren des 3-S-Modell distanzieren sich davon, die drei Informationsmerkmalscluster exakt den entsprechenden zwei Routen des HSM zuzuordnen. Sie führen lediglich an, dass sie die Bewertung von Oberflächenmerkmale als eindeutig heuristisch, die von Semantik- und Quellen-Merkmale als sowohl heuristisch als auch systematisch betrachten (Lucassen & Schraagen, 2011). Dennoch werden ihre Vorüberlegungen sowie die Einteilung der Merkmale in die drei Kategorien für das vorliegende Forschungshaben diskutiert und an entsprechender Stelle angewandt.

2.3.2. Das IPS-I-Modell

Das "Information Problem Solving regarding Internet Usage"-Modell beschreibt das Lösen eines Informationsproblems durch Nutzung des Internets (Brand-Gruwel et al., 2009, S. 1208)

und kann als Erweiterung gängiger IPS-Modelle (z. B. Eisenberg & Berkowitz, 1990; Kuhlthau, 2004; Stripling & Pitts, 1988) angesehen werden. Bisherige Modelle zur Beschreibung des Problem-Löse-Prozesses, wie z. B. das *Information Search Process*-Modell (ISP) (Kuhlthau, 1991), *Task-Based Information Retrieval Process*-Modell (Vakkari, 2001) oder das *Search Engine Information-Seeking Behavior*-Modell (Sharit et al., 2008) ziehen nicht die Rolle des Internets in Betracht¹⁰ (Kammerer & Brand-Gruwel, 2020, S. 215). Das IPS-I hingegen wird der Rolle des Internets gerecht und erlaubt einen tieferen und detaillierten Einblick in die Prozesse (Brand-Gruwel et al., 2009, S. 1215), was somit die Verwendung des Modells in dieser Forschungsarbeit rechtfertigt.

Die folgenden Ausführungen in Anlehnung an Brand-Gruwel et al. (2009) zeigen auf, welche Komponenten zum IPS-I-Modell gehören: Grundsätzlich besteht das Modell aus drei unterschiedlichen Sets an Fähigkeiten – den aufbauenden Fähigkeiten, den Regulationstätigkeiten sowie den bedingten Fähigkeiten – welche im Folgenden beschrieben werden (Abbildung 5).

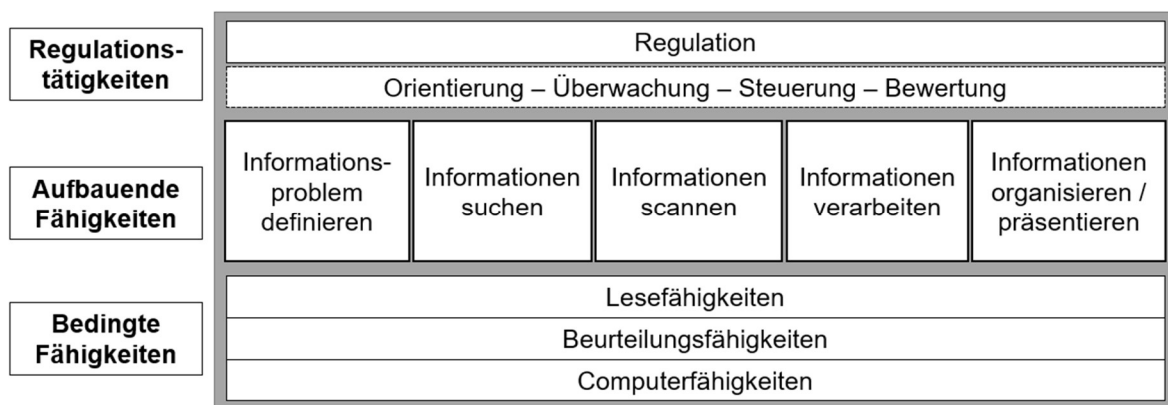


Abbildung 5: Das IPS-I-Modell, eigene Abbildung in Anlehnung an Brand-Gruwel et al. (2005, 2009) sowie Eisenberg und Berkowitz (1990).

Die aufbauenden Fähigkeiten stellen die Kern-Fähigkeiten im Rahmen des gesamten Modells dar. Der Problemlöse-Prozess beginnt zunächst damit, dass das Informationsproblem bestimmt bzw. definiert werden muss (S. 1208). Zur Definition müssen vorhandenes Vorwissen aktiviert sowie Fragen formuliert werden (S. 1208). Im zweiten Schritt (Informationen suchen) wird der eigentliche Suchprozess initiiert, wobei eine Suchstrategie gewählt werden muss (S. 1208). Die gängigste Strategie ist die Eingabe eines Suchbegriffes in eine Suchmaschine, wobei auch das Eintippen einer gezielten URL als Strategie angeführt wird (S. 1208 f.). Der ersten Strategie zufolge gehört hierzu auch die Fähigkeit, den Suchbegriff zu formulieren bzw. spezifizieren und die aufgelisteten Suchergebnisse zu bewerten (S. 1209). Der dritte Schritt (Informationen scannen) beinhaltet Informationen auf einer ausgewählten Webseite zu

¹⁰ Eine Ausnahme bildet das Modell von (Kiili et al., 2018), welches ebenfalls im Online-Kontext angesiedelt ist und eine sechs-Faktor-Struktur postuliert.

scannen und zu bewerten (S. 1209). Die Bewertung kann anhand des Inhalts, zusätzlicher Informationen zu dieser Webseite aus anderen Webseiten oder anhand von Vorwissen erfolgen (S. 1209). Dieser Schritt dient zur Feststellung, ob die gefundenen Informationen relevant sind und somit zur Lösung des Informationsproblems beitragen könnte (S. 1209). Wenn dem so ist, wird der vierte Schritt der Informationsverarbeitung (Informationen verarbeiten) angesteuert (S. 1209), bei welchem die Informationen im Hinblick auf den Aufbau eines tiefen, z. B. inhaltlich durchdringenden, Verständnisses verarbeitet werden (Schmeck & Geisler-Brenstein, 1989, S. 114). Weiterhin sollte das Ziel sein, die Informationen so zu verarbeiten, dass sie an vorhandenes Wissen anknüpfen (Brand-Gruwel et al., 2009, S. 1209). Die fünfte und letzte Phase (Informationen organisieren und präsentieren) beschreibt die Synthese, im Gegensatz zu den vorherigen analysierend geprägten Phasen (Brand-Gruwel et al., 2009). Weiterhin gilt in Anlehnung an Brand-Gruwel et al. (2009): Die zuvor als relevant erachteten und elaborierten Informationen werden organisiert, kombiniert und anschließend präsentiert, um das Ausgangsproblem zu lösen (S. 1209). Die möglichen Formen der Präsentationen umfassen beispielsweise Poster, eine Präsentation an sich oder einen Aufsatz (S. 1209).

Als Grundlage für den Prozess dienen die so genannten Regulationstätigkeiten, welche sich über den gesamten Prozessverlauf strecken (S. 1209). In wechselseitigem Austausch mit der jeweiligen eigentlichen Phase finden Orientierung, Überwachung, Steuerung und Bewertung statt (S. 1209). Diese Regulationstätigkeiten sind besonders von Bedeutung, um die Lösung des Informationsproblems als Ziel im Auge zu behalten und bei Abweichungen ggf. auf Veränderungen flexibel reagieren zu können (S. 1209).

Zuletzt sind die bedingten Fähigkeiten zu nennen, welche von einem Individuum an sich abhängen und somit je nach seiner Bedingung ausgeprägt sind (Brand-Gruwel et al., 2009). Ein grundsätzliches Leseverständnis sollte bei jedem vorhanden sein, um ein Informationsproblem online lösen zu können (Brand-Gruwel et al., 2009). Hierbei unterscheidet sich das analoge vom digitalen Lesen (insbesondere Lesen von Hypertext) (Coiro, 2003; Sutherland-Smith, 2002). Aufgrund der Vielzahl an Informationen und Verlinkungen in der digitalen Welt muss sich der Leser selbst eine Organisation und Struktur der Informationen schaffen und ist nicht gezwungen einer Leserichtung zu folgen, wie z. B. bei einem ausgedruckten Text (Brand-Gruwel et al., 2009). Aus diesem Grund kommen den Lesefähigkeiten eine besondere Bedeutung zu und verlangen Leseverständnisstrategien, wie zusammenfassen, unklare oder schwierige Wörter nachschauen oder sich selbst inhaltliche Rückfragen stellen (Brand-Gruwel et al., 1998, S. 66, 2009, S. 1210).

Die zweite bedingte Fähigkeit umfasst die Beurteilungsfähigkeit, bei dem ein Individuum in der Lage sein sollte, die Quellen und Informationen beurteilen zu können (Brand-Gruwel et al.,

2009). Insbesondere mit der aktuellen Entwicklung, dass Informationen frei und von jedem nicht verifiziert in rasanter Geschwindigkeit im Internet veröffentlicht werden können, um beispielsweise eine Manipulation der Meinung zu erzielen und zum Kauf eines Produktes anzuregen, wird dieser Fähigkeit zunehmend mehr Bedeutung beigemessen werden.

Die letzte Fähigkeit, einen Computer zu bedienen und sich im Internet zurecht zu finden, liegt ebenfalls als Bedingung zu Grunde (Brand-Gruwel et al., 2009; Sutcliffe & Ennis, 1998).

Zusammenfassend lässt sich sagen, dass das IPS-I-Modell einen guten Überblick über die stattfindenden Prozesse beim Lösen eines Informationsproblems online gibt. Das IPS-I wurde auch bereits empirisch eingesetzt, indem es z. B. als theoretischer Rahmen für Scaffolding¹¹-Maßnahmen (Argelagós & Pifarré, 2012; Raes et al., 2012) oder als Grundlage für die Ableitung und weiterhin Messung der entsprechenden IPS-Fähigkeiten innerhalb eines Trainingsdiente (Frerejean et al., 2016).

Das IPS-I-Modell war überwiegend die Basis für das Konstrukt des kritischen Online-Denkens (COR), welches nachstehend erläutert wird und weiterhin die Grundlage für das in Kapitel 4.1 dargestellte Forschungsprojekt BRIDGE bildet.

2.3.3. Das Konstrukt des Critical Online Reasoning

Die Stanford History Education Group entwickelte das ursprüngliche Konstrukt, das Civic Online Reasoning (Wineburg et al., 2018). Die Idee war, das kritische Denken im Rahmen eines zivilgesellschaftlichen Informationsproblems zu messen, z. B. bei der Beurteilung einer kommerziellen Webseite (Wineburg et al., 2018). Aufgrund soziokultureller und historischer Unterschiede war das ursprüngliche Civic Online Reasoning Assessment nicht auf den deutschen Kontext anwendbar (Molero et al., 2020). Daher wurde das länderspezifische *Critical Online Reasoning* (COR) Konstrukt auf Basis verwandter Konstrukte wie multiple Quellennutzung bzw. -verständnis (Braasch et al., 2018; Goldman & Brand-Gruwel, 2018) Internet-Glaubwürdigkeit (Wierzbicki, 2018) oder dem IPS-I-Modell (Brand-Gruwel et al., 2009) entwickelt (Molero et al., 2020).

Kritisches Online-Denken beinhaltet u. a. die Fähigkeiten Online-Informationen zu suchen, auszuwählen, abzurufen, zu verarbeiten und zu nutzen mit dem Ziel, ein Informationsproblem zu lösen oder Wissen aufzubauen (Molero et al., 2020). Weiterhin ist die Fähigkeit inkludiert, glaubwürdige von unglaubwürdigen Informationen unterscheiden zu können und relevante glaubwürdige Informationen zu nutzen, um eine begründete Argumentation zur Lösung des Informationsproblems zu verfassen (Molero et al., 2020). Die Definition stellt ähnlich zum IPS-

¹¹ Scaffolding beschreibt einen Prozess der Lernunterstützung, bei dem Lehrpersonen dem Lernenden kleine didaktische Hilfestellungen zur Bearbeitung einer Aufgabe geben (Lipowsky, 2020).

I-Modell einen Prozess dar, wie Menschen im Online-Kontext Informationen idealerweise kritisch hinterfragen (sollten) (Molero et al., 2020).

Es können dabei drei Facetten unterschieden werden, welche die Einzelprozesse detailliert abbilden (Molero et al., 2020).

Die erste Phase wird als *Online Information Acquisition* (im Folgenden abgekürzt mit OIA) bezeichnet und stellt den Beginn der Online-Suche dar (Molero et al., 2020). In der Phase finden die Suche und der Zugriff auf Online-Informationen statt: z. B. durch die Nutzung allgemeiner und spezialisierter Suchmaschinen und Datenbanken, die Spezifikation von Suchanfragen, das Öffnen bestimmter Webseiten, Suche nach bestimmten Platzeinträgen und Passagen und Begriffen auf einer Webseite, soweit sie zu einem (effizienten) Zugriff auf relevante und vertrauenswürdige und zur Vermeidung von nicht vertrauenswürdigen Informationen beitragen (Molero et al., 2020). Weiterhin sollten die Fähigkeiten vorhanden sein, relevante Texte auf einer Webseite auszuwählen, zu verstehen, bewerten und zu beurteilen, ob eine Quelle glaubwürdig ist, indem zusätzliche online verfügbare Ressourcen genutzt und mit anderen Suchergebnissen abgeglichen werden (Nagel et al., 2022).

Die zweite Phase wird als *Critical Information Evaluation* (im Folgenden abgekürzt mit CIE) bezeichnet (Molero et al., 2020). Ihr Fokus liegt auf der Auswahl und ersten Evaluation von Informationsquellen (Molero et al., 2020). Als Grundlage für die Evaluation der Glaubwürdigkeit einer Information werden Cues (Deutsch: Hinweise) herangezogen, welche beispielsweise in Form von textlichen oder grafischen Hinweisen vorliegen können (Molero et al., 2020). Mithilfe der z. B. in Kapitel 2.2.2.1 dargestellten Heuristiken können diese Hinweise auf Glaubwürdigkeit beurteilt werden. Weiterhin sollten Argumente und ihre Bestandteile, die in den online gefundenen Quellen verwendet werden, erkannt und im Hinblick auf Beweiskraft, Objektivität, Gültigkeit und Konsistenz bewertet werden (können) (Nagel et al., 2022).

Die letzte Phase stellt das *Reasoning with Evidence, Argumentation and Synthesis* (im Folgenden abgekürzt mit REAS) dar (Molero et al., 2020). In REAS werden alle anfänglich bewerteten Informationen zusammengetragen und gegeneinander und mit zusätzlichen Informationen abgewogen (Molero et al., 2020). Anschließend werden die als nützlich empfundenen Informationen verwendet, um eine Argumentation zu schreiben und das Ausgangsproblem zu lösen (Molero et al., 2020). Die Schlussfolgerung sollte dabei auf verlässlichen Quellen beruhen und eine gut strukturierte und logisch zusammenhängende Erklärung bzw. Entscheidung beinhalten (Nagel et al., 2022).

Eine letzte darüber hinaus gehende Phase, die der metakognitiven Selbstregulation, beschreibt die Fähigkeit, die COR-Kompetenzen in entsprechenden Situationen überhaupt

aktivieren zu können (Molero et al., 2020). Alle Bestandteile von COR werden in Abbildung 6 vereint.

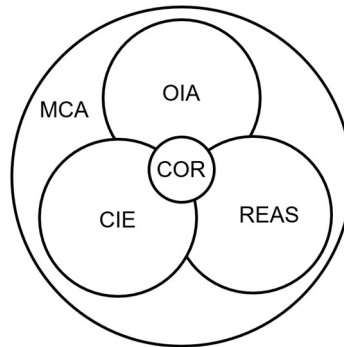


Abbildung 6: COR-Konstrukt, eigene Abbildung in Anlehnung an Molero et al. (2020).

Das theoretisch entwickelte COR-Konstrukt dient in einem weiteren Schritt zur Messung der COR-Performanz. Hierbei wird die Unterscheidung zwischen generischem COR (Englisch: generic COR, im Folgenden abgekürzt mit GEN-COR) und domänen-spezifischen COR (Englisch: domain-specific COR, im Folgenden abgekürzt mit DOM-COR) relevant. Bei erst genannten geht es um das allgemeine kritische Online-Denken, welches sich auf ein generelles, meist alltägliches Problem bezieht und für junge Erwachsene nach einem zweiten Bildungsabschluss (Molero et al., 2020) kein Problem darstellen sollte. Das domänenspezifische kritische Online-Denken erfordert hingegen fachspezifisches Domänenwissen (Molero et al., 2020). Vorliegende Forschung konzentriert sich stets auf das allgemeine, generische COR (GEN-COR).

Wie eine neuere Forschung zeigt, konnte das COR-Konstrukt auch bereits empirisch erfolgreich erprobt werden (Nagel et al., 2022).

2.4. Transfer und Anwendung der theoretischen Grundlagen in den Forschungskontext

Wie anhand aller Modelle ersichtlich, spielen bei der Informationsverarbeitung die Merkmale der zu verarbeitenden Informationen sowie individuelle kognitive Prädispositionen eine entscheidende Rolle. Welchen Einfluss sie auf die einzelnen Prozessschritte haben, soll in vorliegender Arbeit untersucht werden. Aus diesem Grund ist es notwendig, zunächst die Prozessschritte und Zusammenhänge mit der allgemeinen Informationsverarbeitung und Selektion bzw. mit dem IPS-I-Modell und COR-Konstrukt zu verstehen und auf vorliegende Situation zu transferieren. Daraus wird eine überblicksartige Prozessabfolge abgeleitet, welche die Grundlage für die Herleitung der Performanz- und Prozessindikatoren im darauffolgenden Kapitel darstellen wird.

Wiederholung. Zum jetzigen Stand wurden die Begriffe der Glaubwürdigkeit und verwandte Konstrukte definiert sowie die Grundlagen der klassischen Informationsverarbeitung und Selektion geschaffen. Weiterhin wurde sich begründet für das HSM entschieden, welches auch für die spätere Hypothesenherleitung dienlich sein wird. Damit zusammenhängend wurden wesentliche Heuristiken thematisiert. Für die Überleitung zum Online-Kontext wurden die Grundlagen aus der Informationsverarbeitung mithilfe des 3-S-Modells und dem IPS-I-Modell bzw. COR-Konstrukt dargestellt.

Grundlagen der Prozessmodellierung. Wie anhand der Ausführungen ersichtlich, überschneidet sich das COR-Konstrukt konzeptuell mit dem IPS-I-Modell (Nagel et al., 2022), denn COR resultiert überwiegend aus dem IPS-I. Für vorliegende Forschungsarbeit wird sich hauptsächlich auf das COR-Konstrukt bezogen, jedoch werden vereinzelt theoretische Aspekte mithilfe des IPS-I-Modell ergänzt, sodass erst die Kombination beider Informationsmodelle die vollständige theoretische Grundlage bildet. Für die Darstellung der Prozesse werden zudem die allgemeinen Informationsverarbeitungsmodelle berücksichtigt und unterliegen diesen Prozessen. Jedoch erhalten sie im Vergleich zu COR und IPS-I eine geringere Beachtung. Der Fokus auf das IPS-I-Modell sowie COR wird einerseits mit dem digitalen Kontext und weiterhin mit kleinen konzeptionellen Unterschieden gerechtfertigt:

Das IPS-I-Modell wird in der Literatur im Themengebiet der Informationsverarbeitung im Digitalkontext als das am verständlichsten und meist anerkannte Modell angesehen (Zhou & Lam, 2019) und wurde auch bereits erfolgreich empirisch eingesetzt (z. B. Frerejean et al., 2016; Raes et al., 2012). Auch COR erwies sich bislang als geeignetes Konstrukt zur Erfassung des kritischen Online-Denkens (Molero et al., 2020; Nagel et al., 2022). Da sich vorliegende Forschungsarbeit im Digitalkontext befindet und sich für die Forschungsfrage das IPS-I und COR ergänzen, stehen sie somit hauptsächlich im Fokus der Arbeit.

Zweitens unterscheiden sich die Modelle hinsichtlich ihres Gegenstandes bzw. ihrer Ausrichtung. Das HSM als „Prozessmodell“ bezieht sich auf die Informationsverarbeitungsprozesse, die innerhalb der Person geschehen und auf kognitiver Ebene anzusehen sind (S. Chen & Chaiken, 1999). Dies verdeutlicht erneut, dass kognitive Dispositionen bei jedem Prozessschritt eine Rolle spielen könnten und deshalb mitunter in den Vordergrund der Untersuchung zu stellen sind. Das IPS-I-Modell wie auch das COR-Konstrukt beschreiben hingegen darüberliegende Prozessschritte, die eine Person im Online-Kontext tätigen. So gesehen kann jeder einzelne Prozessschritt (z. B. Informationssuche in Phase zwei des IPS-I) auch auf heuristischem oder systematischem Wege innerhalb der Person stattfinden.

Ein Beispiel für die heuristische Herangehensweise in Phase zwei wäre, die Suchergebnisse aus der Trefferliste der Suchmaschine aufgrund von Heuristiken, beispielsweise der

Bekanntheitsheuristik (Novin & Meyers, 2017), zu selektieren. Bei der systematischen Verarbeitung ist es denkbar, dass die Suchergebnisse und Textausschnitte, sofern die Textlänge es ermöglicht, auf Inhaltliches überprüft (z. B. relevante Schlagwörter) und mit vorhandenem Vorwissen abgeglichen werden.

Was anhand dieser beispielhaften Anschauung verdeutlicht werden soll, ist, dass das IPS-I-Modell und das COR-Konstrukt ganzheitlich die Prozesse im Online-Kontext darstellen, wohingegen die allgemeinen Prozesstheorien lediglich die Personenebene thematisieren. Diese sind somit zwar auch beteiligt und bedeutsam, da sie für das Verständnis der kognitiven Prozesse innerhalb einer Person und für die spätere Hypothesenherleitung essenziell sind. Das Gros der theoretischen Basis wird allerdings die soeben dargestellten Informationsmodelle und die daraus abgeleiteten Prozess- und Performanzindikatoren bilden.

Transfer und Anwendung. Alle zu diesem Zeitpunkt dargestellten Theorien werden nachfolgend in Abbildung 7 vereint, um den kleinschrittigen Prozess der Informationsverarbeitung und zum Lösen des Informationsproblems für das vorliegende Forschungsvorhaben zu veranschaulichen. Eines muss jedoch im Vorfeld verdeutlicht werden: An jeder Stelle des Prozesses kann die Person entscheiden, wieder zurück zum „Start“ zu gehen oder einen vorherigen Schritt (z. B. neue Formulierung der Suchbegriffe) zu machen. Dies wurde in der Abbildung an nur wenigen Stellen exemplarisch dargestellt (durch Rückwärtspfeile). Eine Vereinheitlichung der Prozessschritte ist aufgrund der vielen möglichen Verzweigungen und individuellen Unterschiede hinsichtlich des Vorwissens, Kognitionen etc. nicht möglich. Der Prozess ist lediglich eine vereinfachte Darstellung eines möglichen, idealtypischen Verlaufs.

Eine Person wird mit einem Informationsproblem konfrontiert. Sie soll ein offenes (Englisch: ill-defined) Problem (Jonassen, 1997, S. 69) lösen, zu dem keine klare Ja/Nein-Antwort existiert. Hierzu soll sie Informationen aus dem Internet zusammentragen und eine begründete Entscheidung zur Ursprungsproblematik abgeben. Laut IPS-I repräsentiert das die *erste Phase*, nämlich das Informationsproblem zu definieren. Anschließend entwickelt die Person ein *Selektionsinteresse*, denn sie benötigt zur Lösung des Problems Informationen, die sie zunächst sichten muss. Aus diesem Grund wird sie eine erste Suchanfrage tätigen (*zweite Phase IPS-I, OIA-Phase von COR*). Hierbei kann sie z. B. oberflächliche oder spezifische Suchbegriffe in die Suchmaschine eingeben. In der *Selektionsorientierungs-Phase* orientiert sich die Person an Merkmalen, die sie aus der Umwelt wahrnimmt. Der Wahrnehmungsraum beschränkt sich hierbei auf die Merkmale aus dem Internet, sprich Merkmale der Suchmaschine sowie der Suchergebnisseite, die allerdings per se nicht viele sind und auch in vorliegender Forschungsarbeit nicht untersucht werden (z. B. Ranking der Links, Inhalt der Ausschnitte der Suchergebnisseite etc.). Es erfolgt an dieser Stelle bereits eine Informationswahrnehmung und -verarbeitung (Beurteilung der Suchergebnisse in *zweiter Phase IPS-I, OIA-Phase von COR*). Die wahrgenommenen Suchergebnisse werden entschlüsselt, in mentale Repräsentationen umgewandelt und mit vorhandenem Wissen aus dem Langzeitgedächtnis verglichen bzw. interpretiert. Daran könnte sich eine erste *Selektionshandlung* anknüpfen, weil der Abgleich möglicherweise ergeben hat, dass die wahrgenommenen Informationen neu und relevant sind. Die erste Selektionshandlung ist die Webseite aus der Suchergebnisseite auszuwählen. Nachdem eine erste Webseite geöffnet wird, werden neue Informationen und eine Vielzahl an Webseitenmerkmalen präsentiert, die aufs Neue wahrgenommen und verarbeitet werden müssen (*dritte Phase IPS-I, OIA-CIE-Phase von COR*). Das sind letztlich die zu untersuchenden Webseitenmerkmale für das Forschungsvorhaben. Die Prozesse aus der Mikroperspektive innerhalb der Selektionsorientierung laufen dann erneut ab. Ein wahrgenommenes Merkmal, wie die Angabe eines Autors auf einer Webseite, kann aufgrund einer Heuristik, wie der Bekanntheitsheuristik, hinsichtlich der Qualität oder Glaubwürdigkeit beurteilt werden. Hier kann anschließend eine *zweite Selektionshandlung* folgen, wenn die Evaluation der Webseite und ihren Merkmalen zu Gunsten dieser ausgefallen ist (z. B. die Merkmale wurden als glaubwürdig eingeschätzt) und sie für die eigentliche Beantwortung der Aufgabenstellung herangezogen werden kann (*vierte Phase IPS-I, CIE-Phase von COR*). Ist dies nicht der Fall, so wird die Person möglicherweise zur Suchergebnisseite zurückkehren, entweder erneut suchen, die Glaubwürdigkeit anhand zusätzlicher Informationen prüfen, eine andere Webseite auswählen oder eine Kombination dieser Handlungen tätigen. Hierdurch werden die Schritte der Selektionsorientierung, Informationswahrnehmung und -verarbeitung, sowie Selektionshandlung und Evaluation immer wieder neu durchgeführt, bis die *fünfte IPS-I-Phase bzw. REAS-Phase aus COR* erreicht wird und die gefundenen Informationen organisiert und

synthetisiert werden. Mithilfe der zusammengeführten theoretischen Grundlagen und der dargestellten Prozessschritte lassen sich im Folgenden die einzelnen Indikatoren als abhängige Variable ableiten.

2.5. Theoretische Herleitung der Prozess- und Performanzindikatoren als abhängige Variablen

Das Konstrukt COR (Molero et al., 2020; Nagel et al., 2020; Zlatkin-Troitschanskaia, Brückner et al., 2021) und das IPS-I-Modell (Brand-Gruwel et al., 2009) beschreiben und erfassen den idealtypischen Prozess, wie Menschen mit Online-Informationen umgehen. Im Nachfolgenden werden anhand theoretischer Überlegungen aus dem IPS-I und COR Performanz- und Prozessindikatoren hergeleitet, welche als abhängige Variablen in die Modelle mit eingehen werden. Die einzelnen Indikatoren basieren dabei auf den theoretischen Vorüberlegungen und können als beobachtbares Verhalten abgeleitet werden. Alle darüberhinausgehenden, weiteren Prozessschritte, welche zwar im vorherigen Kapitel in Abbildung 7 dargestellt wurden, fließen nicht in das Forschungsvorhaben ein. Grund hierfür ist, dass die vorliegende Arbeit ausschließlich die Prozess- und Performanzdaten analysieren wird, was bereits einen ausreichend breiten Fokus darstellt. Die Abbildung 7 zeigt zudem die abgeleiteten Indikatoren aus der Such- und Evaluationsphase.

Prozessindikatoren beschreiben, welche Abfolge von Prozessschritten eintreten (können). Die drei abgeleiteten abhängigen Variablen stellen Prozessschritte dar, die sich interindividuell unterscheiden und im Suchprozess verankert sind. *Performanzindikatoren* messen das Ergebnis eines abgeschlossenen Prozesses¹². Die abgeleiteten abhängigen Variablen sind im Verwertungs- bzw. Evaluationsprozess verankert und messen die Performanz hinsichtlich von drei Indikatoren.

2.5.1. Prozessindikatoren aus dem Suchprozess

Spezifität des Suchbegriffes. In der zweiten Phase des IPS-I-Modell wird beschrieben, dass Individuen Suchbegriffe spezifizieren sollten (Brand-Gruwel et al., 2009), was im COR-Konstrukt in der OIA Phase mit der Spezifikation von Suchanfragen bezeichnet ist (Molero et al., 2020). Aus der Kombination der beiden theoretischen Annahmen wird abgeleitet, dass die Festlegung des Suchbegriffs als ein wesentlicher Schritt verstanden wird, wobei die Spezifität des ausgewählten Suchbegriffes variieren kann. Personen könnten sich hinsichtlich der Formulierung eines spezifischen oder weniger spezifischen Suchbegriffs unterscheiden.

¹² Die beschriebenen Performanzindikatoren könnten auch als Prozessvariablen mit einfließen, da die Schritte immer analytisch trennscharf voneinander zu unterscheiden sind. Dennoch wird für vorliegende Forschungsarbeit der Begriff der Performanz mit der Bedeutung, dass dies ein Endergebnis eines vorangegangenen Prozesses darstellt, verwendet.

Nachrecherche. In der dritten Phase im IPS-I-Modell elaborieren Informationsverarbeiter gefundene Inhalte und kombinieren sie mit ihrem Vorwissen oder anderen Informationen (Brand-Gruwel et al., 2009). Im COR-Konstrukt wird weiter ausgeführt, dass die Quellenglaubwürdigkeit "anhand zusätzlicher online verfügbarer Ressourcen und durch Quervergleiche mit anderen Suchergebnissen" (Nagel et al., 2020, S. 3) beurteilt wird. Die Untersuchung der Quellenglaubwürdigkeit mithilfe einer neuen Suche wird als „laterales Lesen“ bezeichnet (Breakstone, Smith, Connors, et al., 2021; Wineburg & McGrew, 2019). Andere online verfügbare Ressourcen bzw. Informationen können beispielsweise Informationen über die Quelle selbst liefern, indem nach erstem Aufruf der ursprünglichen Webseite in einer weiteren Suche über die Quelle recherchiert wird. Die gefundenen Informationen über die Quelle können dann mit den ersten Informationen kombiniert werden, um zu einem Urteil über die Quellenglaubwürdigkeit (d. h. durch laterales Lesen) oder der gefundenen Informationen selbst zu gelangen. Alternativ können Personen auch anhand von weiteren Angaben auf einer Webseite prüfen, ob es sich um eine glaubwürdige Quelle handelt, z. B. mit Hilfe des Impressums oder der „Über-Uns“-Seite. Dies stellt einen zusätzlichen Aufwand dar, der zur Glaubwürdigkeitsbeurteilung und Einschätzung der Informationen hilfreich sein kann. Die Quelle erneut zu suchen oder sich durch Anschauen der Impressums-Seite darüber zu informieren und dadurch die Quellenglaubwürdigkeit und/oder die gefundenen Informationen zu überprüfen, ist daher nicht zwingend notwendig, aber eventuell förderlich. Es ist wichtig zu erfahren, ob Personen dieses Verhalten zeigen und darüber hinaus, ob es interindividuelle Unterschiede gibt.

Anklicken der Weiterverlinkung. Die Phase der Suche und Evaluation der Suchergebnisse beinhaltet, gefundene Inhalte mit zusätzlich verfügbaren Quellen oder anderen Informationen zu kombinieren (Brand-Gruwel et al., 2009; Nagel et al., 2020, S. 3). Das Konzept „SIFT“ beschreibt „Stop, Investigate the source, Find better coverage, and Trace claims, quotes, and media to the original context“ (Caulfield, 2019, S. 1). Zusätzlich online verfügbare Quellen zu Rate zu ziehen, könnte gemäß dem SIFT-Konzept somit implizieren, die Quellen nach ihren Originalen zurückzuverfolgen und dadurch die Informationen bzw. Webseitenglaubwürdigkeit zu evaluieren. Dies ist ratsam, aber nicht zwingend notwendig. Es ist folglich wichtig zu erfahren, ob Individuen dieses Verhalten zeigen und ob sie sich intraindividuell unterscheiden.

2.5.2. Performanzindikatoren aus dem Evaluationsprozess

Korrekte Quelleneinschätzung. In der vierten Phase des IPS-I-Modells werden die gefundenen Informationen analysiert und auf Qualität und Nützlichkeit bewertet (Brand-Gruwel et al., 2009). Innerhalb des COR-Konstruktes wird dieser Prozess unter CIE beschrieben und umfasst das Bewerten von Informationen aufgrund von Webseitenmerkmalen oder Hinweisen (wie Texte oder Grafiken) (Molero et al., 2020). Wie die Glaubwürdigkeit und somit Qualität

einer Quelle bewertet wird bzw. ob die Einschätzung korrekt ist, könnte interindividuell unterschiedlich sein und ist für vorliegende Forschung von Interesse. Eine korrekte Quelleneinschätzung bedeutet hierbei, eine glaubwürdige Quelle als korrekt glaubwürdig oder eine unglaubwürdige Quelle als korrekt unglaubwürdig zu erkennen.

Systematische Begründungen. Auf Basis der selben theoretischen Grundlage – vierte Phase des IPS-I-Modells (Brand-Gruwel et al., 2009) und CIE von COR (Molero et al., 2020) – wird angenommen, dass sich die Bewertungen von Informationen oder Quellen auf der Grundlage von Merkmalen oder Hinweisen und folglich die Begründungen für bzw. gegen die Quellenglaubwürdigkeit unterscheiden können. Weiterhin können diese Begründungen Rückschlüsse auf die Verarbeitungstiefe i. S. v. heuristisch oder systematisch ermöglichen. Das bereits in Kapitel 2.3.1 vorgestellte 3-S-Modell findet in für den Kontext abgewandelter Form eine Anwendung. Die Autoren teilen Hinweisreize zur Beurteilung des Vertrauens als vorgelagerten Prozess zur Glaubwürdigkeit in drei Kategorien (Semantikmerkmale, Oberflächenmerkmale und Quellenmerkmale) ein, welche auch hier zum Einsatz kommen sollen, sodass Rückschlüsse auf die Informationsverarbeitung – heuristisch oder systematisch – gezogen werden können. Jedoch distanzieren sich die Autoren des 3-S-Modells davon, ihre Strategien eindeutig den zwei Routen des HSM zuzuordnen (Lucassen & Schraagen, 2011, S. 1236): Sie folgern lediglich, dass die Bewertung von Oberflächenmerkmalen eindeutig heuristisch ist, die Beurteilung der Quellen- und Semantikmerkmale allerdings sowohl heuristisch als auch systematisch ablaufen kann. Tabelle 1 zeigt die ursprüngliche Konzeption der Einteilung der Merkmalskategorien sowie das angepasste Modell.

Tabelle 1: Anwendung des 3-S-Modell für den Forschungskontext.

3 S Modell (Lucassen et al., 2013, S. 256)

Semantik		Quelle		Oberfläche	
Systematisch/heuristisch		Systematisch/heuristisch		Heuristisch	

Angepasstes Modell für das Forschungsvorhaben

Text		Autor		Medium	
Systematisch	Heuristisch (+ Oberflächenmerkmale)	Systematisch	Heuristisch (+ Oberflächenmerkmale)	Systematisch	Heuristisch (+ Oberflächenmerkmale)

Für die Herleitung der abhängigen Variablen wird dieser Einschätzung der Autoren jedoch nicht gänzlich gefolgt. Da die Literatur zeigt, dass die Quelle im Online-Kontext nicht immer eindeutig identifizierbar oder voneinander (Autor vs. Medium) trennbar ist (Sundar, 2008, S.

83; Sundar & Nass, 2001, S. 53), wird „Quelle“ per se bereits zu Beginn in „Autor als Quelle“ und „Medium als Quelle“ separiert. Damit soll eine korrekt zugeschriebene und detailliertere Einteilung ermöglicht werden. Zudem werden die als heuristisch betrachteten Oberflächenmerkmale als Unterkategorien zu den je thematisch passenden Kategorien Text, Autor oder Medium subsumiert. Wie Lucassen et al. (2013, S. 255) feststellen, beziehen sich die Oberflächenmerkmale auf die Art und Weise, wie Informationen präsentiert werden. Jedoch zeigt sich in Anbetracht der ihnen verwendeten Beispiele (wie Bilder, Schreibstil oder Referenzen), dass die präsentierten Informationen unterschiedlichen Bewertungsobjekten (also Text, Medium oder Autor) zugrunde liegen. Beispielsweise ist die Angabe von Referenzen oder der Schreibstil basierend auf der Beurteilung des Textes, wohingegen die Darstellung von Bildern tendenziell dem Medium als Betreiber der Webseite zuzuschreiben ist. Wird dieser Einteilung gefolgt, so würde sich eine Redundanz ergeben, da z. B. das Merkmal des Schreibstils laut ihrer Einteilung als Oberflächenmerkmal zählt, es jedoch auch den textlichen Merkmalen zugeordnet werden könnten. Die Kategorie der Oberflächenmerkmale wird daher in vorliegender Forschung aufgebrochen und den anderen Kategorien (Autor bzw. Medium bzw. Text) zugeordnet.

Die Anpassung des ursprünglichen Modells führt schließlich zu der in (s. Tabelle 7 Abschnitt „Einordnung und Begründung“).

) dargestellten konzeptionellen Einteilung. Folglich werden die von den Probanden angegebenen Hinweise gemäß dieser Abbildung in die Kategorie Text, Medium und Autor eingeordnet, dass daraus weiterhin eine Schlussfolgerung auf die systematische oder heuristische Informationsverarbeitung geschlossen werden kann. Zu erwartende Begründungen könnten bezüglich folgender Beispiele genannt werden:

Qualität der Argumente im Text (Text – systematisch); Anzahl der Argumente im Text (Text – heuristisch), Länge des Textes (Text – heuristisch); beruflicher Werdegang des Autors (Autor – systematisch), Autor angegeben (Autor – heuristisch), Überprüfung des Mediums durch unabhängigen Dritten, z. B. Mediabiasfactcheck (*Media Bias/Fact Check News*, 2023)¹³ (Medium – systematisch), Genre des Mediums (Medium – heuristisch).

Konkret wird angenommen, dass Personen, die Begründungen aus den systematischen Kategorien angeben, sich tiefergehend mit den Inhalten auf der Webseite auseinandergesetzt

¹³ Eine systematische Auseinandersetzung mit dem Medium als Quelle zur Überprüfung stellt sich als problematisch dar. Die Überprüfung eines Mediums durch einen unabhängigen Dritten, wie beispielsweise der Webseite *Mediabiasfactcheck* könnte ebenfalls als Heuristik, nämlich als die Bandwagon-Heuristik (s. Kapitel 2.2.2) verstanden werden. Insofern wäre dies ebenfalls eine heuristische Informationsverarbeitung. Jedoch ist es im Vergleich zur einfachen Beurteilung des Genres eines Mediums aufwändiger, wird als systematischer bezeichnet und demnach als Beispiel hierfür aufgeführt.

haben werden. Denn systematische Informationsverarbeitung bedeutet die Argumentationsstruktur bzw. den Inhalt einer Nachricht genauer zu verarbeiten (Dillard & Marshall, 2003). Wenn der Webseiteninhalt und die Argumentation genau und somit systematisch verarbeitet wurde, könnte die Begründung für bzw. gegen die Glaubwürdigkeit ebenfalls aus der systematischen Kategorie sein. Umgekehrt könnten Begründungen aus den heuristischen Kategorien als Indiz für die heuristische Informationsverarbeitung gesehen werden. Dieser Performanzindikator sagt daher aus, ob ein Proband mind. eine Begründung aus der systematischen Kategorie angegeben oder nur Begründungen aus der heuristischen Kategorie verwendet hat.

Argumentationsstruktur. Mit „Synthese“ in Phase fünf im IPS-I-Modell (Brand-Gruwel et al., 2009) nur kurz angerissen, wird innerhalb des COR-Konstruktes eine ausführlichere Darstellung der Prozesse gewählt: Denn Individuen sollen im letzten Schritt die anfänglich bewerteten Informationen zusammenführen, mithilfe weiterer Indizien gewichten und als Beweismittel nutzen, um ein überzeugendes Argument zu konstruieren (Molero et al., 2020). Weiterhin ist die Anforderung, zu einer Schlussfolgerung zu gelangen, die eine gut strukturierte und logisch zusammenhängende Entscheidungsbegründung beinhaltet (Nagel et al., 2020). Es wird daraus angenommen, dass die gesamte Schlussfolgerung bzw. Argumentation bestenfalls gut strukturiert ist. Daher sind die Struktur dieser Argumentation (i. S. v. gut strukturiert und logisch zusammenhängend) und mögliche interindividuelle Unterschiede für diese Forschungsarbeit von Interesse.

2.6. Merkmale der Akteure im Informationsverarbeitungsprozess

Da die Glaubwürdigkeitsbeurteilung eine subjektive Zuschreibung darstellt, sind uneindeutige Aussagen über die objektive Glaubwürdigkeit von Informationen oder Personen nur schwer möglich. Denn das Konstrukt der Glaubwürdigkeit ist, anders als direkt beobachtbare und somit manifeste Variable, wie das Einkommen, nicht direkt messbar und folglich latent (Eid et al., 2017, S. 84). In einem Prozess, bei dem subjektives Empfinden eine Rolle spielt, müssen viele verschiedene Faktoren in Betracht gezogen werden, wie z. B. personenbedingte (Trait) sowie auch situative (State) Faktoren (Steyer et al., 1999). Vorliegende Forschung versucht daher, zumindest zwei Perspektiven mit diversen Faktoren abzudecken. Im Nachfolgenden werden daher die Perspektive der Informationsanbieter (Medien; State) und die der Informationskonsumenten (Personen; Trait) theoretisch erläutert, in den aktuellen Forschungsstand eingebettet und mit Glaubwürdigkeit in Zusammenhang gebracht.

2.6.1. Für Informationsverarbeitung relevante Merkmale im Informationsangebot

Um eine Annäherung an ein vollumfängliches Konstrukt der Glaubwürdigkeit erhalten zu können, wurden die Merkmale der Informationsanbieter in Vorbereitung auf diese Arbeit mittels einer ausgiebigen Literaturrecherche herausgearbeitet, was in Tabelle 2 dargestellt ist.

Auffällig ist, dass sich aus der Literatur zwei Wege zur Erfassung von subjektiv empfundener Glaubwürdigkeit herauskristallisierten. In qualitativen Arbeiten werden z. B. Kommentare zur Glaubwürdigkeitsbeurteilung sowie Think-Aloud-Interviews¹⁴ von Probanden inhaltlich analysiert. Außerdem können thematisch passend hierin auch konzeptionelle Daten, wie theoretische Modelle über die Glaubwürdigkeitswahrnehmung, oder Reviews eingeordnet werden. Bei quantitativen Arbeiten wird zum einen der Einfluss der Merkmale durch bewusste Manipulation auf diverse abhängige Variablen z. B. Evaluation von Webseiten bzw. Information (de Meo, 2012; Scharrer et al., 2013) oder Glaubwürdigkeitsbeurteilung (u. a. Flanagin & Metzger, 2007; Denner et al., 2016) experimentell erfasst. Weiterhin können hierzu auch Befragungen, bei denen Probanden explizit nach Merkmalen für bzw. gegen die Glaubwürdigkeit gefragt werden (z. B. Metzger et al., 2003), zugeordnet werden.

Tabelle 2: Kategorisierung der Merkmale nach unterschiedlichen Erhebungen in Untersuchungen.

Webseiten-merkmal	Experimente, Befragungen	Qualitative Analyse von Interviews, Kommentaren und Inhalten von als glaubwürdig erachteten Webseiten, Reviews, konzeptionelle Modelle
Aktualität	- Metzger, Flanagin & Zwarun, 2003 [Umfrage] <i>Merkmal als AV als Teil der Verifikationsstrategie</i> - Unkel, 2019 [Experiment] <i>Suchresultate</i> T. Hong, 2006 [Inhaltsanalyse der als glaubwürdig bewerteten Webseiten und Fragebogen]	- Tate, 2019 [Review] - Walraven et al. 2009 [qualitative Analyse v. TA und Gruppendiskussion]; - Hope 2007 [qualitative Analyse von TA und Videos]
Verständlichkeit	- De Meo, 2012 [Experiment] <i>Verständlichkeit von Gehörtem</i> - Scharrer et al., 2013 [Experiment] - Metzger, Flanagin & Zwarun, 2003 [Umfrage] <i>Merkmal als AV als Teil der Verifikationsstrategie</i>; - Scharrer, 2021 [Experiment]; - Scharrer, 2022 [Experiment]; - Lev-Ari (2010) [Experiment] <i>Verständlichkeit von Gesprochenem</i>	
Transparenz	- Thomm & Bromme, 2012 [Experiment] - Curry & Stroud, 2021 [Umfrage, Experiment] <i>kumuliert mit weiteren, sich überschneidenden Merkmalen</i>	Hilligoss & Rieh, 2008 [qualitative Analyse von Tagebucheinträgen und Interviews]
Quellenherkunft		- Hilligoss & Rieh, 2008 (qualitative Analyse von Tagebucheinträgen und Interviews) - Walraven et al. 2009 [qualitative Analyse v. TA und Gruppendiskussion];
Angabe von Quellen	Fogg, Marshall, Lakari, et al., 2001 [Umfrage]	Walraven et al. 2009 [qualitative Analyse v. TA und Gruppendiskussion];

¹⁴ Die Think-Aloud-Methode (van Someren et al., 1994), oder auch „Methode des lauten Denkens“, ist eine Methode zur Erhebung und Sichtbarmachung von Denkprozessen (McGrew, 2021b, S. 77), bei der Probanden ihre Gedanken mit einer weiteren Person laut besprechen, ohne dass jedoch durch das Verbalisieren die kognitiven Prozesse und Strukturen stark verändert werden (Ericsson & Simon, 1980, S. 235).

	• Thomm & Bromme, 2012 [Experiment] • Henke et al., 2019 [Experiment] • Johnson & Wiedenbeck, 2009 [Experiment] <i>Hyperlinks als Teil der Quellenangabe</i> • Borah, 2014 [Experiment] <i>Hyperlinks als Teil der Quellenangabe</i> • Liu & Huang, 2005 [Umfrage] • Liu, 2004 [Umfrage] • Curry & Stroud, 2021 [Umfrage, Experiment] <i>zusammengefasst unter Transparenz</i> • Karlsson & Clerwall 2018 [Umfrage und Fokusgruppeninterview] <i>Hyperlinks als Teil der Quellenangabe</i> • T. Hong, 2006 [Inhaltsanalyse der als glaubwürdig bewerteten Webseiten und Fragebogen]	• Dissen et al., 2022 [qualitative Analyse von Antworten] <i>Glaubwürdigkeit wird als Augenscheinvalidität operationalisiert;</i>
Struktur der Information bzw. Webseite	• Sharp & McClung, 1966 [Experiment] <i>AV war Einstellungsänderung, nicht explizit Glaubwürdigkeit;</i> • McCroskey & Mehrley, 1969 [Experiment] <i>Struktur von Rede</i> • Liu, 2004 [Umfrage]	• Fogg et al., 2003 [qualitative Analyse von Kommentaren]; • Rieh 2002 [qualitative Analyse von TA, Logs und Interviews] <i>aber Informationsqualität, nicht Glaubwürdigkeit</i>
Faktische Korrektheit	• Lucassen & Schraagen, 2011 [Experiment] <i>Vertrauen, nicht explizit Glaubwürdigkeit</i>	• Wathen & Burkell, 2002 [Review] • Hope 2007 [qualitative Analyse von TA und Videos] • Fogg et al., 2003 [qualitative Analyse von Kommentaren] • Tate, 2019 [Review]
Statistiken	• Middaugh, 2018 [Erinnerungs-Manipulation der Experimentalgruppe; qualitative Analyse der TA] • Henke et al., 2019 [Experiment] • Koetsenruijter, 2011 [Experiment] <i>analoge Medien</i> • Allen & Preis, 1997 [Meta-Analyse] <i>Persuasion</i> • Freling et al., 2020 [Meta-Analyse] <i>Persuasion</i> • T. Hong, 2006 [Inhaltsanalyse der als glaubwürdig bewerteten Webseiten und Fragebogen]	• Hornikx 2005 [Review] <i>Persuasion</i>
Vollständigkeit	• Metzger, Flanagin & Zwarun, 2003 [Umfrage] <i>Merkmal als AV als Teil der Verifikationsstrategie</i>	• Hope 2007 [qualitative Analyse von TA und Videos]
Sprachstil	• Schaewitz, 2020 [Experiment] • König & Jucks, 2019a [Experiment] • König & Jucks, 2019b [Experiment]	• Kattenbeck & Elsweiler (2019) [qualitative Analyse von TA] <i>aber nur SERP-Snippets, d. h. Suchphase</i>
Wissenschaftlichkeit	• Jonas et al., 2024 [Experiment] • Thon & Jucks, 2017 [Experiment] • Joiner et al., 2002 [Experiment] <i>Schnittmenge mit Verständlichkeit und Videoaufnahme</i>	
Argumentqualität	• Stoltenberg & Davis, 1988 [Experiment] <i>Printmedien</i> • K. Z. K. Zhang et al., 2014 [Fragebogen] <i>umgekehrter Effekt: Quellenglaubwürdigkeit hat positiven Effekt auf Argumentqualität</i>	

Narrative	• Cheung et al., 2012 [Fragebogen] <i>Konsumentenbereich</i> • Hamilton 1998 [Meta-Analyse] • Boukes & LaMarre (2020) [Experiment, Fragebogen]	
	• Clementson (2020) [Experiment]; • Huang & Wang, 2020 [Experiment]	
Spezialisierungsgrad	• Koh & Sundar (2010) [Experiment] <i>Kundenvertrauen/ nicht Glaubwürdigkeit</i>	
Design	• Robins & Holmes, 2008 [Experiment] • Syrief & Hibino, 2007 [Experiment] • Fogg, Marshall, Lakari, et al., 2001 [Umfrage] • Liu & Huang, 2005 [Umfrage] • Liu, 2004 [Umfrage]	• Fogg et al., 2003 [qualitative Analyse von Kommentaren] • Hilligoss & Rieh, 2008 [qualitative Analyse von Tagebucheinträgen und Interviews]; • Rieh 2002 [qualitative Analyse von TA, Logs und Interviews] <i>aber Informationsqualität, nicht Glaubwürdigkeit</i> • Hope 2007 [qualitative Analyse von TA und Videos] • Dissen et al., 2022 [qualitative Analyse von Antworten] <i>Glaubwürdigkeit wird als Augenscheinvalidität operationalisiert;</i>
Genre	• Flanagin & Metzger, 2007 [Experiment]; • Denner et al. 2016 [Experiment]	• Rieh 2002 [qualitative Analyse von TA, Logs und Interviews] <i>aber Informationsqualität, nicht Glaubwürdigkeit</i>
Kommerzielles Interesse	• Metzger, Flanagin & Zwarun, 2003 [Umfrage] <i>Merkmal als AV als Teil der Verifikationsstrategie</i> • Fogg, Marshall, Lakari, et al., 2001 [Umfrage] • Liu, 2004 [Umfrage] • Haas & Unkel, 2017 [Beobachtungsstudie] <i>als Neutralität bezeichnet</i>	• Tate, 2019 [Review] • Walraven et al. 2009 [qualitative Analyse v. TA und Gruppendiskussion]; • Fogg et al., 2003 [qualitative Analyse von Kommentaren] • Hope 2007 [qualitative Analyse von TA und Videos] • Kattenbeck & Elswiler (2019) [qualitative Analyse von TA] <i>aber nur SERP-Snippets, d. h. Suchphase</i> • Dissen et al., 2022 [qualitative Analyse von Antworten] <i>Glaubwürdigkeit wird als Augenscheinvalidität operationalisiert;</i>
Bekanntheit	• T. Hong, 2006 [Inhaltsanalyse der als glaubwürdig bewerteten Webseiten und Fragebogen] • Denner et al. 2016 [Experiment] <i>jedoch ohne Ergebnis</i> • Barzilai et al., 2020 [Experiment]	• Hilligoss & Rieh, 2008 [qualitative Analyse von Tagebucheinträgen und Interviews] • Fogg et al., 2003 [qualitative Analyse von Kommentaren]
Reputation	• McKnight & Kacmar, 2007 [Experiment]; • Fogg, Marshall, Lakari, et al., 2001 [Umfrage] • Liu & Huang, 2005 [Umfrage] • Liu, 2004 [Umfrage] • Haas & Unkel, 2017 [Beobachtungsstudie]	• Wathen & Burkell, 2002 [Review] • Walraven et al. 2009 [qualitative Analyse v. TA und Gruppendiskussion]; • Fogg et al., 2003 [qualitative Analyse von Kommentaren] • Rieh 2002 [qualitative Analyse von TA, Logs und Interviews] <i>aber Informationsqualität, nicht Glaubwürdigkeit</i>

Angabe der Autorenschaft	• Unkel, 2019 [Experiment] <i>wird aber mit Bekanntheit operationalisiert; Suchresultate</i> • Fogg, Marshall, Lakari, et al., 2001 [Umfrage] • Metzger, Flanagan & Zwarun, 2003 [Umfrage] <i>Merkmal als AV als Teil der Verifikationsstrategie</i> • Liu & Huang, 2005 [Umfrage] • Liu, 2004 [Umfrage]	• Dissen et al., 2022 [qualitative Analyse von Antworten] <i>Glaubwürdigkeit wird als Augenscheinvalidität operationalisiert;</i> • Wathen & Burkell, 2002 [Review] • Rieh 2002 [qualitative Analyse von TA, Logs und Interviews] <i>aber Informationsqualität, nicht Glaubwürdigkeit</i> • Walraven et al. 2009 [qualitative Analyse v. TA und Gruppendiskussion]; • Fogg et al., 2003 [qualitative Analyse von Kommentaren]; • Dissen et al., 2022 [qualitative Analyse von Antworten] <i>Glaubwürdigkeit wird als Augenscheinvalidität operationalisiert;</i>
Foto des Autors/der Autorin	• T. Hong, 2006 [Inhaltsanalyse der als glaubwürdig bewerteten Webseiten und Fragebogen] • Fogg, Marshall, Kameda et al., 2001 [Umfrage] • Johnson & Wiedenbeck, 2009 [Experiment] • Edwards et al., 2015 [Experiment]; • Liu & Huang, 2005 [Umfrage] • Curry & Stroud, 2021 [Umfrage, Experiment] <i>zusammengefasst unter Transparenz</i>	

Anmerkung.

Kommentare in kursiv stellen Einschränkungen dar

Abkürzung TA = Think-Aloud

Abkürzung SERP = Search Engine Result Page, im Deutschen: Suchergebnisseite

Begriff SERP-Snippet = Ausschnitt der Suchergebnisseite

Die nachfolgenden Unterkapitel geben einen Einblick in die möglichen Merkmale der Webseiten, die für eine Glaubwürdigkeitsbeurteilung herangezogen werden können. Hierzu werden zunächst einführend die jeweiligen Merkmale thematisiert. Weiterhin wird der Zusammenhang mit Glaubwürdigkeit aufgezeigt, indem relevante Heuristiken für das entsprechende Merkmal aufgegriffen werden. Letztlich werden für das Forschungsvorhaben ausgewählte, relevante Forschungen beschreiben, wobei in diesem Zusammenhang die beiden Wegen zur Erfassung der Glaubwürdigkeit betrachtet werden.

2.6.1.1. Aktualität

Einführung und Glaubwürdigkeit. Dass eine Information in einem Online-Medium aktuell ist, könnte in Zusammenhang mit der empfundenen Glaubwürdigkeit des Mediums stehen. Aktualität könnte deshalb zur Glaubwürdigkeit beitragen, da veraltete Informationen auch möglicherweise auf veraltetem Wissen basieren und zwischenzeitlich durch neue Information widerlegt worden sein könnten. Da zudem neue Informationen meistens auf vorherigen (alten) Informationen basieren, ist eine aktuellere Information ggf. von besserer Qualität (und somit glaubwürdiger).

Forschungsstand. Der mögliche Zusammenhang zwischen Aktualität und wahrgenommener Glaubwürdigkeit wird beispielsweise aus dem Werk von Tate (2019) ersichtlich, die in ihrem Handbuch zur Evaluation der Informationsqualität von Online-Informationen u. a. Aktualität thematisieren. Sie empfehlen, dass Informationskonsumenten bezüglich der Aktualität darauf achten sollten, ob das Datum auf dem Material vorhanden ist oder wann es zuletzt aktualisiert wurde (Tate, 2019). Hope (2007) wies im Rahmen ihrer Dissertation mithilfe von Think-Aloud-Interviews nach, dass ihre Studienteilnehmer das Merkmal Aktualität verwenden, um die Glaubwürdigkeit zu beurteilen (Hope, 2007). Walraven et al. (2009) führten mit einer anderen Begrifflichkeit („Topicality“) an, dass die empfundene Glaubwürdigkeit mit der Aktualität der Informationen zusammenhängen könnte. Sie fanden in ihrer qualitativen Studie heraus, dass die Aktualität zwar als wichtig erachtet, von den Probanden in der eigentlichen Evaluation an sich aber nicht verwendet wird (Walraven et al., 2009). Im Rahmen eines Gruppenvergleichs zwischen Studierenden und Nicht-Studierenden und ihrer Befragung zeigte sich, dass innerhalb der Studierendenteilgruppe die Aktualität zur Evaluation von Informationen am häufigsten genannt wird (Metzger, Flanagin & Zwarun, 2003). Hong (2006) nutzte eine Kombination aus qualitativer Inhaltsanalyse von als glaubwürdig bewerteten Webseiten und einen Fragebogen, um herauszufinden, ob u. a. Aktualität die wahrgenommene Glaubwürdigkeit vorhersagen kann. Sie kann bestätigen, dass u. a. Aktualität, Quellenangabe und Autorenangabe die wahrgenommene Glaubwürdigkeit signifikant voraussagen können (T. Hong, 2006).

Zusammenfassend ist daher davon auszugehen, dass Probanden bei der Beurteilung von Online-Quellen auf das Datum, sofern angegeben, achten und es somit notwendig ist, das Merkmal der Aktualität in die Betrachtung zu Glaubwürdigkeitsbeurteilungen mit einfließen zu lassen.

2.6.1.2. Verständlichkeit

Einführung und Glaubwürdigkeit. Wie verständlich ein Text oder eine Information ist, könnte dazu führen, dass eine Information aufgrund der Verständlichkeit als glaubwürdig oder unglaubwürdig erachtet wird. Ursächlich dafür könnte dafür die Verarbeitungsleichtigkeit sein: Ein Stimulus wird aufgrund dieser als vertrauenswürdiger (und somit glaubwürdiger, da Vertrauenswürdigkeit als Bestandteil der Glaubwürdigkeit angesehen wird) eingeschätzt (Reber & Schwarz, 1999, S. 342). Zu beachten ist jedoch, dass Verständlichkeit in diesem Zusammenhang nicht bedeutet, wie gut eine Person einen Text verstanden hat. Die Rezipienten-Seite wird außer Acht gelassen. In dieser Arbeit meint Verständlichkeit, welche textlich-formalen Merkmale ein Text bereitstellt, die für eine niedrige/hohe Verständlichkeit sprechen können und nicht, wie und ob die Informationen beim Rezipienten als verständlich ankommen. Es wird also lediglich die Seite des Senders einer Information berücksichtigt.

Forschungsstand. Lev-Ari und Keysar (2010) untersuchten in einem Experiment, ob Personen mit Akzent als weniger glaubwürdig wahrgenommen werden im Vergleich zu Muttersprachlern, weil sie möglicherweise weniger verständlich sind. Sie fanden heraus, dass die Probanden die Verständnisschwierigkeit fälschlicherweise dem Wahrheitsgehalt der Aussage zuschreiben und die Person mit Akzent als weniger glaubwürdig beurteilen (Lev-Ari & Keysar, 2010). Sie bestätigen somit ihre Vermutung, dass Personen mit Akzent als weniger glaubwürdig wahrgenommen werden im Vergleich zu Muttersprachler, weil die Unverständlichkeit die Verarbeitungsleistung reduziert (Lev-Ari & Keysar, 2010). De Meo (2012) fand heraus, dass die Glaubwürdigkeit generell mit der Verständlichkeit einer Nachricht verbunden ist. Denn die Verständlichkeit der Person, die die Nachricht überbringt, wird überwiegend aufgrund von sprachlichen Gegebenheiten (z. B. unflüssige Aussprache, falsche Laute oder tonale Variationen) bewertet (de Meo, 2012). Metzger, Flanagin und Zwarun (2003) fanden heraus, dass die Stichprobe der Studierenden die Verständlichkeit der Information zur Evaluation dieser überdurchschnittlich oft heranziehen. Scharrer et al. (2013) untersuchten, inwieweit Laien einen medizinischen Text als glaubwürdig einschätzen und von dem Inhalt überzeugt werden, wobei die Verständlichkeit des Textes manipuliert wurde. Sie zeigten, dass verständlichere im Gegensatz zu weniger verständlichen Informationen von Laien als glaubwürdiger bewertet werden (Scharrer et al., 2013). Scharrer et al. (2021, 2022) bestätigten in zwei weiteren Experimenten ihre früheren Ergebnisse und zeigten erneut für Laien auf, dass zumindest die wahrgenommene Autoren-Glaubwürdigkeit (2021, S. 7, 9) und Informations-Glaubwürdigkeit (2022, S. 10) mit zunehmender Verständlichkeit der Informationen steigt.

Für die vorliegende Forschung wird daher angenommen, dass die Verständlichkeit eine Rolle bei der Bewertung der Glaubwürdigkeit spielt und verständlichere Informationen bzw. Nachrichten als glaubwürdiger erachtet werden.

2.6.1.3. Transparenz

Einführung und Glaubwürdigkeit. Die Problematik um eine Transparenz ist im Bereich des Journalismus längst nichts Neues und eine genaue Definition eben dieser ist nicht uneindeutig möglich (Curry & Stroud, 2021). Für vorliegendes Forschungsvorhaben wird sich aber an der gängigsten orientiert, was von Karlsson (2010) beschrieben wird als „Disclosure Transparency“. Sie beschreibt die Bereitschaft eines Journalisten oder Informationsbereitstellers, aufzuzeigen, wie Informationen erstellt bzw. beschafft worden sind (Karlsson, 2010). Zu wissen, wie Informationen zu Stande kommen durch beispielsweise Beschreibung der Methodik und Datenerhebung (bei Primärquellen) oder wie der Prozess zur Überprüfung stattgefunden hat (bei Sekundärquellen), könnte ebenfalls die wahrgenommene Glaubwürdigkeit beeinflussen. Dies könnte mit einer Kombination aus Expertenheuristik sowie Reaktanz erklärt werden.

Zunächst könnte die freiwillige Zusatzinformation, wie Daten erhoben oder Prozesse zur Informationsauswahl überprüft wurden, die Reaktanz von Menschen reduzieren. Das Bereitstellen von transparenten Informationen könnte vermeiden, dass Menschen sich gezielt manipuliert fühlen und dadurch negative Emotionen empfinden. Transparent zu „scheinen“, wirkt möglicherweise ehrlich und authentisch, weshalb Menschen transparente Informationen weniger als eine Manipulation empfinden und positive Qualitätsmerkmale, wie Glaubwürdigkeit, attribuieren könnten. Weiterhin könnte die Glaubwürdigkeitsbeurteilung auch deshalb positiv ausfallen, weil Medien, die transparente Informationen liefern, eine höhere Expertise zugeschrieben bekommen. Wenn sie offen darlegen, welche Methodiken zur Datenerhebung genutzt wurden, zeigen sie indirekt auch, wie sie arbeiten. Sie geben folglich etwas zur Arbeitsweise und zum Know-how Preis. Insofern könnten beide Entscheidungsregeln dafür verantwortlich sein, dass transparente Informationen als glaubwürdiger erachtet werden.

Forschungsstand. Der Zusammenhang zwischen Transparenz und Glaubwürdigkeit wird z. B. in der qualitativen Studie von Hilligoss und Rieh (2008) berichtet. Sie untersuchten Tagebucheinträge über tägliche Informationssuchaktivitäten und analysierten diese qualitativ in Zusammenhang mit geführten Interviews der Teilnehmenden (Hilligoss & Rieh, 2008). Anhand der Ergebnisse wird deutlich, dass die empfundene Glaubwürdigkeit mit dem Wissen darüber, dass beispielsweise ein Artikel einen Peer-Review-Prozess durchlaufen hat, positiv zusammenhängt (Hilligoss & Rieh, 2008). Thomm und Bromme (2012) zeigten experimentell, dass ein Text, der unter anderem transparent die statistische Methode der Datenerhebung beschrieb, als wissenschaftlicher empfunden wird und ihm dadurch eine höhere Glaubwürdigkeit zuteilwird. Ebenfalls experimentell erprobt haben Curry und Stroud (2021), wie verschiedene Transparenz-Elemente auf die wahrgenommene Glaubwürdigkeit von einer Webseite wirkten. Den Probanden wurden dabei entweder Artikel mit allen fünf oder ohne diese Elemente gezeigt. Sie zeigten, dass die Artikel mit den Transparenz-Elementen als glaubwürdiger eingeschätzt werden im Vergleich zu den Artikeln ohne (Curry & Stroud, 2021, S. 910 f.). Diese Erkenntnisse sind für vorliegendes Forschungsvorhaben jedoch mit Vorsicht zu genießen. Denn die Forschenden fassen u. a. folgende Elemente unter dem Merkmal Transparenz zusammen: Foto des Autors, Angabe der Quellen via Hyperlink sowie Hintergrund zum Datenerhebungsprozess. Lediglich das Letztgenannte ist laut der herausgearbeiteten Definition für vorliegendes Forschungsvorhaben als Transparenz zu bezeichnen. Die Angabe eines Fotos des Autors oder der Quellen werden als gesonderte Merkmale betrachtet. Insofern spricht das kumulierte Ergebnis der Autoren zunächst dafür, dass transparentere Informationen zu mehr Glaubwürdigkeit führen. Dennoch muss dies hinsichtlich der einzelnen Elemente differenziert betrachtet werden.

Zusammenfassend lässt sich sagen, dass die Angabe zu Datengenerierungs- oder erhebungsprozessen eine Auswirkung auf die wahrgenommene Glaubwürdigkeit haben könnte.

2.6.1.4. Quellenherkunft

Einführung und Glaubwürdigkeit. Die Angabe, ob es sich um Primär- oder Sekundärdaten handelt, könnte ebenfalls zu unterschiedlichen Glaubwürdigkeitswahrnehmungen führen. Erklärt werden könnte diese Assoziation mit der Expertenheuristik. Denn wenn Daten selbst erhoben wurden, spricht das für ausreichende Fähigkeiten auf diesem Fachgebiet, sodass von einem Experten ausgegangen werden kann. Dies könnte dann wiederum positiv für die Glaubwürdigkeitsbeurteilung ausfallen. Wenn hingegen Sekundärquellen angegeben werden, senkt dies möglicherweise die wahrgenommene Expertise und Fachkundigkeit der Webseite respektive der Informationen, sodass ihnen weniger Glaubwürdigkeit zugeschrieben werden könnte.

Forschungsstand. Hilligoss und Rieh (2008) zeigten durch ihre Interviewanalysen, dass die Probanden bei der Suche nach glaubwürdigen Informationen auf die Quellenherkunft achten und Primärquellen als glaubwürdiger beurteilen im Vergleich zu Sekundärquellen. Walraven et al. (2009) stellten ebenfalls fest, dass die Probanden bei der Beurteilung der Informationen auf die Quellenherkunft achten und Primärquellen eine größere Glaubwürdigkeit zuschreiben. In der eigentlichen, praktischen Durchführung der Beurteilung wurde dieses Merkmal allerdings nicht häufig verwendet, was in dieser Forschung aber als generelles Problem thematisiert wird (Walraven et al., 2009, S. 241, 244).

Insofern ist davon auszugehen, dass die Herkunft der Quelle (primär vs. sekundär) ebenfalls eine Rolle in der Wahrnehmung der Glaubwürdigkeit spielen könnte.

2.6.1.5. Angabe von Quellen

Wenn ersichtlich ist, dass die Informationen, über die berichtet werden, klar einer Quelle zuzuordnen ist – sei es durch direkte Zitation des Sprechers oder Interviewpartners oder durch den Verweis zu weiteren Quellen (K. A. Johnson & Wiedenbeck, 2009) – könnte dies ebenso die wahrgenommene Glaubwürdigkeit begünstigen.

Abgrenzung zu Quellenangabe. Es ist an dieser Stelle notwendig, das Merkmal der Transparenz von dem der Angabe von Quellen abzugrenzen. Bei der Transparenz ist es nicht relevant, dass Daten, Referenzen, Links etc. vorhanden sind, sondern wie diese im Vorhinein zu Stande kamen; sprich welchen Prozess sie durchlaufen haben oder mussten, um als entsprechenden Beleg für eine Information angeführt zu werden. Bei der Quellenangabe geht es hingegen lediglich um die schiere Existenz von solchen Quellenangaben und nicht, wie sie zustande kamen.

Glaubwürdigkeit. Der Zusammenhang mit Glaubwürdigkeit könnte mit der Expertenheuristik begründet werden. Die Angabe von Evidenzen und Belegen könnte dafürsprechen, dass es sich bei dem Medium oder dem Verantwortlichen für den Inhalt um eine Person handelt, die ausreichend (Fach-)Kenntnisse bezüglich korrekter Zitation und Quellenverweise aufweist. Auch wenn es sich in diesem Fall nicht um die klassische Expertise bezüglich inhaltlicher Themen handelt, sondern vielmehr um eine methodische Kenntnis, könnte dies dennoch dem Leser vermitteln, dass das Medium (respektive das Team des Mediums in Summe) Expertise aufweist und diesem daher eine hohe Glaubwürdigkeit zugerechnet werden kann.

Forschungsstand. Dieser mögliche Zusammenhang ist unter anderem erkennbar an der Studie von Walraven et al. (2009), die herausgearbeitet haben, ob und inwieweit Studierende bei der Beurteilung der Glaubwürdigkeit darauf achteten, ob „Referenzen“ (S. 242) zu erkennen sind oder nicht. Die Studierenden geben in einer ersten Runde an, dass sie auf Referenzen achten würden. In der eigentlichen Aufgabenbearbeitung wird allerdings ersichtlich, dass sie es in der Praxis nicht zur Beurteilung der Glaubwürdigkeit anwenden. Karlsson und Clerwall (2018) kombinierten Fokusgruppeninterviews mit einer Umfrage und zeigten, dass die Angabe von Quellen in Form von erkennbaren Hyperlinks das zur Beurteilung der Glaubwürdigkeit am meisten wertgeschätzte Tool für die Teilnehmer der Interviews ist. Fogg, Marshall, Lakari, et al. (2001) ermittelten in einer Umfrage, welche Merkmale von einer Webseite aus Sicht der Benutzer für oder gegen die Glaubwürdigkeit sprechen. Aus ihren zuvor geführten Interviews und der anschließenden Faktorenanalyse der gefundenen Items wurden verschiedene Skalen für diese Umfrage abgeleitet. Das Merkmal der Quellenangabe zeigt sich in zwei verschiedenen, aber inhaltlich ähnlichen Items von zwei Subskalen (Expertise und Vertrauenswürdigkeit). Ihre Ergebnisse der Umfrage verdeutlichten, dass die Angabe von Quellen bzw. Zitationen von ihren Probanden überdurchschnittlich oft als Indikator für eine positive Glaubwürdigkeitsbeurteilung genannt werden (Fogg, Marshall, Lakari, et al., 2001). Wie bereits beim Merkmal Transparenz erwähnt, ist die Forschung von Thomm und Bromme (2012) auch hier anzuführen. Denn sie zeigten, dass ein Text, der die Angabe von Zitationen beinhaltete, als wissenschaftlicher empfunden und ihm dadurch eine höhere Glaubwürdigkeit zuteilwird (S. 194, 205). Henke et al. (2019) fanden ebenfalls experimentell heraus, dass die Angabe von wissenschaftlichen Quellen die wahrgenommene Glaubwürdigkeit erhöht. Obwohl ihre Manipulation explizit nur auf „wissenschaftliche“ Quellen abzielte, könnte das Ergebnis dennoch verallgemeinert werden und auf allgemeine Quellen bzw. Referenzen oder Zitationen angewendet werden. Bei der Bewertung der Glaubwürdigkeit einer vorgegebenen Seite, die entweder Hyperlinks und somit Verlinkungen zu weiteren Quellen i. S. der Angabe von Referenzen enthielt oder nicht, führt die Angabe der Hyperlinks in zwei unabhängigen Studien zu einer größeren wahrgenommenen Glaubwürdigkeit (Borah, 2014; K. A. Johnson & Wiedenbeck, 2009).

In Anbetracht der hier angeführten Forschungen ist daher davon auszugehen, dass die Angabe von Quellen i. S. v. Zitationen oder Verlinkungen zu anderen genutzten Quellen in Zusammenhang mit der Glaubwürdigkeitsbeurteilung steht.

2.6.1.6. Struktur der Information bzw. Webseite

Einführung und Glaubwürdigkeit. Unübersichtliche oder wenig strukturierte Informationen oder ganze Webseiten sind für Leser eine Herausforderung, da sie das Gehirn vor einen zusätzlichen kognitiven Aufwand stellen. Die Informationen müssen erst strukturiert werden und sind folglich nicht auf Anhieb bzw. leichter zu verarbeiten. Aufgrund der Heuristik der Verarbeitungsleichtigkeit ist bekannt, dass Menschen Informationen besser abspeichern können und positiver bewerten (Graf et al., 2017, S. 394), wenn sie die Informationen leichter verarbeiten können (Schooler & Hertwig, 2005, S. 612). Wenn Informationen gut präsentiert sind, müssen die Informationen nicht zwingend neu strukturell angeordnet werden, um sie im Gehirn abzuspeichern. In der Information bereits vorhandene Zusammenhänge und Strukturen können in das Gehirn übernommen werden, ohne viel kognitiven Aufwand zu betreiben. Das führt zu einer leichteren Informationsverarbeitung, die wiederum in einer positiven Bewertung, welche beispielsweise Glaubwürdigkeit darstellen könnte, resultiert. Insofern ist es denkbar, dass strukturierte Informationen oder ganze Webseiten als glaubwürdiger wahrgenommen werden.

Forschungsstand. Bereits in den 60er-Jahren wurde in zwei Studien herausgefunden, dass mangelnde Organisation respektive Strukturiertheit dazu führt, dass Studierende einen unorganisierten Redner nach seiner Rede weniger wertschätzen (Sharp & McClung, 1966; McCroskey & Mehrley, 1969). Diese Unterscheidung trifft auf den organisierten Redner nicht zu (Sharp & McClung, 1966, S. 183). Auch eine reduzierte Glaubwürdigkeitszuschreibung kann zumindest für einen ihrer Faktoren der Glaubwürdigkeit aus einer weiteren Studie in Zusammenhang mit Unstrukturiertheit nachgewiesen werden (McCroskey & Mehrley, 1969, S. 19, 21). Der unstrukturierte Stimulus beinhaltete dabei z. B. die zufällige Anordnung von Sätzen (McCroskey & Mehrley, 1969, S. 16). Auch wenn diese beiden Forschungen zeitlich betrachtet sehr lange zurück liegen und auch in einem anderen Kontext (Rede statt Bewertung von Online-Informationen) stattfanden, haben sie dennoch einen Anspruch auf Erwähnung, da sie grundsätzlich die Verbindung zwischen Glaubwürdigkeit und Struktur aufzeigen. Im Online-Kontext ist es daher auch denkbar, dass die Struktur bzw. Unstrukturiertheit von Informationen oder der ganzen Webseite einen Einfluss auf die wahrgenommene Glaubwürdigkeit hat. Ebenso durch qualitative Analyse von Nutzerkommentaren konnten Forschende (Fogg et al., 2003) Erkenntnisse darüber gewinnen, auf was Nutzer bei der Bewertung der Glaubwürdigkeit von Webseiten achten. Sie zeigten, dass 28.5 % der Kommentare (und somit am zweithäufigsten) Bezüge zur Struktur der Informationen enthalten, was die Wichtigkeit des Merkmals

für die Glaubwürdigkeitsbeurteilung betont. Hierbei kann nochmal unterschieden werden, ob die Information rein inhaltlich gut strukturiert ist, beispielsweise durch Absätze, Überschriften oder Hervorhebungen. Es kann aber auch festgehalten werden, ob die Webseite per se eine gute Struktur aufweist i. S. v. Seitenaufbau oder Layout. Da dies in den bisherigen Literaturen nicht näher spezifiziert wurde, wird grundsätzlich von beidem ein Einfluss auf die Glaubwürdigkeit angenommen.

2.6.1.7. Faktische Korrektheit

Einführung und Glaubwürdigkeit. Dass eine Information faktisch korrekt bzw. akkurat ist i. S. v. richtig, könnte ebenfalls bei der Glaubwürdigkeitsbeurteilung eine Rolle spielen. Dies könnte mit der Expertenheuristik erklärt werden. Wenn ein Leser erkennt, dass eine Information faktisch korrekt ist, könnte er davon ausgehen, dass es sich bei dem Autor um eine Person handeln muss, die ein gewisses Fachwissen oder zumindest Recherchewissen, und somit ein hohes Maß an Expertise aufweist. Menschen, die als Experten wahrgenommen werden, erhalten eine höhere Glaubwürdigkeitszuschreibung (Sundar, 2008), weshalb es daher denkbar ist, dass die faktische Korrektheit als ein Hinweismerkmal für die Expertise in Zusammenhang mit der Glaubwürdigkeit steht.

Forschungsstand. Wathen und Burkell (2002) fassten relevante Literatur zur Glaubwürdigkeitswahrnehmung von Informationen im Web zusammen. Sie schlagen ein konzeptionelles Modell vor, in dem u. a. auch die Akkuratheit von Informationen beurteilt wird. Hope (2007) fand heraus, dass Probanden beim Lösen von Informationsproblemen, auf die sie innerhalb der Vorbereitung und Nachbereitung ihrer Studieninhalte stoßen, auf die faktische Korrektheit der Information achten. Fogg et al. (2003) stellten fest, dass 14.3 % der Kommentare zur Glaubwürdigkeitsbeurteilung die faktische Korrektheit thematisieren. Experimentell ermittelt wurde der Unterschied auf Informationsvertrauen zwischen einem Text, der eine hohe faktische Korrektheit aufwies und einem Text, der bewusst faktische Fehler enthielt (Lucassen & Schraagen, 2011). Obwohl es sich bei der abhängigen Variable nicht um Glaubwürdigkeit handelte, ist das Resultat dennoch von Interesse, da Vertrauen und Glaubwürdigkeit, wie in Kapitel 2.1 erläutert, in Zusammenhang stehen. Die Forschenden fanden heraus, dass in Abhängigkeit der Expertise die faktische Korrektheit eine Rolle spielt, wobei Experten bei reduzierter faktischer Korrektheit weniger Vertrauen in Informationen haben und es bei Novizen unabhängig von der faktischen Korrektheit keine Auswirkung auf ihr Vertrauen in Informationen gibt (Lucassen & Schraagen, 2011). Eine aktuellere Forschung (Middaugh, 2018) hierzu konnte weiterhin aufzeigen, dass bei bewusster Erinnerung an die Bedeutung des Merkmals faktische Korrektheit dieses zwar in einer Online-Suchphase in Kommentaren zur Glaubwürdigkeit

angeführt wird, dieses jedoch im späteren Verlauf bei weiteren Schritten im Informationsproblemlösungsprozess (Evaluation und Weitergabe an Dritte) zunehmend in Vergessenheit gerät.

Dennoch ist in Summe davon auszugehen, dass die faktische Korrektheit eine Rolle bei der Glaubwürdigkeitsbeurteilung spielen könnte.

2.6.1.8. Statistiken

Einführung und Verbindung zur Glaubwürdigkeit. Statistische Angabe in Form von Zahlen könnten ebenfalls in Zusammenhang mit Glaubwürdigkeit gebracht werden. Dies könnte mit der Expertenheuristik erklärt werden: Laut Artz und Tybout (1999) wird ein Experte als überzeugender wahrgenommen, wenn seine Nachricht Zahlen enthält. Wenn die Nachricht hingegen lediglich verbale Ausdrücke enthält, wird kein Zusammenhang mit Expertise festgestellt. Insofern könnte dem Autor aufgrund numerischer Ausdrücke in einer Nachricht eine gewisse Expertise zugeordnet werden, die im Weiteren auch mit positiven Attributen wie Glaubwürdigkeit assoziiert werden könnten.

Forschungsstand. Belege für diesen Zusammenhang liefern unter anderem Arbeiten aus der Persuasionsforschung, die überwiegend statistische mit anekdotischen Angaben verglichen haben. Anekdotische Angaben sind z. B. Fallstudien oder Beispiele (Allen & Preiss, 1997). Wie bereits ausgeführt, hängen Persuasion und Glaubwürdigkeit miteinander zusammen (Wathen & Burkell, 2002), weshalb die nachstehenden Forschungsbeispiele hier angeführt werden können. Hornikx (2005) schlussfolgerte in seinem Literaturreview, dass statistische Angaben überzeugender sind im Vergleich zu anekdotischen Angaben. Zur gleichen Erkenntnis kommen Allen und Preiss (1997) dank ihrer Meta-Analyse. Eine neuere Meta-Analyse verdeutlichte unter Berücksichtigung der emotionalen Involviertheit, dass statistische Angaben bei niedriger emotionaler Involviertheit überzeugender sind als anekdotische Angaben (Freling et al., 2020)

Explizit mit Glaubwürdigkeit in Zusammenhang gebracht wurde die Angabe von Statistiken auch in nachfolgenden Arbeiten: Koetsenruijter (2011) konnte für analoge Textmedien aufzeigen, dass numerische Angaben positiv zur wahrgenommenen Glaubwürdigkeit der Artikel beitragen. Im Online-Bereich wurde in einem Experiment die wahrgenommene Nachrichtenglaubwürdigkeit als höher empfunden, wenn die entsprechende Nachricht mit numerischen Belegen i. S. v. Statistiken im Vergleich zu textlichen Darstellungen angereichert waren (Henke et al., 2019). Weiterhin wurde herausgefunden, dass neben weiteren Nachrichtenmerkmalen auch die Angabe von Statistiken die wahrgenommene Glaubwürdigkeit signifikant voraussagen können (T. Hong, 2006). Es ist daher davon auszugehen, dass die Angabe von Statistiken in Form von Zahlen zu einer höheren Glaubwürdigkeit beitragen könnten.

2.6.1.9. Vollständigkeit

Einführung und Glaubwürdigkeit. Dass Informationen vollständig vorliegen, wird ebenfalls als Merkmal für Glaubwürdigkeitsbeurteilung angeführt (Hope, 2007, S. 84): Laut ihrer Definition bedeutet Vollständigkeit, dass es einer Information nicht an signifikanten Informationen fehlt. Zu erklären wäre die Assoziation vollständiger Informationen mit höherer Glaubwürdigkeit mithilfe der Expertenheuristik. Medien, die vollständige Informationen liefern, könnten eine höhere Expertise zugeschrieben werden, da sie alle relevanten Informationen bedacht, recherchiert und präsentiert haben. Da die Bereitsteller der Informationen dadurch als Experte wahrgenommen werden und die präsentierten, vollständigen Informationen Zeuge einer guten Arbeit und Kompetenz der Informationsbereitsteller sind, könnten ihnen weiterhin positive Attribute, wie Glaubwürdigkeit, zugerechnet werden.

Forschungsstand. Hope (2007) stellte mithilfe der verbalen Protokolle ihrer Studierenden fest, dass die Vollständigkeit zur Beurteilung der Glaubwürdigkeit herangezogen wird. Metzger et al. (2003) fanden heraus, dass ihre Teilstichprobe der Studierenden überdurchschnittlich oft die Vollständigkeit zur Evaluation von Informationen heranzieht (Metzger, Flanagin, & Zwarun, 2003). Obwohl die Vergleichbarkeit dieser Studie mit der hier geplanten geringer ist, da das Merkmal den Autoren als Operationalisierung der Evaluationsstrategie dient und somit nicht direkt zur Glaubwürdigkeitsbeurteilung herangezogen wird, könnte dennoch ein Zusammenhang nachgesagt werden. Es wird daher davon ausgegangen, dass Vollständigkeit bei der Glaubwürdigkeitsbeurteilung eine Rolle spielen könnte.

2.6.1.10. Sprachstil

Einführung und Glaubwürdigkeit. Wie ein Text verfasst ist, im Sinne seines Sprachstils und dem daraus resultierenden Klima, könnte ebenfalls die Glaubwürdigkeit beeinflussen. Der Sprachstil ist als Begriff *Sentiment* (französisch für Stimmung, Gefühl) innerhalb der Linguistik und mit Schnittstelle zur Computerlinguistik zuzuordnen (Taboada, 2016). Eine mögliche Erklärung, dass ein bestimmter Sprachstil zu einer schlechteren oder besseren Glaubwürdigkeitsbeurteilung führt, liefert die *Language Expectancy Theory* (Burgoon et al., 1975). Sie besagt unter anderem, dass Menschen eine Erwartungshaltung gegenüber dem Sprachverhalten eines Dritten entwickeln und das Erfüllen oder Verletzen dieser Erwartungshaltung die Persuasion positiv oder negativ beeinflussen (Burgoon et al., 2002). Es ist daher vorstellbar, dass Menschen eine Erwartung gegenüber dem Sprachstil eines Online-Informations-Bereitstellers haben. Wenn diese Erwartung erfüllt oder verletzt wird, hat das Auswirkungen auf die Persuasion und ggf. auf die Glaubwürdigkeitsbeurteilung.

Forschungsstand. Die Verbindung zur Glaubwürdigkeit fanden beispielsweise König und Jucks (2019a; 2019b) in zwei verschiedenen Experimenten heraus: Zum einen manipulierten

sie eine wissenschaftliche Debatte als entweder aggressiv (durch Hinzufügen von negativen, beleidigenden und zynischen Kommentaren) oder neutral, wobei einmal der aggressivere Autor als weniger vertrauenswürdig (und somit weniger glaubwürdig) und weiterhin seine Information als weniger glaubwürdig eingeschätzt wird (König & Jucks, 2019a). In einem zweiten Experiment manipulierten sie den Sprachstil eines Autors bei einem Forenpost als entweder enthusiastisch (durch zusätzliche Sätze wie „das begeistert mich total“ oder „...was ich sehr interessant finde“) oder neutral (König & Jucks, 2019b). Sie schlussfolgerten, dass der enthusiastischere Sprachstil den Autor als manipulativer, wenig bewandert und seine Information als weniger glaubwürdig wahrnehmen lässt. Eine ähnliche Forschung betrieben Schaewitz et al. (2020), welche unter anderem einen Text hinsichtlich Sensationslust bzw. Aufsehererregung manipulierten (durch Hinzufügen von z. B. „unglaublich, extrem, dramatisch“) und den Einfluss auf die Glaubwürdigkeit untersuchten. Sie fanden jedoch heraus, dass der sensationelle Text keinen Einfluss auf die Glaubwürdigkeit hat.

Es zeigt sich dennoch im Gros der Forschungen, dass der Sprachstil einen Einfluss auf die wahrgenommene Glaubwürdigkeit eines Autors oder seiner Information (somit indirekt) haben könnte.

2.6.1.11. Wissenschaftlichkeit

Einführung und Glaubwürdigkeit. Ein weiteres Merkmal, welches ebenfalls mithilfe von linguistischen Markern und als Teil der Semantik erfasst werden könnte, ist die Verwendung von Fachbegriffen bzw. Fachsprache, welche hier als Wissenschaftlichkeit bezeichnet wird. Warum Wissenschaftlichkeit in Zusammenhang mit wahrgenommener Glaubwürdigkeit steht, könnte mit der Expertenheuristik erklärt werden: Wird eine Information oder eine Person als wissenschaftlich eingeschätzt, könnte diesen eine gewisse Expertise in dem Fachgebiet zugeschrieben werden. Aufgrund der wahrgenommenen Expertise bzw. der Person als Experte werden weitere positive Attribute, wie Glaubwürdigkeit, zugeordnet (Sundar, 2008).

Abgrenzung zu Verständlichkeit und Forschungsstand. Wissenschaftlichkeit muss konzeptionell von Verständlichkeit abgegrenzt werden, denn Forschende sehen einen Zusammenhang der Wissenschaftlichkeit mit dem Merkmal der wahrgenommenen Verständlichkeit: Joiner et al. (2002) untersuchten die Verwendung von technischer Sprache, also Fachtermini, und die Verständlichkeit eines gesprochenen Textes. Sie manipulierten dabei die Häufigkeit von Fachtermini in einem Video und fanden heraus, dass die Fachsprache (und somit Wissenschaftlichkeit) die Verständlichkeit verringert und dies zu einer geringeren Expertise und Zuverlässigkeit (folglich Glaubwürdigkeit) führt. Vorliegende Arbeit wird sich allerdings nicht diesem Weg anschließen, denn eine jüngste Forschung von Jonas et al. (2024) zeigt, dass die beiden Merkmale nicht zwingend miteinander verbunden sind: Sie manipulierten die Verständlichkeit

und Wissenschaftlichkeit von zusammengefassten Texten aus unterschiedlichen Themenbereichen. Sie fanden heraus, dass die Verständlichkeit keinen, die Wissenschaftlichkeit jedoch einen Effekt auf die wahrgenommene Zuverlässigkeit (und somit Glaubwürdigkeit) hat und dass es entgegen ihrer Hypothese auch keinen Interaktionseffekt gibt. Es zeigt sich also, dass Verständlichkeit und Wissenschaftlichkeit getrennt betrachtet werden müssen. Eine weitere Forschung untersuchte die Verwendung der Fachtermini und somit Wissenschaftlichkeit: Thon und Jucks (2017) manipulierten einen Text entweder mit Fachtermini oder alltäglicher Sprache und erfragten u. a. die Glaubwürdigkeit der Autoren. Sie konnten signifikante Effekte nachweisen und bestätigen, dass sich die Verwendung von Fachtermini negativ auf die Glaubwürdigkeit der Autoren und ihren Informationen auswirkt.

In vorliegender Forschung wird daher von einem Zusammenhang zwischen Wissenschaftlichkeit und Glaubwürdigkeit ausgegangen.

2.6.1.12. Argumentqualität

Einführung und Glaubwürdigkeit. Argumentqualität wird bereits im Rahmen des ELM (Petty & Cacioppo, 1984) eine entscheidende Rolle zugeschrieben. Die Qualität eines Argumentes bedeutet dabei, dass ein Argument für eine Behauptung entweder als stark oder schwach¹⁵ wahrgenommen wird (Petty & Cacioppo, 1986, S. 188, 1996, S. 259 ff.). Dass es als stark oder schwach wahrgenommen bzw. beurteilt wird, erfordert eine tiefergehende, systematische Auseinandersetzung mit dem Argument bzw. der Nachricht und kann nicht mithilfe von Heuristiken beurteilt werden. Insofern ist der Zusammenhang mit der Wahrnehmung der Glaubwürdigkeit und der Argumentqualität nur durch eine systematische, elaborierte Auseinandersetzung und nicht durch Heuristiken, wie bei den übrigen Merkmalen, zu erklären.

Forschungsstand. Stoltenberg und Davis (1988) fanden beispielsweise heraus, dass es Interaktionseffekte von Argumentqualität und Quellenglaubwürdigkeit bei der Beurteilung von Printmedien gibt: Ihre Teilnehmer sind eher bereit, einer Empfehlung zu folgen, wenn starke Argumente von einer sehr glaubwürdigen Quelle vorgebracht werden, und am wenigsten bereit, zu handeln, wenn schwache Argumente von einer sehr glaubwürdigen Quelle vorgetragen werden (Stoltenberg & Davis, 1988). In einem Experiment im Konsumentenbereich evaluierten Probanden Kundenrezensionen, wobei die Forschenden die Argumentqualität manipulierten (Cheung et al., 2012). Sie konnten ihre Hypothese, dass Argumentqualität einen positiven Effekt auf die Glaubwürdigkeit einer Rezension hat, bestätigen (Cheung et al., 2012). Es wird

¹⁵ Stark und schwach sind beispielhaft genannt und stellen die Extrema der Skalierung dar. Abstufungen sind möglich, werden hier aufgrund der Kürze nicht explizit aufgeführt.

daher unterstellt, dass die Argumentqualität einen Einfluss auf die Glaubwürdigkeitsbeurteilung hat

2.6.1.13. Narrative

Einführung und Glaubwürdigkeit. Narrative können bei der Glaubwürdigkeitsbeurteilung von Informationen eine Rolle spielen. Ein Narrativ ist eine Erzählung bzw. eine Art Orientierungsrahmen für das menschliche Denken (Clementson, 2020, S. 1), und trägt zum Verständnis der Welt bei (Fisher, 1984). Eine Information oder ein Text ist nie gänzlich¹⁶ frei von Narrativen (Banerjee et al., 2020, S. 3) und transportiert daher auch immer automatisch eine Botschaft mit, die einen Einfluss auf die wahrgenommene Glaubwürdigkeit haben könnte. Warum narrative (vs. weniger narrative) Texte mit einer größeren wahrgenommenen Glaubwürdigkeit assoziiert werden, könnte mit der Verarbeitungsleichtigkeit zusammenhängen. Erzählungen werden laut Forschenden leichter im Gehirn verarbeitet, was im Umkehrschluss zu einer größeren Persuasion führt (Bullock et al., 2021). Es ist daher denkbar, dass hieraus auch weitere positive Effekte die Folge sind, wie eine höhere zugeschriebene Glaubwürdigkeit aufgrund der Verarbeitungsleichtigkeit des Narrativs. Es ist aber auch denkbar, dass ein Narrativ die wahrgenommene Glaubwürdigkeit reduzieren könnte, was mit einer Kombination aus Reaktanz und Expertenheuristik zu erklären ist. Wenn eine Person das Gefühl hat, durch eine Geschichte beeinflusst zu werden und gleichzeitig der beeinflussenden Person aufgrund der Erzählung der Geschichte statt beispielsweise Fakten eine geringere empfundene Expertise zuschreibt, könnte eine geringere Glaubwürdigkeitswahrnehmung resultieren. Denn die Wahrnehmung als Experte sowie die Abwesenheit einer bewussten Beeinflussung (und somit Reaktanz) kann zu einer positiven Glaubwürdigkeitsbeurteilung beitragen. Umgekehrt könnte aber die Wahrnehmung einer gezielten Beeinflussung sowie die Abwesenheit einer Expertise negativ für die Glaubwürdigkeitsbeurteilung ausfallen. Ob narrative oder weniger/nicht-narrative Informationen in positiven Zusammenhang mit der Glaubwürdigkeit stehen, ist in der Forschung umstritten (Zebregs et al., 2015). Dennoch werden einige aktuellere Forschungen exemplarisch angeführt, welche zumindest einen Zusammenhang (egal ob positiv oder negativ) untersuchten.

Forschungsstand. Clementson (2020) ermittelte u. a., wie sich narrative oder nicht-narrative Informationen auf die Beurteilung der Vertrauenswürdigkeit/Zuverlässigkeit (Trustworthiness) eines Pressesprechers auswirkte. In der Version der narrativen Informationen waren Elemente des Storytellings, wie zeitliche Abfolge von Ereignissen, handelnde Figuren und Auflösung eines unglücklichen Ereignisses enthalten. Der Forscher konnte nachweisen, dass nicht-narrative Informationen die Vertrauenswürdigkeit und somit die Glaubwürdigkeit einer Person

¹⁶ Diese Auffassung, dass ein Text nie frei von Narrativen ist, wird von dieser Forschungsarbeit vertreten. Es gibt jedoch auch andere Meinung, wie die der nachstehenden Autoren, welche Narrative im Gegenpol mit „Nicht-narrative Information“ (Clementson, 2020) oder „kein Narrativ“ (Huang & Wang, 2020) beschreiben.

begünstigt (Clementson, 2020). Huang und Wang (2020) untersuchten die Glaubwürdigkeit einer Nachricht, wenn diese ein oder kein Narrativ enthielt. Sie fanden heraus, dass die Nachricht als glaubwürdiger empfunden wird, wenn sie kein Narrativ enthält im Vergleich zu der Nachricht mit Narrativ (Huang & Wang, 2020). Boukes und LaMarre (2021) manipulierten ebenso eine Nachricht im Bereich der Corporate Social Responsibility mit unterschiedlichen Levels an narrativer Reichhaltigkeit. Sie postulierten, dass ein narrativ reichhaltiger Text als weniger glaubwürdig empfunden wird. Ihre Hypothese kann jedoch nicht, wie im Vergleich zur vorherigen Forschung, bestätigt werden (Boukes & LaMarre, 2021). Dennoch ist davon ausgehen, dass Narrative in Zusammenhang mit der empfundenen Glaubwürdigkeit stehen.

2.6.1.14. Spezialisierungsgrad

Einführung und Glaubwürdigkeit. Der Grad der Spezialisierung beschreibt, inwieweit ein Medium in einem gewissen Fachgebiet spezialisiert ist. Hierbei können zwischen wenig spezialisierten, z. B. allgemein-informativen Webseiten, und speziell für das Fachgebiet ausgelegte Webseiten unterschieden werden. Einerseits kann ein hoher Spezialisierungsgrad einer Webseite mit der wahrgenommenen Glaubwürdigkeit in Verbindung stehen, was mit der Expertenheuristik zu erklären ist. Wenn eine Webseite bzw. der Betreiber einer Webseite aufgrund seines Fachwissens, wie bei fachthematisch spezialisierten Webseiten, als Experte für das Thema wahrgenommen wird, könnte ihm mehr Glaubwürdigkeit geschenkt werden (Sundar, 2008). Andererseits könnten allgemeine, nicht-spezifische Webseiten als glaubwürdiger wahrgenommen werden, weil sie Objektivität vermitteln und somit keine Reaktanz verursachen.

Der aktuelle Forschungsstand um den Spezialisierungsgrad ist innerhalb der Informationsverarbeitung und Bildung bislang nicht gut erforscht. Im Konsumentenbereich lässt sich allerdings z. B. eine Forschung anführen: Koh und Sundar (2010) ermittelten experimentell, wie sich der Spezialisierungsgrad der Webseite auf die Vertrauenszuschreibung und Kaufintention der Probanden auswirkte. Sie zeigten, dass der Spezialisierungsgrad einen positiven Einfluss auf das Vertrauen der Probanden hat (Koh & Sundar, 2010). Da Vertrauen in Verbindung mit Glaubwürdigkeit steht (s. Kapitel 2.1), könnte auch hier von einem Zusammenhang mit dem Spezialisierungsgrad einer Webseite auszugehen sein.

2.6.1.15. Design

Einführung und Glaubwürdigkeit. Eine Webseite kann ein unterschiedliches Design haben und dementsprechend als ästhetisch bzw. schön wahrgenommen und folglich bewertet werden. „Design“ in dem Sinne meint beispielsweise Layout, Bilder oder Farbschemata einer Webseite, welche eine Rolle bei der Glaubwürdigkeit spielen könnten. Warum eine Person einer Webseite aufgrund eines als gut empfundenen Designs eine höhere Glaubwürdigkeit zuschreibt, könnte mit dem individuellen ästhetischen Empfinden und dem Halo-Effekt (Forgas & Laham,

2016) zusammenhängen. Personen, die etwas als ästhetisch ansprechend empfinden, weil sie das Design, die Farbgestaltung oder das Layout als schön empfinden, schlussfolgern dann von einer Eigenschaft des zu Bewertenden auf eine andere Eigenschaft, was den Halo-Effekt beschreibt. Sie ziehen möglicherweise irrtümlich die von ihnen empfundene subjektive Ästhetik oder Schönheit einer Webseite heran und übertragen diese auf eine weitere Eigenschaft, die Glaubwürdigkeit. Eine weitere Erklärung liefern Robins und Holmes (2008): Sie benennen und erklären sich diesen Zusammenhang mithilfe des „Amelioration-Effekts“ (Robins & Holmes, 2008, S. 397). Dieser würde bedeuten, dass sich z. B. die wahrgenommene Glaubwürdigkeit durch eine subjektiv schön gestaltete Webseite verbessert.

Forschungsstand. Das Autorenteam um Fogg zeigte anhand einer Umfrage (Fogg, Marshall, Lakari, et al. 2001) sowie einer qualitativen Analyse von Kommentaren zur Glaubwürdigkeitsbewertung (Fogg et al., 2003) auf, dass Layout und Farbgebung eine Rolle bei der Glaubwürdigkeitsbeurteilung spielen. Ein positiver Zusammenhang mit Farbe (vs. Schwarz-weiß) und Glaubwürdigkeit konnte in einem Online-Experiment von Syarief und Hibino (2007) bestätigt werden. Robins und Holmes (2008) stellten dar, dass eine ästhetisch gestaltete Webseite (farbig, kompakt, mit Illustrationen) im Vergleich zu einer weniger ästhetisch gestalteten (grau, einfach) Webseite zu einer höher empfundenen Glaubwürdigkeit führt.

Es ist somit naheliegend, dass Design als Merkmal ebenfalls eine Rolle zur Beurteilung der Glaubwürdigkeit spielen könnte.

2.6.1.16. Genre

Das Genre des Mediums selbst kann in Zusammenhang mit der wahrgenommenen Glaubwürdigkeit stehen. Genre meint, welcher Art das Medium angehört, also beispielsweise Zeitung, Zeitschrift, Nachrichtenorganisation, Online-Shop, offizielle Webseite einer Regierung, eines Unternehmens etc.

Glaubwürdigkeit. Menschen nehmen das Genre des Mediums, auf dem die Information gelesen oder gesehen wird, wahr und beurteilen dahingehend die Glaubwürdigkeit. Sie könnten zum einen ein gewisses Genre als glaubwürdiger wahrnehmen, da sie mit dem Genre als eine Oberkategorie von Medien eine gewisse Reputation verbinden oder gar Erfahrungen mit anderen Medien aus der gleichen Oberkategorie gemacht haben. Dies könnte sowohl mit der Reputations- als auch Bekanntheitsheuristik erklärt werden. Informationen oder Medien, die bekannter oder angesehener sind, werden gemäß der Heuristik als glaubwürdiger bewertet. Es ist zum anderen auch möglich, dass hierbei die Experten- und Autoritätsheuristik greift, denn beispielsweise könnte ein offizieller Verband für ein spezielles Fachgebiet aufgrund der zugeschriebenen Autorität (als offizieller Verband) und der Expertise (spezielles Fachgebiet) als glaubwürdiger erachtet werden. Wenn aus der gleichen Oberkategorie weitere

(unbekannte) Verbände zur Beurteilung in Frage kommen, könnten die Heuristiken gemeinsam wirken, die bereits getätigten Erfahrungen aktiviert werden und die Glaubwürdigkeitsbeurteilung für die neue Erfahrung positiv ausfallen. Insofern ist ein Zusammenhang des Genres mit der subjektiven Glaubwürdigkeitsbeurteilung begründbar und somit denkbar.

Forschungsstand. In einer qualitativen Studie (Rieh, 2002) beantworteten Probanden eine medizinische Fragestellung mithilfe von Informationen aus dem Web. Zudem wurden Think-Aloud Protokolle und Interviews durchgeführt, um Aufschluss über die verwendeten Kriterien zur Beurteilung der Informationsqualität und somit Glaubwürdigkeit zu erhalten (Rieh, 2002). Die Forschenden analysierten die Think-Alouds zusammen mit den Probanden und ermittelten die Verteilung der zur Beurteilung herangezogenen Merkmale (Rieh, 2002). Sie können u. a. das Merkmal Genre (bei ihnen „Art der Quelle“) identifizieren, was zeigt, dass die Probanden bei der Beurteilung der Informationsqualität und somit der Glaubwürdigkeit auf das Genre der Information bzw. Webseite achten (Rieh, 2002). In einer experimentellen Studie manipulierten Flanagin und Metzger (2007) das Genre, indem sie Probanden inhaltlich gleiche Informationen von vier verschiedenen Medien (Nachrichtenorganisation *CNN*, E-Commerce Webseite *Baby-Center.com*, Webseite einer besonderen Interessensgemeinschaft *www.childrensdefense.org*, sowie eine individuelle Webseite) hinsichtlich verschiedener Glaubwürdigkeitsarten (Sponsoren-, Nachrichten- und Seiten-Glaubwürdigkeit) der Informationen bewerten ließen. Sie fanden heraus, dass für die Sponsoren- und Nachrichten-Glaubwürdigkeit die Nachrichtenorganisation als am glaubwürdigsten erachtet wird (Flanagin & Metzger, 2007). Denner et al. (2016) berichten ähnliche Ergebnisse in ihrer experimentellen Studie, in der sie eine 2x2-Manipulation des Merkmals Genre vornahmen. Sie unterschieden zwischen einer Boulevard-Zeitung (*Bild*) und einer angesehenen Qualitätszeitung (*Zeit*) und untersuchten, ob sich Unterschiede in der Glaubwürdigkeitsbeurteilung ergaben (Denner et al., 2016). Ihre Studie ergab, dass bezüglich vier von fünf Dimensionen ihres Glaubwürdigkeitskonstrukts die Qualitätszeitung im Vergleich zur Boulevard-Zeitung als glaubwürdiger wahrgenommen wird (Denner et al., 2016).

Es zeigt sich daher, dass das Genre ebenfalls in Zusammenhang mit wahrgenommener Glaubwürdigkeit stehen könnte.

2.6.1.17. Kommerzielles Interesse

Abgrenzung des kommerziellen Interesses von Objektivität. Die Voraussetzung für die Glaubwürdigkeit bzw. Qualität einer Quelle sollte sein, dass diese unabhängig, d. h. objektiv ist. Objektivität wird in der quantitativen Forschung als die „Unabhängigkeit“ definiert (Moosbrugger & Kelava, 2008, S. 8 ff.) und bedeutet, dass eine andere Person unter gleichen Voraussetzungen zum gleichen Ergebnis kommen würde (Mack, 2021, S. 99). Das impliziert, dass

eine Person zuvor, also bei der Erhebung der Ergebnisse oder angewandt auf den Online-Kontext – beim Zusammenstellen der Informationen – ihre subjektiven Einflüsse (z. B. Biases) erfolgreich ausschließen konnte (s. Definition von Tate, 2019, S. 14)¹⁷. Objektivität kann daher durch eine Vielzahl an subjektiven Einflüssen bzw. Interessen (z. B. kommerzielles oder ideologisches) gefährdet sein und hängt zudem von diesen und weiteren Faktoren, wie Vorkenntnisse oder Erfahrungen, ab (American Psychological Association, o. J.). Insbesondere im Online-Kontext hat ein Informationssuchender nur wenig Wissen über die Vorkenntnisse oder Erfahrungen der Person, die die Informationen bereitstellt oder veröffentlicht. Was allerdings als subjektiver Einfluss meist einfach zu identifizieren ist, ist das Vorliegen eines kommerziellen Interesses, durch z. B. das Schalten von thematisch passenden Werbeanzeigen oder das Werben für einen thematisch passenden Online-Shop. Wenn das subjektive bzw. kommerzielle Interesse ausgeschlossen werden kann, könnte eine Information bzw. Webseite im weiten Sinn als objektiv bezeichnet werden. Metzger, Flanagin und Zwarun (2003, S. 283) stützen diese Überlegung, denn nach ihnen beinhaltet die Beurteilung der Objektivität das Identifizieren des Zweckes der Webseite. Die Abwesenheit von Interesse bzw. kommerziellem Interesse speziell im Online-Kontext ist daher notwendig, um möglicherweise zu schlussfolgern, dass etwas objektiv ist. Um aber mit 100 %-iger Sicherheit etwas als objektiv bezeichnen zu können, müssten neben dem kommerziellen Interesse oder der Intention auch die weiteren subjektiven Einflüsse (Vorkenntnisse, Erfahrungen, ideologische Überzeugungen) ausgeschlossen werden. Mangels Messbarkeit dieser weiteren subjektiven Einflüsse, die die Objektivität einer Information oder Webseite limitiert und weil sich für Suchende im Online-Kontext das kommerzielle Interesse leichter erkennen lässt, soll für diesen Forschungszweck daher das kommerzielle Interesse in den Vordergrund gestellt werden. Aufgrund der soeben dargestellten Komplexität und der Schwierigkeit, die Begriffe trennscharf voneinander abzugrenzen, werden jedoch auch alle Forschungen einbezogen, die beide Begrifflichkeiten thematisieren, sofern sie der Bedeutung nach den oben dargestellten Ausführungen zur Definition entsprechen.

Glaubwürdigkeit. Wenn Personen Webseiten bezüglich des Merkmales des kommerziellen Interesses beurteilen, und zur Schlussfolgerung gelangen, dass die Information oder Webseite nicht gänzlich frei von besagtem Interesse ist, kann dies die Glaubwürdigkeitsbeurteilung beeinflussen. Erklärung über diesen Sachverhalt gibt die bereits in Kapitel 2.2.2.1 dargestellte Reaktanz. Dieses Gefühl der Reaktanz wird durch das Wahrnehmen der Manipulation (durch z. B. Erkennen des kommerziellen Interesses) hervorgerufen und verursacht eine negative Bewertung, die sich in einer geringen empfundenen Glaubwürdigkeit niederschlagen könnte.

¹⁷ „Objectivity is the extent to which material expresses facts or information without distortion by personal feelings or other biases“ (Tate, 2019, S. 14).

Forschungsstand. Tate (2019) berichtet in ihrem Buch über essenzielle Merkmale, wie das der Objektivität, die die Vertrauenswürdigkeit (als eine der beiden Komponenten von Glaubwürdigkeit) beeinflussen kann und auf die sich Informationskonsumenten einstellen müssen. Basierend auf ihren Ausführungen zur Objektivität leiten sie Empfehlungen und Leitfragen ab, wie Nutzer zukünftig der Objektivität bzw. dem kommerziellen Interesse begegnen können. Sie bringen daher indirekt kommerzielles Interesse bzw. Objektivität mit der Vertrauenswürdigkeit (und somit Glaubwürdigkeit) in Verbindung. Walraven et al. (2009) eruierten mithilfe von Gruppeninterviews, auf welche Merkmale die interviewten Teilnehmer bei der Lösung einer Online-Aufgabe achten würden, und teilten dies anhand eines zuvor entwickelten Codesystems (Brand-Gruwel et al., 2005; Wopereis et al., 2008) den entsprechenden Phasen und Merkmalen zu. Auch hieraus ergab sich anhand der analysierten Interviews die Erkenntnis, dass ihre Teilnehmer auf Objektivität achten werden. Die tatsächliche Anwendung des Merkmals von Seiten der Interviewten bleibt in einer zweiten, praktischen Runde aus. Dennoch verdeutlicht diese Forschung, dass kommerzielles Interesse, wenn auch hier nur durch qualitative Analyse ermittelt, für die Probanden eine bedeutende Rolle spielte. Fogg, Marshall, Lakari, et al. (2001) zeigten anhand ihrer Umfrage, dass die Abwesenheit von kommerziellem Interesse überdurchschnittlich oft als Indikator für eine positive Glaubwürdigkeitsbeurteilung genannt wird. Anhand der Analyse von Nutzerkommentaren zur Glaubwürdigkeitsbeurteilung fanden Forschende (Fogg et al., 2003) heraus, dass sich unter den Top 10 der genannten Merkmale u. a. Kommentare zu „Motiv des Unternehmens“ sowie „Bias von Informationen“ befinden. Ersteres und im entfernten Sinne auch letzteres können unter das Merkmal kommerzielles Interesse subsumiert werden. Im Rahmen eines Gruppenvergleichs zwischen Studierenden und Nicht-Studierenden und ihrer Befragung zeigte sich, dass innerhalb der Studienteilgruppe das Berücksichtigen der Ziele (und somit Intention) des Autors zur Evaluation von Informationen überdurchschnittlich oft genannt wird (Metzger, Flanagin & Zwarun, 2003).

Anhand der dargestellten Forschungen wird deutlich, dass kommerzielles Interesse bei der subjektiven Bewertung von Glaubwürdigkeit eine Rolle spielen könnte, weshalb es für vorliegende Forschung in Betracht gezogen wird.

2.6.1.18. Bekanntheit

Einführung und Glaubwürdigkeit. Bekannte Quellen oder Inhalte, mit denen bereits Erfahrungen gemacht wurden, erscheinen für Menschen ebenfalls als glaubwürdiger als unbekannte Quellen oder Inhalte, was mit der Bekanntheitsheuristik und der Verarbeitungsleichtigkeit erklärt werden kann. Wie bereits an anderer Stelle dargestellt, lernt eine Person aufgrund (positiver) Erfahrungen mit einer Quelle das Muster, dass Bekanntes mit Gutem verbunden wird, wie mit Glaubwürdigkeit (Gigerenzer & Gaissmaier, 2011). Wenn die Inhalte bekannt sind,

kann an vorhandenes Wissen angeknüpft werden, was eine schnellere Verarbeitung im Gehirn ermöglicht. Aufgrund der empfundenen Verarbeitungsleichtigkeit werden positive Eigenschaften (Graf et al., 2018), wie z. B. Glaubwürdigkeit, zugeschrieben.

Forschungsstand. Auf Basis ihres umfangreichen Literaturreviews postulieren Wathen und Burkell (2002) unter anderem, dass die Nachrichten bzw. Informationsbekanntheit einen Einfluss auf die Glaubwürdigkeitsbeurteilung haben kann. Fogg et al. (2003) fanden bei der Analyse von Kommentaren zur Beurteilung der Glaubwürdigkeit zweier Live-Webseiten heraus, dass die Probanden neben weiteren Merkmalen in 14.1 % der Kommentare Anmerkungen bezüglich des Bekanntheitsgrades (Original: „name recognition“) anführen. Dies zeigt somit, dass in diesem Fall die Bekanntheit des Mediums als Merkmal zur Beurteilung der Glaubwürdigkeit herangezogen wird. Daran anschließend ist die Studie von Hilligoss und Rieh (2008) anzubringen, die mittels Tagebucheinträgen und Interviewanalysen herausfanden, dass die Probanden die Glaubwürdigkeit der gefundenen Informationen auch bezüglich der Bekanntheit beurteilen. In einem Experiment manipulierten Forschende die Bekanntheit des Mediums und ermittelten, ob sich die (Un-)Bekanntheit auf die Glaubwürdigkeitsbeurteilung auswirkte (Denner et al., 2016). Sie können die Hypothese des Einflusses der Bekanntheit jedoch nicht bestätigen. Auch die Bekanntheit des Inhaltes bzw. des Themas wurde gezielt manipuliert: Barzilai et al., (2020) zeigten, dass die Bekanntheit des Themas die wahrgenommene Zuverlässigkeit (und somit Glaubwürdigkeit) eines Experten erhöht, sofern die Aussagen des Experten glaubenskonform mit der eigenen Einstellung sind.

Es zeigt sich somit, dass Bekanntheit der Quelle, des Autors oder des Themas mit der wahrgenommenen Glaubwürdigkeit in Verbindung stehen könnte.

2.6.1.19. Reputation

Einführung und Glaubwürdigkeit. Die Reputation bzw. der Ruf eines Mediums bzw. einer Nachrichtenorganisation beeinflusst, welchen Wert und welche Glaubwürdigkeit Menschen diesem Medium beimessen, was unter anderem auf die Reputationsheuristik zurückzuführen ist. Eine gute Reputation baut sich beispielsweise dadurch auf, dass andere Personen zuvor positive Erfahrungen mit dem Medium/der Quelle gemacht haben und dies weiter kommunizieren. Wenn ein Medium daher eine Reputation aufweist und dies von einem Leser erkannt wird (durch die entsprechenden Hinweisreize), assoziiert dieser damit auch indirekt, dass andere bereits gute Erfahrungen mit der Quelle gemacht haben, weshalb der gute Ruf überhaupt zu Stande kam. Der Bandwagon-Effekt ist daran ebenfalls beteiligt, da andere durch die Weiterkommunikation des Rufs eines Mediums indirekt eine Empfehlung ausgesprochen haben. Der Bandwagon-Effekt besagt genau das, nämlich dass ein Medium oder eine Webseite deshalb als glaubwürdig erachtet werden kann, weil ihnen durch andere (z. B. Empfehlung)

signalisiert wird, dass sie glaubwürdig sei (Sundar, 2008). Der Bandwagon-Effekt hat in den Online-Medien oder im Speziellen, den sozialen Medien, sogar einen weitaus größeren Stellenwert aufgrund z. B. den damit einhergehenden *Like*-, *Share*- und *Comment*-Funktionen. In Kombination mit der Reputationsheuristik an sich ist es daher denkbar, dass Quellen oder Medien mit einem höheren Ruf auch eine größere Glaubwürdigkeitsbeurteilung erhalten. In einem Fokusgruppeninterview zu Glaubwürdigkeitsbeurteilungen wurde dieser Zusammenhang mit dem Bandwagon-Effekt bestätigt: Die Probanden aus der Studie nehmen an, dass die Wiedererkennung im Laufe der Zeit durch positive Interaktionen, die über soziale Medien verbreitet wurden, erworben werden (Metzger et al., 2010). Allein die positiven Interaktionen über soziale Medien, die andere machen (Bandwagon-Effekt), schüren somit die positiven Beurteilungen einer Quelle.

Forschungsstand. Die Autoren (Fogg et al., 2003) stellten bei der Analyse von Kommentaren zur Glaubwürdigkeitsbeurteilung fest, dass Kommentare in Bezug auf die Reputation vergleichsweise häufig (14.1 %) enthalten sind. Walraven et al. (2009) merken in ihrer Forschung kritisch an, dass die Reputation der Quelle eigentlich bei der Beurteilung der Glaubwürdigkeit zu beachten sein sollte. Fogg, Marshall, Lakari, et al. (2001) veranschaulichten, dass eine gut respektierte Webseite, welche somit eine gewisse Reputation hat, von ihren Probanden überdurchschnittlich oft als Indikator für eine positive Glaubwürdigkeitsbeurteilung genannt wird. McKnight und Kacmar (2007) fanden heraus, dass eine Webseite als glaubwürdiger wahrgenommen wird, wenn sie Zusatzinformationen zur Wahrnehmung einer gewissen Reputation enthielt im Vergleich zu Webseiten ohne diese Informationen.

Es ist somit davon auszugehen, dass die Reputation mit der wahrgenommenen Glaubwürdigkeitsbeurteilung in Verbindung steht.

2.6.1.20. Angabe des Autors

Einführung und Glaubwürdigkeit. Die Identität in Form vom Namen eines Autors zu erkennen, könnte ebenfalls vertrauenserweckend wirken und in einer hohen wahrgenommenen Glaubwürdigkeitsbeurteilung resultieren. Dies kann mit sowohl der Autoritäts-, Wiedererkennungs- als auch der Bekanntheitsheuristik erklärt werden. Um die Autorität einer zu beurteilenden Person zunächst einmal feststellen zu können, ist es zwingend erforderlich, dass der Autor auch angegeben wird. Insofern spricht eine Autorenangabe an sich schon für die Wahrnehmung einer ausreichenden Glaubwürdigkeit. Weiterhin könnten im Zuge der Wiedererkennungs- und Bekanntheitsheuristik auch ähnlich klingende Namen, die an eine bekannte Person der beurteilenden Person erinnern, positive Erinnerung an diese wecken und den Autor in einem positiven Licht (z. B. Glaubwürdigkeit) erscheinen lassen. Auch hierfür ist obligatorisch,

dass zunächst Autorennamen erkennbar sind. Insofern könnten Probanden die Angabe von Autoren als glaubwürdiger erachten.

Forschungsstand. Mit ihrem konzeptuellen Modell zur Beurteilung der Glaubwürdigkeit von Online-Informationen von 2002 komprimierten die Autoren Wathen und Burkell viele Merkmale der Quelle, der Nachricht, der Oberfläche oder des Inhaltes. Darunter zu finden ist ebenfalls die Angabe von Autorennachweisen (Name) (Wathen & Burkell, 2002). In der Studie von Rieh (2002) wurde ebenfalls das Merkmal der Autorenschaft untersucht. Die Probanden nutzen in der medizinischen Aufgabenstellung die Hinweise zum Autor (bei ihnen „author credentials“), um zu einer Glaubwürdigkeitsbeurteilung der Information zu gelangen. Hope (2007) stellte fest, dass die Autorenangabe von ihren Probanden zur Beurteilung der Glaubwürdigkeit herangezogen wird. Weiterhin zeigt sich an der an mehreren Stellen bereits angeführten qualitativen Studie von Walraven et al. (2009), dass auch bei ihnen die Angabe der Autorenschaft ein wichtiges Merkmal zur Bewertung der Glaubwürdigkeit ist. Die Probanden nennen u. a. dieses Merkmal in den Gruppeninterviews. Dennoch wird es in der eigentlichen Anwendung der Aufgabe nicht verwendet. Ebenso durch qualitative Analyse von Kommentaren von Nutzern konnten Forschende (Fogg et al., 2003) Erkenntnisse darüber gewinnen, auf was Nutzer bei der Bewertung der Glaubwürdigkeit von Webseiten achten. Unter den Top 10 der genannten Merkmale befinden sich Kommentare zu „Identität des Autors“, was somit unter das Merkmale der Autorenangabe fällt und als weiteren Beleg aus der Forschung mit aufgenommen werden kann. Schließlich konnten Fogg, Marshall, Lakari, et al. (2001) auch mittels einer Umfrage darstellen, dass die Angabe von Autorennachweisen von ihren Probanden überdurchschnittlich oft als Indikator für eine positive Glaubwürdigkeitsbeurteilung genannt wird.

Obwohl das Gros der bisherigen Forschungen die Angabe des Autors als Merkmal nur durch qualitative oder konzeptuelle Forschungen darstellt, zeigt sich dennoch eine Relevanz hierfür. Personen achten bei der Beurteilung der Glaubwürdigkeit auf die Angabe von Autoren. Es ist folglich essenziell, dies in die Betrachtung einfließen zu lassen, da erwartet wird, dass auch die Probanden der vorliegenden Forschung solche Angaben in diese Richtung machen werden.

2.6.1.21. Foto eines Autors

„Bilder sagen mehr als tausend Worte“ besagt eine Metapher, die vermutlich ursprünglich auf Fred R. Bernard aus dem Jahre 1921 zurückgeht. Nach wie vor sind sich Menschen heutzutage dessen bewusst und versuchen durch visuelle im Vergleich zu textlichen Eindrücken positive Assoziationen hervorzurufen. Denn trotz der neusten Entwicklungen um so genannte Deepfake-Anwendungen (*Deepfakes - Gefahren und Gegenmaßnahmen*, o. J.) wird der

Großteil der Menschen der Aussage zustimmen, dass Bilder nicht lügen können und zu Vertrauen und folglich Glaubwürdigkeit betragen können (Sundar, 2008).

Glaubwürdigkeit. Diese Folgerung könnte mit der Verarbeitungsleichtigkeit erklärt werden. Denn Bilder werden im Gedächtnis automatischer und mit geringerer kognitiver Anstrengung aufgenommen und verarbeitet (Kroeber-Riel & Esch, 2015) und sogar erinnert (Kirkpatrick, 1894); letzteres wird in der Konsumentenforschung als der Picture-Superiority-Effekt bezeichnet (Paivio, 1991). Aufgrund der leichteren Verarbeitung könnte der zu verarbeitenden Information eine positivere Bewertung, z. B. Glaubwürdigkeit, beigemessen werden. Insofern ist es denkbar, dass, wenn Bilder eines Autors vorhanden sind, dies mit einer höheren Glaubwürdigkeit assoziiert wird.

Forschungsstand. In einer Studie (Fogg, Marshall, Kameda, et al., 2001) wurde untersucht, wie sich u. a. der Name und das Bild eines Autors auf die Bewertung der Glaubwürdigkeit auswirkten. Auf unterschiedlichen Dimensionen ihres Glaubwürdigkeitskonstrukts (u. a. Vertrauenswürdigkeit, Expertise oder Kompetenz) erzielt ein formales Foto im Vergleich zu einem zwanglosen oder keinem Foto die höchsten Werte (Fogg, Marshall, Kameda, et al., 2001). Die Forschenden schlussfolgerten, dass der Stil des Fotos einen Einfluss auf die wahrgenommene Glaubwürdigkeit der Information hat (Fogg, Marshall, Kameda, et al., 2001). Johnson und Wiedenbeck (2009) ließen die Probanden Nachrichten auf Glaubwürdigkeit beurteilen und manipulierten hierbei unter anderem, ob das Foto eines Autors einen Einfluss auf die Glaubwürdigkeitsbeurteilung hatte. Der Einfluss kann bestätigt werden. Diese Resultate konnten in einer anderen Forschung nicht gänzlich wiederholt werden, da im Zuge der Bewertung der Glaubwürdigkeit eines LinkedIn-Profiles lediglich die Glaubwürdigkeitsdimension der „Kompetenz“ in einem statistisch signifikanten positiven Zusammenhang mit dem Vorhandensein eines Profilbildes (vs. kein Profilbild) steht (Edwards et al., 2015).

Dennoch wird für die vorliegende Forschung ein Zusammenhang mit der wahrgenommenen Glaubwürdigkeit und das Vorhandensein eines Fotos angenommen.

2.6.2. Theoretische Zusammenfassung der Webseitenmerkmale

Alle Webseitenmerkmale haben einen möglichen Einfluss auf die wahrgenommene Glaubwürdigkeit. Dies wurde für jedes Merkmal mithilfe der bereits dargestellten Heuristiken begründet. Bedeutsam zu verstehen ist jedoch, dass die Merkmale und Heuristiken nicht immer den gleichen Bezugsobjekten unterliegen. So wird beispielsweise die faktische Korrektheit eines Textes bzw. einer Information beurteilt, jedoch wird dieser Information eine entsprechende Glaubwürdigkeit/Unglaubwürdigkeit zugeschrieben, weil der Autor der Information als Experte wahrgenommen wird. Das heißt, es fließt indirekt eine Beurteilung des Autors in die Beurteilung der Glaubwürdigkeit der Information mit ein. Die uneindeutige Trennung der Bezugsobjekte der

Merkmale bei der Anwendung von Heuristiken ist daher nicht gänzlich gewährleistet. Da der Fokus der Forschungsarbeit jedoch auf der Beurteilung der Glaubwürdigkeit auf Basis der Merkmale liegt, wird sich auf das Bezugsobjekt konzentriert, welches das eigentliche Merkmal bereitstellt. Im Beispiel wäre das Merkmal die faktische Korrektheit und diese wird anhand der Information bzw. des Textes ersichtlich. Warum das Merkmal der faktischen Korrektheit als glaubwürdig empfunden würde, durch z. B. die Assoziation des Autors als Experte, ist somit nicht Untersuchungsgegenstand. Deshalb wird die faktische Korrektheit lediglich dem Bezugsobjekt Information/Text zugeschrieben.

Auch wenn der Fokus der Arbeit nicht auf den Heuristiken zur Glaubwürdigkeitsbeurteilung liegt, zeigt das Beispiel zudem auf, dass es hilfreich sein kann, thematisch ähnliche Webseitenmerkmale (also solche mit gleichem Bezugsobjekt) zusammenzufassen. Dies vereinfacht unter anderem das Verständnis der Forschungsarbeit. Die Dreier-Einteilung wurde bereits in Kapitel 2.2.1 in Verbindung mit den drei Subkategorien der Glaubwürdigkeit angedeutet. Die Einteilung in diese drei Bezugsobjekte ist außerdem deshalb zweckdienlich, weil die einzelnen Webseitenmerkmale funktionale Äquivalente für die wahrgenommene Glaubwürdigkeit darstellen. Das bedeutet, sie können alle den gleichen Effekt, die Glaubwürdigkeit zu steigern, haben¹⁸. Für die Integration in die Modelle und um die Modellkomplexität zu reduzieren, reicht es daher die thematisch bzw. inhaltlich ähnlichen Merkmale zu betrachten. Der daraus resultierende Datenverlust aufgrund der Aggregation ist der Forschenden bewusst und wird an entsprechender Stelle kritisch diskutiert.

Die Bezugsobjekte der Webseitenmerkmale richten sich nach dem zu beurteilenden Gegenstand hinsichtlich der wahrgenommenen Glaubwürdigkeit, sprich nach Quellen-/Medienmerkmalen, Informations-/Nachrichtenmerkmalen sowie Autorenmerkmalen. Sie sind somit korrespondierend zu den drei Subkategorien der Glaubwürdigkeit. Einige Merkmale sind entweder vollständig in einer Kategorie, andere wiederum als Schnittmenge in mehreren Kategorien zu finden sind. Die Einteilung der Webseitenmerkmale erfolgte nach einer Literaturrecherche ähnlicher Forschungen zu bisherigen Einteilungen (Choi, 2015; Flanagin & Metzger, 2007; T. Hong, 2006; Metzger, Flanagin, Eyal, et al., 2003; Metzger & Flanagin, 2015; Walraven et al., 2009; Wathen & Burkell, 2002). Abbildung 8 zeigt die hier anzuwendende Einteilung.

¹⁸ Mit Ausnahme von Narrativen und kommerziellem Interesse. Kommerzielles Interesse muss umgekehrt betrachtet werden: Je höher das kommerzielle Interesse, desto geringer die wahrgenommene Glaubwürdigkeit. Die Richtung der Glaubwürdigkeitszuschreibung der Narrative bleibt vorerst unklar.

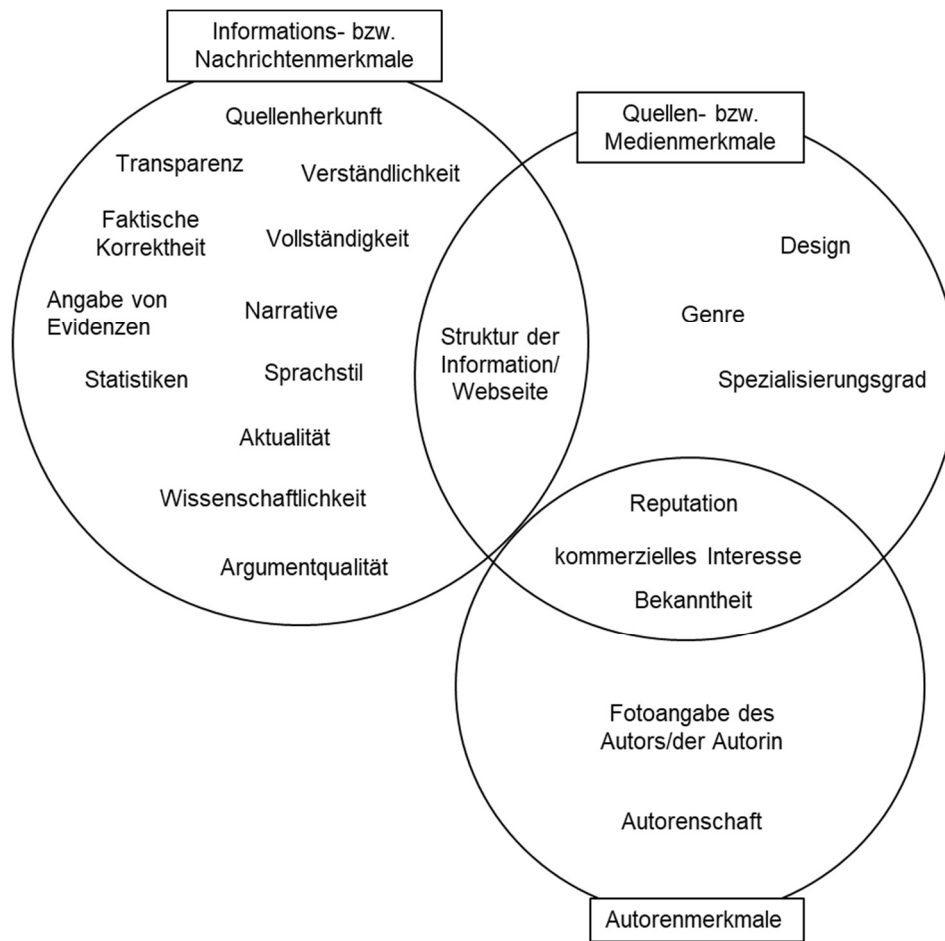


Abbildung 8: Einteilung der Webseitenmerkmale in drei Kategorien.

2.6.3. Individuelle Merkmale

Da der Informationsverarbeitungsprozess innerhalb des Gehirns eines Individuums stattfindet und somit automatisch von dem Individuum, seinen Erfahrungen und Wissen abhängt, spielen Personenmerkmale eine entscheidende Rolle. Einige Forschende legen dar, dass unterschiedliche Personenmerkmale, wie Erfahrungen mit dem Informationsgegenstand bzw. der Domäne (Experte/Novize) (Brand-Gruwel et al., 2017) oder dem Medium (Experte/Novize aus 3S Modell) (Flanagin & Metzger, 2000, 2007), Motivation (Chaiken, 2014; Metzger, 2007; Petty & Cacioppo, 1986), Kognitionen (Kim & Allen, 2002; Metzger et al., 2015), kognitive Stile (Amelang & Bartussek, 1997; Brünken et al., 2019), Vorwissen (Werth et al., 2020), generelle Fähigkeit (Chaiken, 2014; Metzger, 2007; Petty & Cacioppo, 1986) aber auch soziodemografische Merkmale wie Alter (Chin et al., 2009; Sharit et al., 2015) oder Geschlecht (Arcand & Nantel, 2012; Forzani, 2018) den Informationsverarbeitungsprozess (oder Teile des Gesamtprozesses) beeinflussen.

Für vorliegende Forschung wird daher untersucht, ob ausgewählte Personenmerkmale an diesem Prozess beteiligt sind. Hierzu werden diese im nachstehenden Kapitel zunächst definiert und beschrieben. Weiterhin werden sie in den allgemeinen Forschungskontext in

Zusammenhang mit der Informationsverarbeitung eingeordnet. Der allgemeine Forschungskontext dient ausschließlich dafür, Grundlagen für die später folgende Hypothesenherleitung zu schaffen. Letztlich wird der kontextbezogene Forschungskontext thematisiert und liefert einen Einblick, inwieweit das Personenmerkmal in der Online-Umgebung untersucht wurde. Im darauffolgenden Kapitel 2.7 wird ausführlich eruiert, warum die Personenmerkmale und die Kombination dieser für die Forschung ausgewählt wurden. Soziodemografische Merkmale werden als Kontrollvariablen in die Modelle mit aufgenommen, es ist jedoch nicht auszuschließen, dass auch sie aufgrund der bisherigen Literatur eine entscheidende Rolle tragen. Dies wird allerdings für diese Forschung nicht zum Fokus gemacht, sondern lediglich bei entsprechenden Ergebnissen und deren Notwendigkeit zur Diskussion thematisiert.

2.6.3.1. Need for Cognition

Mitte der 1950-er Jahre wurde erstmalig das so genannte Kognitionsbedürfnis (Englisch: Need for Cognition; im weiteren Verlauf mit NfC abgekürzt) thematisiert, als Forschende (A. R. Cohen et al., 1955; A. R. Cohen, 1957) aufzeigen wollten, dass Menschen unterschiedlich in dem Bedürfnis sein können, Erfahrungen sinnvoll zu organisieren (Cohen et al., 1955, S. 291). Sie konstatierten, dass NfC existieren, gemessen werden und unabhängig von anderen Bedürfnissen sein kann (A. R. Cohen et al., 1955, S. 293). Ihr Konsens war, dass es Unterschiede zwischen Personen geben muss, die auf NfC zurückzuführen sind. Diese Unterschiede zeigen sich darin, dass Personen mit hohem NfC mit höherer Wahrscheinlichkeit Informationen, denen sie begegnen, organisieren, bearbeiten und beurteilen (A. R. Cohen, 1957). Basierend auf diesen Ergebnissen entwickelten Cacioppo und Petty (1982) einige Jahre später eine Skala zur Messung des NfC (Cacioppo & Petty, 1982). Ihr Ziel war es, zu überprüfen, ob Personen mit unterschiedlichem NfC auch unterschiedlich viel Freude an kognitiven Anstrengungen erleben. Diese damalige Annahme, dass ein Kognitionsbedürfnis existiert und Menschen sich dahingehend unterscheiden können, konnte verifiziert werden (Cacioppo & Petty, 1982).

Forschungsstand. Aufgrund dessen, dass NfC zum jetzigen Zeitpunkt gut erforscht ist (Petty et al., 2009, S. 318), konnte bereits eine Vielzahl an Zusammenhängen nachgewiesen werden: Menschen mit hohem im Vergleich zu niedrigem NfC empfinden größere Freude an kognitiver Anstrengung und an komplexen Aufgaben (Cacioppo & Petty, 1982), wie z. B. das Markieren von bestimmten Ziffern aus einer Vielzahl von Ziffern nach komplexen Regeln (Cacioppo & Petty, 1982). Außerdem konnten die Autoren belegen, dass Menschen mit hohem NfC eine größere Motivation haben, Informationen zentral (lt. ELM) zu verarbeiten, da sie sich tiefergehend mit Argumenten auseinandersetzen (Cacioppo et al., 1983). Bestätigt wurde dies weiterhin von Chaiken (2014), die bewies, dass Personen mit hohem NfC Informationen systematischer verarbeiten im Vergleich zu Personen mit niedrigerem NfC. Hinsichtlich allgemeiner

Informationsverarbeitung wurde aufgezeigt, dass sie ein höheres Erinnerungs- und Wiedererkennungsvermögen aufweisen (Diehl & Wyrick, 2015). Auch ein Zusammenhang mit Performanz wurde nachgewiesen, denn hohes NfC führt bei Medizinstudenten zu einer besseren Performanz (Heijne-Penninga et al., 2010) und steht generell in positivem Zusammenhang mit einer höheren intellektuellen Aufgabenperformanz (Coutinho, 2006; Coutinho et al., 2005). Zudem empfinden Menschen mit hohem NfC eine größere Freude am Argumentieren (Sanders et al., 1992) und bei der Generierung komplexer und elaborierter Argumente (Nussbaum, 2005).

Kontextbezogener Forschungsstand. Für vorliegende Forschung wird angenommen, dass NfC ebenfalls bei der Suche, Selektion und Evaluation von Informationen eine bedeutsame Rolle spielt, da in diesen Prozessen, wie bereits in den Kapiteln 2.2.1 und 2.4 aufgezeigt, Kognitionen beteiligt sind. In diesem Zusammenhang wurde bereits die Rolle des NfC bei der Bewertung der Glaubwürdigkeit von Online-Informationen untersucht, was zu folgenden Erkenntnissen führte: Menschen mit hohem NfC haben größere Ressourcen der Aufmerksamkeitszuteilung (Fleischhauer et al., 2010), investieren generell mehr kognitive Anstrengung bzw. Ressourcen in eine Suchaufgabe (Verplanken et al., 1992) sowie in die Informationsverarbeitung (Enge et al., 2008) und suchen mehr Informationen (Verplanken et al., 1992). Metzger et al. (2015) fanden mithilfe einer Umfrage heraus, dass hohes NfC ihrer Probanden die Verwendung von analytischen Strategien (z. B. die Informationen sorgfältig zu prüfen) zur Beurteilung der Glaubwürdigkeit voraussagt. Zudem wurde gezeigt, dass die Probanden mit hohem NfC bei der Nutzung von Online-Informationen weniger leichtgläubig sind (Metzger et al., 2015). Ähnlich fanden Schaewitz et al. (2020) in ihrer Studie heraus, dass NfC ein signifikanter Prädiktor für die wahrgenommene Glaubwürdigkeit und Genauigkeit von Fehlinformationen ist. Das bedeutet, dass die Probanden mit hohem NfC Fehlinformationen (korrekterweise) als unglaubwürdig bewerteten und einen Großteil dieser Informationen als falsch/inakkurat einschätzten (Schaewitz et al., 2020).

Obgleich sich bereits viele Arbeiten mit dem Konstrukt und auch im Zusammenhang von Online-Informationen beschäftigten, ist bis zum jetzigen Stand wenig darüber erforscht, wie sich NfC und mehrere, verschiedene Webseitenmerkmale gleichzeitig auf die Prozesse der Online-Informationsverarbeitung auswirken.

2.6.3.2. Need for Cognitive Closure

Das Bedürfnis nach kognitiver Abgeschlossenheit (NfCC) (Englisch: Need for Cognitive Closure; im weiteren Verlauf mit NfCC abgekürzt) wurde erstmal von Webster und Kruglanski (1994) thematisiert und bezeichnet das Bedürfnis nach einem schnellen Abschluss und die Abneigung gegen Unsicherheit oder Ambiguität. Obwohl auch situative Faktoren, wie z. B.

Zeitdruck (Freund et al., 1985; Kruglanski & Freund, 1983), oder Geräusche (Kruglanski & Webster, 1991) – für eine Übersicht, siehe (Kruglanski, 2013) – Auslöser eines solchen Gefühls sein können, argumentieren die Autoren, dass jeder eine gewisse individuelle Dispositionen für NfCC hat (Webster & Kruglanski, 1994). Menschen mit hohem NfCC haben grundlegend zwei Tendenzen inne: die Dringlichkeitstendenz und die Beständigkeitstendenz (Kruglanski & Fishman, 2009; Schlink, 2009). Die Dringlichkeitstendenz beschreibt den Wunsch, schnell und sofort einen Abschluss (i. S. v. abhaken) hinsichtlich einer Problematik oder Thematik zu erzielen („seize“), wobei das Aufschieben bzw. Nicht-Erreichen eines schnellen Abschlusses als unangenehm empfunden wird (Kruglanski & Fishman, 2009). Die Beständigkeitstendenz beschreibt den Wunsch, den Abschluss möglichst lange aufrecht zu erhalten bzw. einzufrieren („freeze“) (Kruglanski & Fishman, 2009). Neue, andere oder gar widersprüchliche Informationen zu berücksichtigen, die den Abschluss gefährden, werden ignoriert (Kruglanski & Fishman, 2009).

Webster und Kruglanski (1994) entwickelten eine Skala zur Messung, inwieweit ein Individuum hinsichtlich NfCC veranlagt sein kann. Die Disposition wird dabei als Kontinuum betrachtet, wobei hohes und niedriges NfCC die beiden Ende bilden. Die Autoren (Webster & Kruglanski, 1994) vermuteten, dass NfCC und NfC korreliert sein könnten, da NfC an Denkprozessen beteiligt ist und die Schritte zu einem Endzustand eines Denkprozesses (was somit das wünschenswerte Ziel unter Beachtung von NfCC ist) beeinflussen könnte (Webster & Kruglanski, 1994). Aus diesem Grund führten sie eine Zusammenhangsanalyse durch und konnten ihre Vermutung bestätigen, dass NfC und NfCC leicht negativ korrelieren (Webster & Kruglanski, 1994).

Forschungsstand. Im Vergleich zu NfC ist NfCC weniger gut untersucht, jedoch existieren einige Forschungen¹⁹, welche nachstehend dargestellt werden: Beispielsweise konnte aufgezeigt werden, dass Personen mit hohem NfCC abstrakte Beschreibungen bevorzugen (Fortier & Burkell, 2014) und bei der Analyse von Beweisen ein geringeres Maß an Genauigkeit aufweisen (Klein & Webster, 2000). Hinsichtlich allgemeiner Informationsverarbeitung wurde herausgefunden, dass Menschen mit hohem NfCC weniger Webseiten mit Informationen zu Rate ziehen als Menschen mit niedrigem NfCC (Webster & Kruglanski, 1994). Sie versuchen stattdessen sehr schnell Informationen zu finden (Dringlichkeitstendenz, „seize“) und diesen Zustand zu erhalten (Beständigkeitstendenz, „freeze“) (Kruglanski & Fishman, 2009). Außerdem benötigen Menschen mit hohem NfCC weniger Zeit (Webster & Kruglanski, 1994) und einen geringeren Umfang an Informationen (Kruglanski & Fishman, 2009), um zu einem Urteil zu

¹⁹ Weitere, für diesen Kontext jedoch nicht relevante Forschungen sind beispielsweise aus dem Bereich Konsumentenforschung (Amichai-Hamburger et al., 2004; Jung & Kellaris, 2004), Integration (Kosic et al., 2004) oder Stereotypenforschung (Dijksterhuis et al., 1996; Shah et al., 1998).

gelangen als Menschen mit niedrigem NfCC. Menschen mit hohem NfCC haben jedoch darüber hinaus ein erhöhtes Vertrauen in ihre eigenen Urteile (Kruglanski & Fishman, 2009). Außerdem neigen sie dazu, Informationen heuristisch zu verarbeiten, während Menschen mit niedrigem NfCC systematisch verarbeiten (Klein & Webster, 2000; Kossowska & Bar-Tal, 2013).

Kontextbezogener Forschungsstand. Hart et al. (2012) untersuchten NfCC in einer Online-Informationsproblemsituation und stellten fest, dass Personen mit hohem NfCC mehr entscheidungsunterstützende und weniger entscheidungshemmende Informationen suchen. Dieses Ergebnis spiegelt ihren Wunsch wider, schnell eine eindeutige Antwort zu erhalten. Darüber hinaus zeigte Fortier (2016), dass Personen mit hohem NfCC Strategien verwenden, die sie dabei unterstützen, Informationsaufgaben so effizient wie möglich zu erledigen, wie z. B. die Nutzung nur bekannter Informationsquellen oder die Begrenzung der Informationsquellen auf eine überschaubare Anzahl.

Es ist daher naheliegend, dass Unterschiede in NfCC zu einem unterschiedlichen Verhalten in der Online-Umgebung, also bei der Suche und Evaluation von Online-Informationen, führen könnte. Bisher gibt es nur wenig Forschende, die NfCC innerhalb dieses Rahmens und zusätzlich mit mehreren Webseitenmerkmalen gleichzeitig untersucht haben, weshalb dies die Integration von NfCC rechtfertigt.

2.6.3.3. Information Overload

Die Informationsüberlastung (Englisch: Information Overload²⁰; im Folgenden abgekürzt mit IO) beschreibt einen in vielen Fällen unerwünschten Zustand, an dem zu viele Informationen als schädlich empfunden wird: „Information overload occurs when information received becomes a hindrance rather than a help when the information is potentially useful²¹“ (Bawden et al., 1999, S. 249). Es wird hierbei von einem Schwellenpunkt ausgegangen, bei dem mit seiner Überschreitung ein Zustand der Informationsüberlastung eintritt (Roetzel, 2019). Die Ursache für das Erreichen des Punktes können zwei unterschiedliche sein: Zum einen könnten limitierte Ressourcen, wie beispielsweise geringe Zeit oder geringes Budget, dafür verantwortlich sein, dass die bereitgestellten Informationen nicht effizient genutzt werden können (Roetzel, 2019). Zum anderen könnten kognitiv bedingte Ressourcen nicht in ausreichender Menge vorhanden

²⁰ In der Literatur oftmals auch gleichgesetzt mit Cognitive Overload (Belabbes et al., 2023). Die beiden Begrifflichkeiten kommen zwar aus unterschiedlichen Fachgebieten - Kognitionspsychologie (für Einführung s. z. B. Sweller, 1988) und Informationswissenschaft (für Einführung s. z. B. Bawden & Robinson, 2015) - beschreiben jedoch prinzipiell den selben Zustand, nämlich der einer Überlastung aufgrund zu vieler Informationen. Für vorliegende Arbeit wird sich begrifflich für Information Overload entschieden, da sich mit NfC und NfCC bereits zwei Begrifflichkeiten rund um „Kognitionen“ in der Arbeit befinden und Verwechslungen so ggf. ausgeschlossen werden könnten.

²¹ Übersetzung ins Deutsche: „Informationsüberlastung tritt auf, wenn die erhaltenen Informationen eher ein Hindernis als eine Hilfe darstellen, obwohl sie potenziell nützlich sind“ (Bawden et al., 1999, S. 249).

sein, sodass die Informationen überladend wirken (Roetzel, 2019). Wie in Kapitel 2.2.1 bereits vorgestellt, liegt die menschliche Kapazität der Informationsverarbeitung bei etwa sieben Informationen (G. A. Miller, 1956). Dies ist jedoch ein Durchschnittswert, weshalb davon ausgegangen werden kann, dass es interindividuelle Unterschiede hinsichtlich der Informationsverarbeitungskapazität gibt. Da sich die vorliegende Arbeit mit kognitiv-bedingten Unterschieden zwischen Personen befasst, wird die zweite Ursache für eine Informationsüberlastung zu Grunde gelegt. Zusammengefasst wird die Informationsüberlastung zum Zwecke dieser Arbeit folgendermaßen definiert: Eine Informationsüberlastung liegt vor, wenn die Anforderungen an die Informationsverarbeitung größer sind oder werden, als die Kapazität, die Informationen verarbeiten zu können (Eppler & Mengis, 2004, S. 277).

Der Zustand, in einer Entscheidungssituationen aufgrund zu vieler Informationen überfordert zu sein, ist in Anbetracht der langjährigen Geschichte des Menschen als Entscheider und den damals zu treffenden Entscheidungen (z. B. Flucht oder Kampf) nicht neu. Welche Entwicklungen allerdings neu sind und daher auch die Betrachtung von IO in dieser Forschung rechtfertigt, sind zum einen, dass sich mit dem Web 2.0²² die Geschwindigkeit zur Verbreitung von Informationen massiv beschleunigt (B. Oswald, 2013, S. 66; Rainsberger, 2022, S. 23) und weniger Gatekeeper (z. B. Reviewer, Editor etc.) zur Verfügung stehen, die eine Qualität gewährleisten können (Badke, 2021; Koltay, 2017). Gleichzeitig könnten Individuen aber aufgrund digitaler Medien und Informationen erschöpft sein (Kramp & Weichert, 2022, S. 7 f.). Zudem verkürzt die Informationsflut ihre Aufmerksamkeitsspanne (Lorenz-Spreen et al., 2019, S. 6). Das bedeutet, sie geraten schneller und früher an den Schwellenpunkt, an dem eine ausreichend adäquate Informationsverarbeitung gerade noch gewährleistet sein kann, da viel mehr Informationen in kürzester Zeit, aber weniger Kapazität zur Verfügung steht.

Forschungsstand. Personen mit hohem IO erreichen im Vergleich zu Personen mit niedrigem IO früher einen Schwellenwert, bei dem weitere Hinweise überladend sind (Chewning & Harrell, 1990; Roetzel, 2019). Außerdem sind Entscheidungen von Personen, die eine Informationsüberlastung verspüren, von geringerer Qualität sind als die von Personen, die keine Informationsüberlastung empfinden (Chan, 2001; Chervany & Dickson, 1974). IO steht darüber hinaus in negativem Zusammenhang mit der Aufgabenperformanz (Scheiter et al., 2009). Menschen, die aufgrund von z. B. häufig wiederholten oder diversen Informationen eine Informationsüberlastung empfinden, treffen Entscheidungen von geringer Qualität (Hwang & Lin, 1999) und evaluieren kaum kritisch bzw. sind leichtgläubiger (Shenk, 1997). Außerdem zeigt

²² Web 2.0 meint ein Stadium in der Entwicklung des Internets, in dem Nutzer die Inhalte aktiv mitgestalten können, weshalb sie auch als Prosumenten, ein Neologismus aus Produzent und Konsument, bezeichnet werden (*Medienpolitik: Web 2.0*, o. J.).

sich, dass mit zunehmenden IO die Suchstrategien weniger systematisch (Swain & Haka, 2000) und oberflächlicher (Simons, 1988) werden.

Kontextbezogener Forschungsstand. Im Hinblick auf die aktuelle Forschung auf dem Gebiet der Online-Informationsproblemlösung ist festzustellen, dass sich viele Autoren bereits mit Informationsüberlastung und dem IPS-I-Modell beschäftigt haben: Hong und Kim (2020) wiesen in einer Umfrage zu Konsequenzen von Informationsüberlastung nach, dass Personen mit hohem IO Informationen heuristisch verarbeiten im Vergleich zu Personen mit niedrigem IO. Ji et al. (2014) untersuchten den Zusammenhang zwischen Soziodemografie, wahrgenommener Effizienz der Informationssuche, Nutzung von Nachrichtenmedien und IO. Sie fanden heraus, dass vor allem Frauen und junge Menschen häufiger von IO betroffen sind und, dass IO und die wahrgenommene Effizienz der Informationssuche in negativem Zusammenhang stehen (Ji et al., 2014). Durch die Erhöhung des psychologischen Stresses der Verbraucher führt IO zu einer erhöhten Wahrscheinlichkeit, Fehlinformationen zu teilen (Bermes, 2021). Hong und Kim (2020) untersuchten neben Konsequenzen auch Antezedenzen der Informationsüberlastung und zeigten, dass die kognitive Kapazität sowie die Häufigkeit der Online-Nachrichtennutzung und der zwischenmenschlichen Kommunikation IO signifikant vorhersagen. Es wurde also von Seiten ihrer Probanden eine Informationsüberlastung empfunden, wenn sowohl die zwischenmenschliche Kommunikation als auch die Online-Nachrichtennutzung zunehmen und wenn die kognitive Kapazität zur Verarbeitung nicht ausreichend ist (H. Hong & Kim, 2020).

Trotz der Vielzahl an Forschungen zu IO gibt es bislang keine, die zudem viele verschiedene Webseitenmerkmale zur gleichen Zeit mit in Betracht zieht. Da, wie bereits angemerkt, beim Prozess der Online-Informationsverarbeitung Personenmerkmale und Webseitenmerkmale einen Einfluss haben können (Wirth et al., 2007), wird mit Einschluss von IO (sowie NfC und NfCC) und derartig vielen Webseitenmerkmalen ein alltagsnahes und realistisches Szenario abgebildet. Denn Menschen, die online nach Informationen suchen, können ebenfalls unterschiedlich bzgl. IO (sowie NfC und NfCC) veranlagt sein und werden in der Situation gleichzeitig mit vielen Webseitenmerkmalen konfrontiert. Aus diesem Grund wird IO in das Forschungsvorhaben mit einbezogen.

2.7. Forschungsdesiderat

Das vorangegangene Kapitel erläuterte, welche Webseiten- und Personenmerkmale beim Lösen eines Informationsproblems von Relevanz sind und warum. Zudem wurde deutlich, dass der bisherige aktuelle Stand bereits eine Vielzahl an Forschungen hervorbrachte, welche jedoch nicht alle den gleichzeitigen Einfluss von Personen- und Webseitenmerkmalen bei der Such- und Evaluationsstrategie zur Lösung eines Informationsproblems online betrachten:

Forschende thematisierten bereits im Rahmen eines Informationsproblems beispielsweise NfC (Tarchi & Mason, 2020; Wu et al., 2014), NfCC (Coppola et al., 2020; Hart et al., 2012; Kossowska, 2007) oder IO (Bermes, 2021; H. Hong & Kim, 2020; Ji et al., 2014), jedoch wurden diese jeweils separat und ohne Berücksichtigung von Webseitenmerkmalen untersucht. Einen Schritt weiter gingen Fortier und Burkell (2016; 2014), die NfCC und NfC als Prädiktoren für die Informationsproblem-Lösung kombiniert betrachteten, wobei allerdings ebenfalls keine Webseitenmerkmale mit einfließen.

Andere Personenmerkmale, die entweder als Prädiktoren oder als Ergebnisvariablen in bereits existierenden, ähnlichen Forschungen und in Kombination mit Webseitenmerkmalen Anklang fanden, sind beispielsweise Performanz (Brand-Gruwel et al., 2017; Breakstone, Smith, Wineburg, et al., 2021; Metzger et al., 2015; von der Mühlen et al., 2016; Wiley et al., 2009; Yao & Wang, 2023), Vorwissen (Brand-Gruwel et al., 2017; Gerjets et al., 2011; Kiili et al., 2022; Wineburg & McGrew, 2019), Mediennutzung (H. Hong & Kim, 2020; Ji et al., 2014), Selbstwirksamkeit (Andreassen & Bråten, 2013) oder andere kognitive Fähigkeiten (Brennan et al., 2014; Sharit et al., 2008).

Wurden Webseitenmerkmalen untersucht und diese mit beispielsweise NfC kombiniert (Kaufman et al., 1999; Schaewitz et al., 2020; Winter & Krämer, 2012), wurde lediglich ein Merkmal oder eine Merkmalsgruppe (bei Schaewitz et al., 2020 z. B. Nachrichtenmerkmale) betrachtet und gezielt als unabhängige Variable manipuliert. Mehrere gleichzeitig auftretende Webseitenmerkmale, die durch die qualitative Analyse von Kommentaren (Fogg et al., 2003; Fogg, Marshall, Laraki, et al., 2001) oder Tagebucheinträgen (Hilligoss & Rieh, 2008) gefunden wurden, waren meist das Ergebnis der Studien und nicht, wie hier, der mögliche Einflussfaktor. Zudem fand auch bei ihnen keine Kombination mit Personenmerkmalen statt.

Wenn das Untersuchungsdesign ein reales Szenario abdeckte, insofern keine bewusste Manipulation der vorgegebenen Informationen vorsah und die Probanden tatsächliche Live-Webseiten beurteilen sollten, wurden die Personenmerkmale jedoch nicht in Betracht gezogen (T. Hong, 2006; Hope, 2007).

Unterschiede gibt es weiterhin bezüglich der untersuchten Prozessschritte innerhalb der Strategien. So fokussierten sich einige Forschende auf die allgemeine Suche von Informationen (T. Hong, 2006; Kossowska, 2007; Swain & Haka, 2000), die Entscheidungsfindung (Chewning & Harrell, 1990; Hart et al., 2012), die Evaluation der Ergebnisse der Suchergebnisseite (Gerjets et al., 2011; Kattenbeck & Elweiler, 2019; Wu et al., 2014) oder die Evaluation von Webseiten (Breakstone, Smith, Wineburg, et al., 2021; Craig & Blankenship, 2011; Kaufman et al., 1999; Kiili et al., 2022; von der Mühlen et al., 2016; Wineburg & McGrew, 2019). Nur wenige zogen sowohl Schritte aus der Suche als auch der Evaluation in der gleichen

Forschung in Betracht, wie z. B. die freie Suche und Evaluationsstrategie bei der üblichen Recherche im Rahmen des Studiums (Hope, 2007), zu einem vorgegebenen Thema (Brand-Gruwel et al., 2017; T. Hong, 2006) oder zum Bericht einer frei gewählten, täglichen Informationsaktivität (Hilligoss & Rieh, 2008).

Eine weitere Limitation bisheriger Forschungen bezieht sich auf die untersuchten Populationen, die im wesentlichen Studierende (Breakstone, Smith, Wineburg, et al., 2021; Coppola et al., 2020; Craig & Blankenship, 2011; Fortier & Burkell, 2014; Hart et al., 2012; Hope, 2007; Liu, 2004; Scharrer et al., 2021; Wineburg et al., 2020), Schülerinnen und Schüler (Kiili et al., 2022, 2023; Lei et al., 2013; McGrew, 2020; S. Zhang et al., 2011) oder erwachsene Nicht-Akademiker (Kammerer et al., 2015) waren. Oder aber es wurde von vorneherein ein Personenvergleich zwischen Studierenden unterschiedlicher Studiendauer (Chewning & Harrell, 1990; F. Johnson et al., 2016; Lucassen et al., 2013; McGrew et al., 2018), zwischen Novizen und Experten (Brand-Gruwel et al., 2005, 2017; McGrew, 2021a; von der Mühlen et al., 2016; Wineburg & McGrew, 2019) oder Alt und Jung (Chin et al., 2009; Ji et al., 2014) getätigt. Spezifisch Berufsanfänger, direkt nachdem sie ihren Masterabschluss erzielt haben, waren bislang kein Gegenstand einer Untersuchung beim Lösen eines Online-Informationsproblems.

Wie aus den Erläuterungen zur bisherigen Literatur erkennbar, stellt somit vorliegende Forschungsarbeit eine einzigartige Forschung vor, da sich Personen- und Webseitenmerkmale als State-Trait-Wechselwirkung (Steyer et al., 1999) bewusst gewählt und gleichzeitig in einem Modell befinden, welches sowohl die Informationssuch- als auch die Informationsverwertungsstrategie betrachtet. Bislang wurden diese drei gewählten kognitiven Dispositionen weder zusammen noch als unabhängige Variable zur Bestimmung ausgewählter Prozess- und Performanzschritte bei der Problemlösung online untersucht. Ähnlich verhält es sich zu den Webseitenmerkmalen, da diese zum besten Wissen noch nie als Einflussfaktor in solch einer Fülle, welche ein möglichst holistisches Modell darstellen soll, zusammen untersucht wurden. Weiterhin werden alle Webseitenmerkmale, die tatsächlich vorkommen, identifiziert, analysiert und im weiteren Verlauf als unabhängige Variable zur möglichen Vorhersage von Prozess- und Performanzindikatoren betrachtet. Daran anknüpfend ist herauszustellen, dass die Probanden in eine natürliche und reale Problemsituation versetzt werden, was die Authentizität des Verhaltens der Probanden und daraus folgernd auch die Übertragbarkeit der Ergebnisse erhöht. Einzigartig ist auch die Tatsache, dass es sich bei den Probanden um Berufsanfänger und nicht, wie bereits häufig vorkommend, um Studierende oder Schüler handelt.

3. Hypothesenherleitung

Das folgende Kapitel 3 gliedert sich in vier Unterkapitel, die jeweils die direkten (in 3.1 und 3.3) sowie Interaktionseffekte (in 3.2 und 3.4) der Such- und Evaluationsstrategie hypothetisch

aufstellen. Es ist anzumerken, dass die Herleitung der Hypothesen dabei auf bisheriger Literatur aus allgemein-gültigen Kontexten basiert, welche für das eher spezifische Forschungsvorhaben (Kombination aus Personen- und Webseitenmerkmalen) sowie für die Daten aus dem Drittmittelprojekt und der damit verbundenen spezifischen Aufgabenstellung bewusst angewandt werden. Grund für diese deduktive Vorgehensweise ist, dass keine bislang vergleichbaren Forschungen zu einem derartigen aufgaben- bzw. situationsspezifischen Kontext existieren, sodass daraus Erkenntnisse für die hier vorliegende Hypothesenherleitung hätten Anwendung finden können.

Die nachstehende Tabelle 3 gibt einen einführenden Überblick darüber, welche Webseiten- und Personenmerkmale beteiligt, welche Hypothese korrespondierend formuliert sind und welche Modelle zugeordnet werden. Die lateinischen Ziffern I-VI der Modelle wurden korrespondierend den sechs verschiedenen abhängigen Variablen zugeordnet. Die Nummerierung der Hypothesen ist aufsteigend und einheitlich den Unterkapitel gewählt. Die Hypothesen der direkten Effekte werden pro Personenmerkmal blockweise bzgl. der abhängigen Variablen berichtet. D. h. in beispielsweise 3.1.1 werden die Hypothesen zu NfC und einem möglichen Einfluss auf die drei abhängigen Variablen dargestellt, im darauffolgenden Unterkapitel 3.1.2 zu NfCC und den drei abhängigen Variablen etc. Von einer generellen Visualisierung der Hypothesen wurde abgesehen. Hypothesen, die einen non-linearen Effekt oder Interaktionen enthalten, sind jedoch zum besseren Verständnis grafisch dargestellt.

Tabelle 3: Übersicht Modelle und Variablen.

	Modell	Prozessindikatoren			Performanzindikatoren		
		I	II	III	IV	V	VI
UVs	AVs	<i>Spezifizität des Suchbegriffs</i>	<i>Nachrecherche</i>	<i>Weiterverlinkung anklicken</i>	<i>Korrekte Quelleneinschätzung</i>	<i>Kategorie Quelleneinschätzung</i>	<i>Argumentationsstruktur</i>
PM	NfC	I-A (H1a)	II-A (H1b)	III-A (H1c)	IV-A (H5a)	V-A (H5b)	VI-A (H5c; H8; H12)
	IO	I-A (H3a)	II-B (H3b)	III-B (H3c)	IV-A (H7a)	V-A (H7b)	VI-B (H7c; H10; H14)
	NfCC	I-B (H2a)	II-C (H2b)	III-C (H2c)	IV-B (H6a)	V-B (H6b)	VI-C (H6c; H9; H13)
WM	Textgläubigkeit besuchter Webseiten		II-A-C (H4a)	III -A-C (H4b)			
	Medienglaubwürdigkeit Autorenglaubwürdigkeit Textgläubigkeit selektierter Webseiten						VI-A-C (H8-14)

3.1. Direkte Effekte des Suchprozesses

3.1.1. Need for Cognition (NfC) und Suchprozess

In der zweiten Phase des IPS-I-Modells sollten idealerweise Suchtermini definiert und spezifiziert werden, die in eine Suchmaschine eingegeben werden können (Brand-Gruwel et al., 2009). Bei der Eingabe eines spezifischen Suchbegriffs muss die Person die ursprüngliche Aussage umformulieren, was im Vergleich zum einfachen Kopieren der Aufgabenstellung in die Suchmaske kognitiven Aufwand bedeutet. Personen mit hohem NfC investieren mehr kognitive Ressourcen in eine Suchaufgabe (Verplanken et al., 1992) und in die Informationsverarbeitung (Enge et al., 2008) und haben größere Ressourcen zur Aufmerksamkeitszuteilung (Fleischhauer et al., 2010). Größeren kognitiven Aufwand zu betreiben kann daher auch bedeuten, eine Anstrengung mehr zu investieren, die Suchanfrage umzuformulieren und nicht zu kopieren. Aus diesem Grund wird postuliert, dass Personen mit hohem im Vergleich zu niedrigem NfC

H1a: eher spezifische Suchbegriffe verwenden.

Die 3. Phase des IPS-I legt nahe, den Inhalt auf einer zuvor angeklickten Webseite mit anderen Informationen zu elaborieren oder zu kombinieren (Brand-Gruwel et al., 2009). Ein hierfür dienlicher Schritt stellt beispielsweise die Überprüfung der angegebenen Links und Referenzen dar (= Anklicken der Weiterverlinkung i. S. einer Überprüfung), indem diese angeklickt und weiterverfolgt werden.

Außerdem soll laut COR-Konstrukt eine Quelle auf Glaubwürdigkeit untersucht werden, indem zusätzlich online verfügbare Ressourcen genutzt und mit anderen Suchergebnissen abgeglichen werden (Molero et al., 2020; Nagel et al., 2020). Die ursprünglich zuerst besuchte Seite zu verlassen und eine neue Suchanfrage zu starten, um Informationen über genau diese Webseite zu finden oder sich mithilfe des Impressums (= Nachrecherche) über die Webseite zu informieren, könnten dies widerspiegeln.

Beide Aktivitäten stellen hinreichende, aber nicht notwendige zusätzliche kognitiv aufwändige Schritte dar. Da Personen mit hohem NfC (im Vergleich zu niedrigem) mehr Ressourcen in eine Suchaufgabe investieren (Verplanken et al., 1992), somit möglicherweise mit höherer Wahrscheinlichkeit nach (Zusatz-)Informationen suchen und sich diese Zusatzschritte als Anklicken der Weiterverlinkung und Nachrecherchieren darstellen könnten, wird postuliert, dass Personen mit hohem NfC

H1b: eher eine Nachrecherche betreiben.

H1c: eher die Weiterverlinkungen anklicken.

3.1.2. Need for Cognitive Closure (NfCC) und Suchprozess

Bei der Eingabe eines Suchbegriffs kann entweder eine allgemeine Formulierung bzw. eine einfache Kopie der ursprünglichen Aussage oder eine spezifische, umformulierte Suchanfrage getätigt werden. Personen mit hohem NfCC bevorzugen abstrakte Beschreibungen und Kategorien, weil sie sie auf verschiedenste Situationen anwenden können (Fortier & Burkell, 2014; Kruglanski & Fishman, 2009). Gemäß der Übersetzung basiert „abstrakt“ auf „allgemeinen Ideen“ (Oxford Learner's Dictionaries; o.J. - a). Eine abstrakte Beschreibung könnte als eine Verallgemeinerung angesehen werden. Weil eine allgemeine Formulierung für Personen mit hohem NfCC, die abstrakte Beschreibungen bevorzugen, ausreichend ist, werden sie sie vermutlich auch nicht umformulieren. Aus diesem Grund wird postuliert, dass Personen mit hohem NfCC

H2a: eher allgemeine Suchbegriffe verwenden.

Personen mit hohem NfCC ziehen zur Entscheidungsfindung eine geringe Anzahl an Informationsquellen zu Rate und meiden andere eingehenden Informationen (Kruglanski & Fishman,

2009). Sie möchten schnellstmöglich einen Zustand des Abschlusses (Closure) erreichen (Houghton & Grewal, 2000), weshalb davon ausgegangen werden kann, dass sie keine zusätzliche Zeit in weitere Schritte bei einer Informationssuche investieren möchten. Zudem haben sie ein größeres Vertrauen in frühe oder unvollständige Informationen (Kruglanski & Fishman, 2009). Aufgrund dieses Vertrauens ist es naheliegend, dass Personen mit hohem NfCC auch mit der zuerst gefundenen Information zufrieden sind und nicht die Notwendigkeit sehen, eine zusätzliche Informationsquelle anzusteuern. Die Referenzen zu überprüfen durch das Anklicken der Weiterverlinkungen impliziert, eine zusätzliche Informationsquelle zu öffnen und zu verfolgen. Auch die ursprünglich gefundene Quelle nochmals zu überprüfen (= Nachrecherche) in Form beispielsweise einer neuen Suchanfrage, wäre gleichbedeutend mit einem Zusatzschritt, der für Personen mit hohem NfCC aufgrund ihres großen Vertrauens nicht notwendig erscheinen könnte. Aus diesem Grund wird postuliert, dass Personen mit hohem NfCC

H2b: eher weniger Nachrecherche betreiben.

H2c: eher weniger die Weiterverlinkungen anklicken.

3.1.3. Information Overload (IO) und Suchprozess

Personen, die sehr schnell von Informationen überlastet werden und daher einen hohen Wert an Information Overload aufweisen, erreichen ihre kognitive Kapazitätsgrenze aufgrund ihrer limitierten kognitiven Gegebenheiten vergleichsweise früher (Roetzel, 2019). Da Suchaktivitäten zur Lösung eines Informationsproblems kognitiven Aufwand darstellen, könnten Personen mit hohem IO weniger in der Lage sein, zusätzliche kognitive Anstrengungen zu betreiben. Denn das Lesen der Aufgabenstellung und erste Gedanken dazu zu machen, könnte sie bereits überfordern und näher an die Schwelle der Überlastung bringen. Wenn sie zu diesem Zeitpunkt bereits eine Überforderung erreichen, könnten zusätzliche (aber nicht zur Lösung der Aufgabenstellung erforderliche) Schritte, beispielsweise sich gut zu überlegen, wie und welche Suchanfrage formuliert wird, nicht gründlich oder gar nicht ausgeführt werden. Dies könnte wiederum dazu führen, dass lediglich die Fragestellung kopiert oder eine allgemeine Suchanfrage gestellt wird. Aus diesem Grund wird postuliert, dass Personen, die hohes IO haben

H3a: eher allgemeine Suchbegriffe verwenden.

Personen, die schnell bei Informationsproblemen überlastet sind, suchen im Vergleich zu Personen mit niedrigem IO oberflächlicher (Simons, 1988). Außerdem ist ihre Suchstrategie weniger systematisch und die relative Menge an gesuchten Informationen nimmt mit steigendem IO ab (Swain & Haka, 2000). Eine oberflächliche Suche könnte somit auch bedeuten, dass nur das Notwendige getan wird und somit keine Referenzen (= Weiterverlinkungen anklicken) oder

Quellen (= Nachrecherche) überprüft werden. Weil sie zudem früher eine Schwelle erreichen, an der neue Informationen aufgrund ihrer verringerten kognitiven Kapazität überfordernd sind (Roetzel, 2019), versuchen sie nur die eigentliche Aufgabenstellung zu lösen, ohne dass zusätzliche Informationen gesucht oder Referenzen bzw. Quellen geprüft werden. Deshalb wird postuliert, dass Personen, bei denen die Informationsüberlastung schneller eintritt (bzw. ein hohes IO haben)

H3b: eher weniger Nachrecherche betreiben.

H3c: eher weniger die Weiterverlinkungen anklicken.

3.1.4. Einfluss der wahrgenommenen Glaubwürdigkeit besuchter Webseiten

Anhand der theoretischen Ausarbeitung (Kapitel 2.4) wird deutlich, dass innerhalb des Evaluationsprozesses nicht nur Personen- sondern auch Webseitenmerkmale in den Einflussraum fallen können. Auch Webseitenmerkmale²³ können Entscheidungsregeln bzw. Heuristiken aktivieren und dadurch einen Einfluss auf die Glaubwürdigkeitsbeurteilung einer Webseite haben (Wirth, 1999; Wirth & Schweiger, 1999). Die Wahrnehmung eines oder mehrerer Merkmale und ihre Interpretation könnten in diesem Schritt dazu beitragen, dass Probanden die Quellen überprüfen (Nachrecherche) oder weiteren Links folgen. Je nachdem, ob sie bereits aufgrund eines für sie eindeutigen Merkmals zu einer Entscheidung über die (Un-)Glaubwürdigkeit gelangt sind, kann dieser Schritt unterschiedlich ausfallen.

Es wird davon ausgegangen, dass das Aufrufen von glaubwürdigen Seiten, die auch als solche identifiziert wurden, Personen eine Sicherheit gibt, sich nicht weiter mit der Glaubwürdigkeit der Webseite zu beschäftigen. Außerdem ist denkbar, dass das Aufrufen und Erkennen von unglaubwürdigen Seiten Probanden dazu veranlasst, diese direkt aus dem Prozess zu verwerfen, weshalb auch hier keine weiteren Schritte getätigt werden. Lediglich Webseiten, die auf einen ersten Blick nicht klar erkennbar glaubwürdig oder unglaubwürdig sind, könnten Probanden ein Unsicherheitsgefühl geben, sodass zusätzliche Schritte unternommen werden, z. B. Referenzen zu überprüfen, nachzurecherchieren oder Links anzuklicken. Aus diesem Grund wird nachfolgend ein non-linearer Effekt der wahrgenommenen Glaubwürdigkeit auf die abhängigen Variablen postuliert:

H4a: Der Aufruf von wahrscheinlich sehr glaubwürdigen und sehr unglaubwürdigen wahrgenommenen Webseiten bezüglich textlicher, medialer sowie Autorenmerkmale

²³ Webseitenmerkmale können Heuristiken aktivieren, sofern sie wahrgenommen und interpretiert werden. Wie bereits in Kapitel 2.1 beschrieben wird unterstellt, dass die Probanden alle zur Verfügung stehenden Merkmale im Sinne der Prominenz-Interpretations-Theorie (Fogg, 2003) erstens wahrgenommen und zweitens als Glaubwürdigkeitsmerkmal identifiziert und interpretiert haben. Eine Überprüfung der tatsächlichen Wahrnehmung und Interpretation erfolgt nicht. Diese Annahme und hieraus resultierende Einschränkung werden an entsprechender Stelle kritisch diskutiert.

führt zu weniger Nachrecherche, wohingegen der Aufruf von Webseiten, die bezüglich textlicher, medialer und Autorenmerkmale eine mittlere²⁴ subjektive Glaubwürdigkeit aufweisen, zu eher mehr Nachrecherche führt (Abbildung 9).

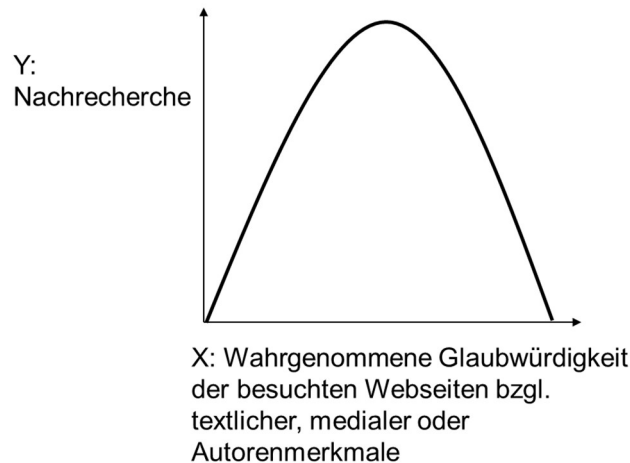


Abbildung 9: Exemplarische Darstellung der Hypothese H4a.

H4b: Der Aufruf von wahrscheinlich sehr glaubwürdigen und sehr unglaubwürdigen wahrgenommenen Webseiten bezüglich textlicher, medialer sowie Autorenmerkmale führt zu eher wenigem Anklicken der Weiterverlinkung, wohingegen der Aufruf von Webseiten, die bezüglich textlicher, medialer und Autorenmerkmale eine mittlere subjektiv wahrgenommene Glaubwürdigkeit aufweisen zu eher mehr Anklicken der Weiterverlinkung führt (analoge grafische Darstellung wie in Abbildung 9).

3.2. Interaktionseffekte in der Suchstrategie

Da Individuen die Glaubwürdigkeit beurteilen, sind an diesem Beurteilungsprozess folglich auch ihre individuellen Prädispositionen beteiligt, weshalb ein Einfluss von NfC, NfCC und IO denkbar ist.

Personen mit hohem NfC evaluieren Informationen genauer bzgl. der Qualität (Cacioppo et al., 1996). Wie eingangs beschrieben, ist die Glaubwürdigkeit ebenfalls ein subjektiv wahrnehmbares Qualitätsmerkmal (siehe Kapitel 2.1.) Das führt zur Folgerung, dass Personen mit hohem NfC Informationen auch bzgl. der wahrgenommenen Glaubwürdigkeit gründlicher beurteilen. Hinzu kommt, dass eben diese Personen Informationen auch systematischer verarbeiten (Cacioppo et al., 1996; Chaiken, 2014). Daraus kann geschlossen werden, dass Personen mit hohem NfC mit höherer Wahrscheinlichkeit glaubwürdige von unglaubwürdigen Webseiten unterscheiden können, weil sie diese systematischer verarbeiten, was bei

²⁴ Das Wort „mittlere“ besagt in diesem Fall, dass die Glaubwürdigkeit oder Unglaubwürdigkeit nicht direkt erkennbar und somit unklar ist.

Personen mit niedrigem NfC nicht der Fall ist, da sie Informationen heuristisch verarbeiten. Da sich zuletzt genannte Gruppe aufgrund der heuristischen Verarbeitung schwerer tun könnte, glaubwürdige von unglaubwürdigen zu unterscheiden, werden sie wenige Unterschiede in der Glaubwürdigkeitsbeurteilung wahrnehmen und folglich auch nicht unterschiedlich mit den gefundenen Informationen umgehen. Das führt bei Personen mit niedrigem NfC dazu, dass sie seltener nachrecherchieren (H1a). Personen mit hohem NfC werden hingegen gemäß dem non-linearen Effekt und in Abhängigkeit von der wahrgenommenen Glaubwürdigkeit unterschiedlich mit den Informationen umgehen, da sie glaubwürdige von unglaubwürdigen besser unterscheiden könnten. Daraus resultiert folgende Hypothese:

Personen mit hohem NfC neigen zu weniger Nachrecherche (H4a) und klicken seltener weiterführende Links an (H4b) bei sehr glaubwürdigen und sehr unglaubwürdigen Webseiten (non-linearer Effekt). Personen mit niedrigem NfC werden ungeachtet der Glaubwürdigkeit weniger nachrecherchieren (H1b) und Links anklicken (H1c) (linearer Effekt) (Abbildung 10).

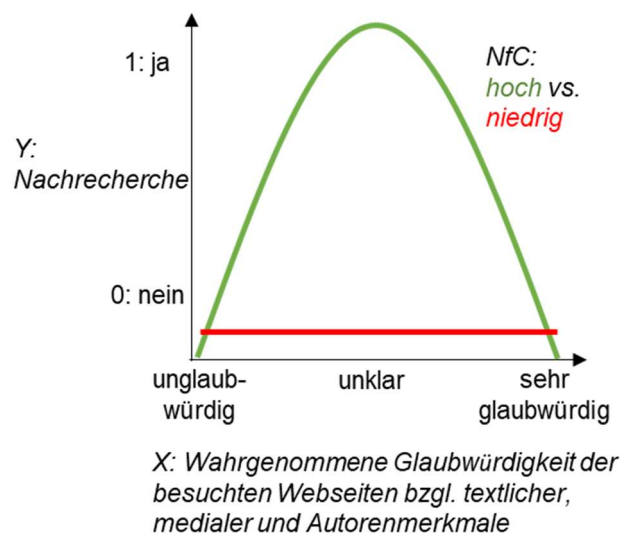


Abbildung 10: Exemplarische Abbildung für Nachrecherche und NfC.

Personen mit hohem NfCC neigen dazu, Suchprozesse frühzeitig zu beenden und greifen dazu auf Hinweisreize z. B. für die Einschätzung einer Information, die dafür dienlich sind, zurück (Dringlichkeitstendenz) (Kruglanski & Fishman, 2009). Dieser Abschluss wird dann so lange wie möglich aufrecht erhalten (Beständigkeitstendenz), indem z. B. keine neuen Informationen berücksichtigt werden oder sich an altem Wissen festgehalten wird, damit eine neue, etwaige Unsicherheit auslösende Situation vermieden wird (Kruglanski & Fishman, 2009). Wenn sie einer Situation ausgesetzt sind, in der sie Webseiten hinsichtlich ihrer Glaubwürdigkeit beurteilen müssen, werden sie sich daher auf schnell greifbare Hinweise stützen (z. B. Oberflächenmerkmale) und diese nicht erneut hinterfragen. Hinzu kommt, dass Personen mit

hohem NfCC ein großes Vertrauen in frühe oder gar unvollständige Informationen aufweisen (Kruglanski & Fishman, 2009). Dies bedeutet somit, dass, sobald sie zu einer ersten Evaluation oder Meinung über die Glaubwürdigkeit einer Webseite gelangt sind, sie bei dieser Evaluation aus Gründen des hohen Vertrauens und der Aufrechterhaltung einer Gewissheitssituation bleiben, selbst wenn dies auf unvollständiger Erkenntnis basiert. Weitere Links anzuklicken oder Quellen durch Nachrecherche zu überprüfen, stellen einen zusätzlichen Schritt dar, wodurch jedoch eine Unsicherheitssituation hervorgerufen werden könnte. Aus diesem Grund werden Personen mit hohem NfCC dies vermeiden wollen und ungeachtet der wahrgenommenen Glaubwürdigkeit nicht nachrecherchieren oder Links weiterverfolgen (linearer Effekt). Personen mit niedrigem NfCC werden abhängig von der wahrgenommenen Glaubwürdigkeit unterschiedlich mit den Informationen umgehen (non-linearer Effekt). D. h., wenn sie eine mittelmäßige oder unklare Glaubwürdigkeit erkennen, werden sie einen Zusatzschritt (nachrecherchieren oder Link anklicken) machen, weil sie keinerlei Probleme mit Unsicherheit oder Ambiguität haben. Aus diesem Grund wird postuliert, dass:

Personen mit niedrigem NfCC neigen zu weniger Nachrecherche (H4a) und klicken seltener weiterführende Links an (H4b) bei sehr glaubwürdigen und sehr unglaubwürdigen Webseiten (non-linearer Effekt). Personen mit hohem NfCC werden ungeachtet der Glaubwürdigkeit weniger nachrecherchieren (H2b) und Links anklicken (H2c) (linearer Effekt) (Abbildung 11)

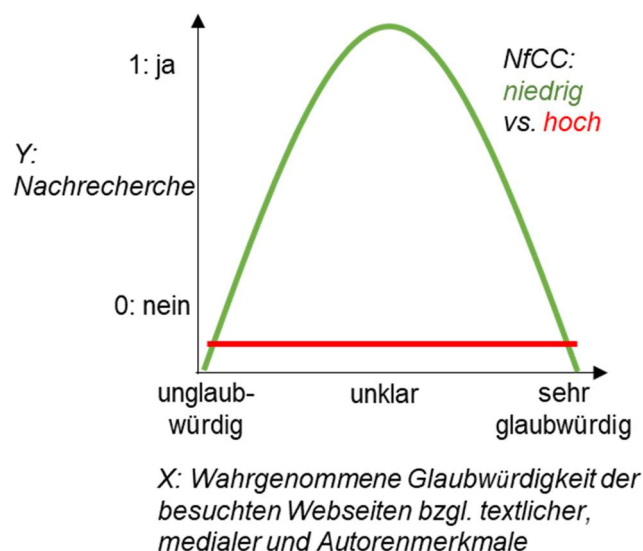


Abbildung 11: Exemplarische Abbildung für Nachrecherche und NfCC.

Personen mit hohem IO evaluieren weniger kritisch (Shenk, 1997). Eine kritische Evaluation von Informationen liegt z. B. dann vor, wenn Personen Quellen oder Referenzen überprüfen. Selbst wenn die Probanden, die hohes IO aufweisen, die aufgerufenen Webseiten in einem ersten Schritt evaluiert haben (weil das beispielsweise die Aufgabenstellung erfordert),

hinterfragen sie aufgrund ihrer Tendenz zur heuristischen Informationsverarbeitung (H. Hong & Kim, 2020) möglicherweise kaum die Glaubwürdigkeit oder Qualität der Webseiten. Denn Informationen oder die Glaubwürdigkeit zu hinterfragen ist gleichbedeutend mit einer systematischen und genauen Auseinandersetzung. Sie sind möglicherweise weniger in der Lage, die Glaubwürdigkeit von Webseiten zu erkennen oder einzuschätzen, weil sie die Informationen aufgrund ihrer empfundenen Überlastung nur heuristisch verarbeiten können. Aus diesem Grund werden sie ungeachtet der Glaubwürdigkeit keine Ressourcen haben nachzurecherchieren oder Links zu verfolgen (linearer Effekt). Personen, die keinen oder weniger IO empfinden, werden bei unterschiedlichen Glaubwürdigkeiten und in Abhängigkeit dieser unterschiedlich mit den Informationen umgehen und möglicherweise nachrecherchieren oder Links weiterverfolgen (non-linearer Effekt). Aus diesem Grund wird postuliert, dass

Personen mit niedrigem IO neigen zu weniger Nachrecherche (H4a) und klicken seltener weiterführende Links an (H4b) bei sehr glaubwürdigen und sehr unglaubwürdigen Webseiten (non-linearer Effekt). Personen mit hohem IO werden ungeachtet der Glaubwürdigkeit weniger nachrecherchieren (H3b) und Links anklicken (H3c) (linearer Effekt) (Abbildung 12)

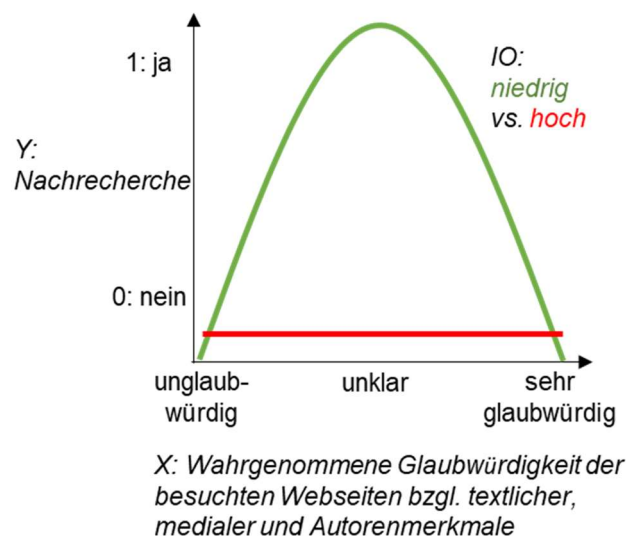


Abbildung 12: Exemplarische Abbildung für Nachrecherche und IO.

3.3. Direkte Effekte im Evaluationsprozesses

Für die direkten Zusammenhänge mit der dritten Variablen im Evaluationsprozess Argumentationsstruktur und den Personenmerkmalen sowie den Interaktionseffekten wird davon ausgegangen, dass lediglich die Textglaubwürdigkeit der selektierten Webseiten von Relevanz sind. Grund hierfür ist, dass der Einfluss von Webseiten, die bezüglich medialer oder Autorenmerkmale eine hohe oder niedrige Glaubwürdigkeit aufweisen, nicht gegeben sein kann. Denn die Argumentationsstruktur wird anhand der verfassten Texte (Stellungnahme) bewertet, die

auf inhaltlichen Aspekten der Webseiten basieren. Mediale Merkmale (z. B., ob die Webseite gut strukturiert ist) oder Autorenmerkmale (z. B., ob der Autor angegeben ist) leisten keinen inhaltlichen Mehrwert, sodass folglich kein Einfluss postuliert werden kann.

3.3.1. Need for Cognition (NfC) und Evaluationsprozess

In der 4. Phase des IPS-I-Modells selektieren Personen idealerweise solche Informationen, die sie anhand von Qualität und Nützlichkeit zuvor als gut bewertet haben (Brand-Gruwel et al., 2009). Ob eine Webseite korrekterweise als qualitativ hochwertig bzw. nützlich oder nicht hochwertig bzw. nicht nützlich eingeschätzt werden konnte, kann von dem Individuum und den entsprechenden kognitiven Fähigkeiten abhängen.

Da Personen mit hohem NfC mehr Ressourcen bzw. kognitive Anstrengung in die Informationsverarbeitung (Enge et al., 2008) oder in die Informationssuche (Verplanken et al., 1992) investieren, könnten sie bei der Bewertung einer Quelle nach mehr Details suchen, was zu einem größeren Erfolg bei der Unterscheidung zwischen guten und schlechten Quellen führen könnte. Zudem ist gut belegt, dass hohes NfC eine bessere Aufgabenperformanz hervorbringen (Coutinho, 2006; Coutinho et al., 2005) und ein höheres Erinnerungs- und Wiedererkennungsvermögen hervorrufen kann (Diehl & Wyrick, 2015). Deshalb können Personen mit hohem NfC auch besser in dieser Aufgabe sein, die darin besteht, die Webseiten hinsichtlich ihrer Qualität (also Glaubwürdigkeit) korrekt einzuschätzen. Korrekte Einschätzung bedeutet, glaubwürdige Webseiten als korrekt glaubwürdig oder unglaubwürdige Webseiten als korrekt unglaubwürdig zu erkennen. Es wird daher postuliert, dass Personen mit hohem NfC

H5a: bei der Bewertung der Korrektheit der Quellen besser abschneiden, d. h. glaubwürdige Quelle korrekt als glaubwürdig und unglaubwürdige Quellen korrekt als unglaubwürdig einschätzen werden.

Da Personen mit hohem NfC Freude an kognitiver Anstrengung empfinden (Cacioppo & Petty, 1982) und bereit sind, mehr Ressourcen in die Informationsverarbeitung zu investieren (Enge et al., 2008), könnten sie sich bei der Analyse von Hinweisen nicht nur auf den ersten Eindruck einer Webseite verlassen, sondern auch mehr Zeit und Mühe in die Suche nach stichhaltigeren Beweisen für die Glaubwürdigkeit einer Seite (z. B. textliche bzw. semantische Merkmale) investieren. Außerdem ist bekannt, dass Personen mit hohem NfC Informationen systematisch verarbeiten (Cacioppo et al., 1996; Chaiken, 2014). Mehr kognitive Ressourcen zu investieren und Informationen systematisch zu verarbeiten, erlaubt möglicherweise vermehrt das Erkennen von textlichen Hinweisen (z. B. Argumentqualität), die eine systematische Informationsverarbeitung voraussetzen. Das daraus resultierende häufigere Erkennen der semantischen Merkmale könnte folglich dazu führen, dass sie sie selbst in höherem Maße verarbeiten und

somit zur Beantwortung der Aufgabe heranziehen. Deshalb wird postuliert, dass Personen mit hohem NfC

H5b: eher systematisch-kategorisierte Begründungen nennen.

Menschen mit hohem NfC genießen sowohl die kognitive Anstrengung (Cacioppo & Petty, 1982) als auch die Argumentation (Sanders et al., 1992). Sie bringen komplexere und elaboriertere Argumente hervor (Nussbaum, 2005). Dadurch können sie eine Aufgabe so lange fortsetzen, bis sie das Gefühl haben, ihr Bestes gegeben zu haben. Da hier davon ausgegangen wird, dass elaborierte Argumente auch logisch nachvollziehbar und gut strukturiert sind, kann angenommen werden, dass Personen mit hohem NfC

H5c: eher eine gute Argumentationsstruktur aufweisen.

3.3.2. Need for Cognitive Closure (NfCC) und Evaluationsprozess

Personen mit hohem NfCC verarbeiten Informationen heuristisch (Kossowska & Bar-Tal, 2013). Eine Quelle korrekt hinsichtlich ihrer Qualität einschätzen zu können, bedarf einer ausgiebigen und systematischen Evaluation. Da Personen mit hohem NfCC schnell frühe (und somit oberflächliche) Hinweisreize ergreifen, um direkt zu einem Urteil zu gelangen und zudem zusätzliche Informationen, die möglicherweise die zuerst gefundenen korrumpieren, vermeiden (Kruglanski & Fishman, 2009), werden sie aller Wahrscheinlichkeit nach keine ausreichende und vollumfängliche Evaluation einer Quelle getätigt haben. Wenn sie daher nicht alle zugrunde liegenden Informationen zur Beurteilung der Quellenqualität zu Rate gezogen haben, werden sie möglicherweise zu einer inkorrekten Einschätzung kommen und folglich schlechter bei der Korrektheit der Quelleneinschätzung abschneiden. Aus diesem Grund wird postuliert, dass Personen mit hohem NfCC

H6a: bei der korrekten Bewertung von Quellen schlechter abschneiden, d. h. glaubwürdige Quelle inkorrekt als unglaubwürdig und unglaubwürdige Quellen inkorrekt als glaubwürdig einschätzen werden.

Personen mit hohem NfCC bevorzugen es, möglichst schnell Gewissheit zu erlangen (Kruglanski & Fishman, 2009). Sie verarbeiten Informationen heuristisch (Kossowska & Bar-Tal, 2013). Schnelle Gewissheit erreichen sie beispielsweise dadurch, dass sie Hinweise wahrnehmen und schnell und somit heuristisch beurteilen. Das bedeutet wiederum, dass Personen mit hohem NfCC weniger Begründungen angeben, die auf einer ausgiebigen und systematischen Beurteilung stützen. Aus diesem Grund wird postuliert, dass Personen mit hohem NfCC

H6b: eher weniger systematisch-kategorisierte Begründungen nennen.

Personen mit hohem NfCC haben ein höheres Vertrauen in frühe oder unvollständige Informationen (Kruglanski & Fishman, 2009) und urteilen voreilig auf der Grundlage nicht eindeutiger Beweise (Kruglanski & Webster, 1996). Personen mit hohem NfCC könnten daher auch zufriedener und vertrauensvoller mit ihrer eigens verfassten Argumentation sein, die auf unvollständigen Informationen basiert oder zu voreilig auf der Grundlage nicht eindeutiger Beweise getätigt wurde und somit selbst unvollständig ist. Sie vermeiden möglicherweise zusätzliche Schritte, wie z. B. die Gegenkontrolle, ob ein geschriebener Text logisch nachvollziehbar oder gut strukturiert ist, weil sie dadurch ihren erwünschten Zustand der Gewissheit gefährden könnten („seize“ und „freeze“) (Kruglanski & Fishman, 2009). Daher wird postuliert, dass Personen mit hohem NfCC

H6c: eher eine schlechte Argumentationsstruktur aufweisen.

3.3.3. Information Overload (IO) und Evaluationsprozess

Personen mit hohem IO suchen oberflächlich (Simons, 1988) und mit der Zeit verhältnismäßig weniger Informationen (Swain & Haka, 2000). Sie verwenden zur Entscheidungsfindung eine geringere Anzahl an Informationen im Vergleich zu Menschen, die wenig oder kein IO empfinden (Chewning und Harrel, 1990), weshalb davon ausgegangen werden kann, dass sie auch eine geringere Anzahl an Cues wahrnehmen, da dies überhaupt eine Voraussetzung zur Verwendung darstellt. Zudem ist bekannt, dass unter hohem IO (operationalisiert durch disaggregierte Informationsdarstellungen vs. zusammengefasste Informationsdarstellungen) die Qualität von Entscheidungen, die Personen treffen (Chervany & Dickson, 1974) und weiterhin auch die Aufgabenperformanz (Scheiter et al., 2009) abnehmen. Somit könnte die geringe Anzahl an wahrgenommenen Cues eine unzureichende Grundlage darstellen, die Qualität einer Quelle korrekt einzuschätzen, zumal die Entscheidungsqualität ebenso verringert sein könnte. Es wird somit postuliert, dass Personen, bei denen die Informationsüberlastung schneller eintritt (ein hohes IO haben)

H7a: bei der korrekten Bewertung von Quellen schlechter abschneiden, d. h. glaubwürdige Quelle inkorrekt als unglaubwürdig und unglaubwürdige Quellen inkorrekt als glaubwürdig einschätzen werden

Personen mit hohem IO betreiben eine oberflächliche Suche (Simons, 1988), verarbeiten Informationen und die entsprechenden Cues heuristisch (H. Hong & Kim, 2020) und suchen mit der Zeit verhältnismäßig weniger Informationen (Swain & Haka, 2000). Sie verwenden eine geringere Anzahl an Cues im Vergleich zu Menschen, die wenig oder keinen IO empfinden (Chewning & Harrel, 1990), weshalb davon ausgegangen werden kann, dass sie auch eine geringere Anzahl an Cues wahrnehmen. Weil Personen mit hohem IO daher tendenziell heuristisch verarbeiten, nur auf wenige Merkmale achten und wahrnehmen, können sie auch nur

diese in ihren Begründungen darstellen. Zudem werden sie aufgrund der heuristischen Verarbeitung auch nur heuristisch-kategorisierte Begründungen angeben, weshalb postuliert wird, dass Personen, bei denen die Informationsüberlastung schneller eintritt (ein hohes IO haben)

H7b: eher weniger systematisch-kategorisierte Begründungen nennen.

Personen mit hohem IO evaluieren Informationen nicht ausreichend kritisch (Shenk, 1997) und ihre Entscheidungsqualität für eine Information nimmt ab (Chan, 2001), wenn z. B. die Informationsdiversität erhöht ist (Hwang & Lin, 1999), was in einer Online-Suchumgebung der Fall ist. Sich für eine ausreichende Menge an Informationen zu entscheiden, die eine gute Informationsgrundlage für eine Argumentation liefern könnte, ist somit bei Personen mit hohem IO unwahrscheinlich. Zudem könnten sie aufgrund der schnell eintretenden Informationsüberlastung im Vergleich zu Personen mit niedrigem IO weniger fähig sein, gut strukturierte und logisch zusammenhängende Argumentationen zu verfassen. Aus diesem Grund wird postuliert, dass Personen, bei denen die Informationsüberlastung schneller eintritt (ein hohes IO haben)

H7c: eher eine schlechte Argumentationsstruktur aufweisen.

3.3.4. Need for Cognition (NfC) und wahrgenommene Glaubwürdigkeit selektierter Webseiten

Da Personen mit hohem NfC mehr Aufwand in die Informationsverarbeitung (Enge et al., 2008) und folglich von Aufgaben stecken und weil sie Informationen systematischer verarbeiten (Cacioppo et al., 1996; Chaiken, 2014), könnten sie im Vergleich zu Personen mit niedrigem NfC solche Webseiten wahrgenommen und folglich selektiert haben, deren textliche Merkmale mehr zu einer positiven Glaubwürdigkeitsbeurteilung beitragen. Es wird somit postuliert, dass Personen mit hohem NfC

H8: Webseiten selektieren, die bezüglich textlicher Merkmale eine höhere Glaubwürdigkeit aufweisen.

3.3.5. Need for Cognitive Closure (NfCC) und wahrgenommene Glaubwürdigkeit selektierter Webseiten

Personen mit hohem NfCC weisen ein geringeres Maß an Genauigkeit bei der Analyse von Beweisen auf (Klein & Webster, 2000). Sie überprüfen weniger die Informationen und lassen sich häufiger von oberflächlichen Hinweisreizen beeinflussen (Klein & Webster, 2000). Weil sie ungenauer sind bei der Analyse von Beweisen, wie beispielsweise Online-Informationen, und zudem sich häufiger von Hinweisreizen beeinflussen lassen, könnten sie sich zur Beurteilung der Glaubwürdigkeit überwiegend auf oberflächliche Hinweisreize stützen, die keine ausreichend gute Grundlage zur Beurteilung der Glaubwürdigkeit darstellt. Deshalb könnten sie

im Vergleich zu Personen mit niedrigem NfCC solche Webseiten wahrgenommen und folglich selektiert haben, deren textliche Merkmale weniger zu einer Glaubwürdigkeitsbeurteilung beitragen. Deshalb wird die Hypothese aufgestellt, dass Personen mit hohem NfCC

H9: Webseiten selektieren, die bezüglich textlicher Merkmale eine geringere Glaubwürdigkeit aufweisen.

3.3.6. Information Overload (IO) und wahrgenommene Glaubwürdigkeit selektierter Webseiten

Personen mit hohem IO stoßen schnell an ihre kognitiven Kapazitätsgrenzen und evaluieren Informationen kaum kritisch (Shenk, 1997). Außerdem nimmt die Qualität der Entscheidungen mit zunehmender Informationsüberlastung ab (Chewning & Harrell, 1990), u. a. auch ausgelöst durch zu große Diversität oder Wiederholung der Informationen (Hwang & Lin, 1999). Personen mit hohem IO neigen also dazu, Entscheidungen von geringer Qualität zu treffen. Da die Unterscheidung zwischen glaubwürdigen und unglaubwürdigen Quellen eine kritisch zu hinterfragende, kognitive Entscheidung ist, die innerhalb einer bestimmten Zeitspanne getroffen werden muss, kann aufgrund der obigen Ausführungen die Hypothese aufgestellt werden, dass Personen mit hohem IO

H10: Webseiten selektieren, die bezüglich textlicher Merkmale eine geringere Glaubwürdigkeit aufweisen

3.3.7. Wahrgenommene Glaubwürdigkeit selektierter Webseiten und Performanzindikatoren

Eine Argumentation sollte gut strukturiert und logisch zusammenhängend sein (Molerov et al., 2020). Damit einhergehend sollte eine gute Argumentationsstruktur die grundlegenden Elemente einer Argumentation beinhalten, d. h. vollständig sein. Nach Toulmin (2003) ist eine Argumentation vollständig, wenn sie These, Begründung und Beleg beinhaltet. Webseiten, die aufgrund textlicher Merkmale eine hohe Glaubwürdigkeit aufweisen, sprich Quellen, die z. B. transparent, verständlich, aktuell und faktisch korrekt sind, sowie strukturierte Informationen und Quellenangaben aufweisen, bieten eine bessere Informationsgrundlage als Webseiten, die bezogen auf diese Merkmale eine geringe wahrgenommene Glaubwürdigkeit aufweisen. Eine bessere Informationsgrundlage zeichnet sich dann beispielsweise dadurch aus, dass unter der Prämisse der inhaltlichen Richtigkeit alle notwendigen Informationen enthalten sind, was somit einer Vollständigkeit entspricht. Im Falle von Primärquellen, die eigene Studien beschreiben und somit Informationen aus erster Hand liefern, beinhalten diese Quellen mit größerer Wahrscheinlichkeit vollständigere Informationen im Vergleich zu Sekundärquellen, was folglich eine fundierte Grundlage für die Argumentation darstellt. Denn Sekundärquellen werden in der Regel die Primärquelle nur insoweit nutzen bzw. sich darauf referieren, als dass

dies für die beabsichtigte Nachricht förderlich ist. Im Gegensatz dazu kann aus akademischer Sicht bei Primärquellen ein Anspruch auf Vollständigkeit zumindest unterstellt werden. Weiterhin könnten gut strukturierte Informationen als weiteres Beispiel für ein textliches Merkmal dazu beitragen, dass die Vollständigkeit gewährleistet wird, da durch eine gute, vorgegebene Struktur dem Probanden das thematische Durchdringen erleichtert wird, was möglicherweise der eigenen Argumentation zugutekommt. Ebenso können Webseiten oder Informationen, die transparenter sind und Quellenangaben und somit mehr Hintergrundinformationen beinhalten, mehr Belege und Begründungen liefern, die in der eigenen Argumentation selbst als entsprechende Begründungen verwendet werden können und somit zur Vollständigkeit beitragen. Durch die Selektion eben solcher Webseiten mit den entsprechenden hier exemplarisch angeführten Merkmalen wird den Probanden eine gute Ausgangsbasis dargeboten, weshalb postuliert wird:

H11: Die Selektion von Webseiten, die hinsichtlich textlicher Merkmale eine hohe Glaubwürdigkeit aufweisen bzw. suggerieren, führt dazu, dass die ausgewählten Webseiten hinsichtlich der einzelnen Merkmale eine bessere Informationsgrundlage für eine vollständige, d. h. bessere Argumentationsstruktur darstellen.

3.4. Interaktionseffekte im Evaluationsprozess: Einfluss der wahrgenommenen Glaubwürdigkeit selektierter Webseiten

Moderation von NFC auf die Glaubwürdigkeitswahrnehmung selektierter Webseiten hinsichtlich textlicher Merkmale und Argumentationsstruktur. Der Effekt, dass die Selektion von Webseiten, die hinsichtlich textlicher Merkmale eine hohe Glaubwürdigkeit aufweisen, zu einer besseren Argumentationsstruktur führt, wird durch NfC in dem Sinne moderiert, als dass hohes NfC diesen Effekt verstärkt, da erstens Personen mit hohem NfC aufgrund ihrer systematischen Verarbeitung tendenziell mehr angeregt (i. S. v. getriggert) werden, Webseiten zu selektieren, die als glaubwürdiger wahrgenommen werden und dies als Idealbild dienlich sein kann, woraus eine bessere Informationsgrundlage und somit Argumentationsstruktur folgen kann. Zweitens werden bei hohem NfC durch die systematische Informationsverarbeitung die Informationen auf der selektierten Webseite besser verarbeitet und möglicherweise auch tiefergehend verstanden, in dem Sinne, dass die Informationen aufgrund des Verständnisses auch besser in eigenen Worten wieder gegeben werden können und/oder auf die vorgegebene Aufgabe transferiert werden könnte. Folglich kann hieraus eine gut strukturierte und vollständige Argumentation resultieren. Im Gegensatz dazu verarbeiten Personen mit niedrigem NfC Informationen heuristisch, d. h. vergleichsweise mit hohem NfC werden sie einerseits Webseiten mit einer geringeren Glaubwürdigkeit selektieren, weshalb ihnen eine Grundlage fehlt für die gute Argumentationsstruktur, sodass in Summe der Effekt hierdurch abgeschwächt wird.

Weiterhin werden sie aufgrund der heuristischen Informationsverarbeitung die Informationen weniger gut verarbeiten und verstehen, sodass ihnen möglicherweise relevante Informationen fehlen, die daher nicht zu einer vollständigen und guten Argumentationsstruktur beitragen können.

Aus diesem Grund wird postuliert, dass

H12: in Abhängigkeit des NfC der Effekt der wahrgenommenen Glaubwürdigkeit der selektierten Webseiten auf die Argumentationsstruktur variiert) (Abbildung 13).

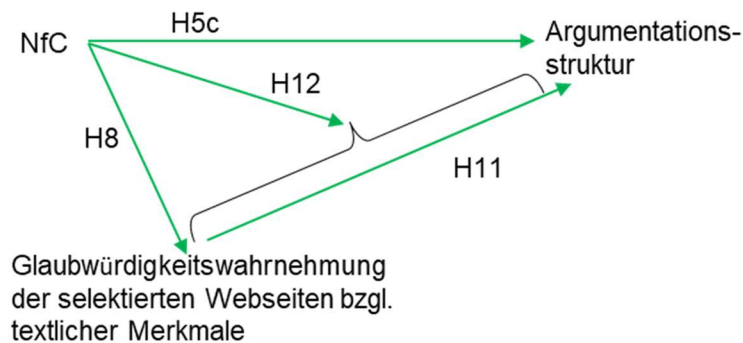


Abbildung 13: Darstellung der Hypothesen aus Modell VI-A.

Moderation von NfCC auf die Glaubwürdigkeitswahrnehmung selektierter Webseiten hinsichtlich textlicher Merkmale und Argumentationsstruktur. Der Effekt, dass die Selektion von Webseiten, die hinsichtlich textlicher Merkmale eine hohe Glaubwürdigkeit aufweisen, zu einer besseren Argumentationsstruktur führt, wird durch NfCC moderiert in dem Sinne, als dass hohes NfCC diesen Effekt abschwächt. Begründet wird dies damit, dass Personen mit hohem NfCC aufgrund ihrer heuristischen Verarbeitung (Kossowska & Bar-Tal, 2013) generell weniger Webseiten als glaubwürdig wahrnehmen, diese selektieren und dies somit unzureichend als Idealbild dienlich sein kann, woraus eine schlechtere Informationsgrundlage und somit Argumentationsstruktur folgen kann. Hinzu kommt, dass Personen mit hohem NfCC Informationen heuristisch verarbeiten, sodass die gefundenen Informationen auf der Webseite möglicherweise nur unzureichend verarbeitet und verstanden werden könnten. Folglich werden sie aufgrund des geringeren Verständnisses nicht alle notwendigen und als relevant erachteten Informationen in der Argumentation anbringen können, was somit zu einer Unvollständigkeit der Argumentation führen könnte. Personen mit niedrigem NfCC verarbeiten Informationen systematisch, d. h., sie werden Webseiten mit einer höheren Glaubwürdigkeit selektieren, weshalb ihnen eine gute Grundlage für eine gute Argumentationsstruktur dargeboten wird, sodass in Summe der Effekt hierdurch verstärkt wird. Zweitens werden bei niedrigem NfCC durch die systematische Informationsverarbeitung die Informationen auf der selektierten Webseite besser verarbeitet und möglicherweise auch tiefergehend verstanden, in dem Sinne, dass die Informationen aufgrund des Verständnisses auch besser in eigenen Worten wieder gegeben

werden können. Folglich könnte hieraus eine gut strukturierte und vollständige Argumentation resultieren. Aus diesem Grund wird postuliert, dass

H13: in Abhängigkeit des NfCC der Effekt der wahrgenommenen Glaubwürdigkeit der selektierten Webseiten auf die Argumentationsstruktur variiert (Abbildung 14)

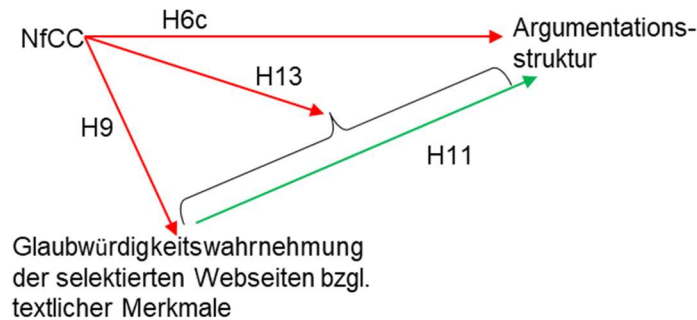


Abbildung 14: Darstellung der Hypothesen aus Modell VI-C.

Moderation von IO auf die Glaubwürdigkeitswahrnehmung selektierter Webseiten hinsichtlich textlicher Merkmale und Argumentationsstruktur. Der Effekt, dass die Selektion von Webseiten, die hinsichtlich textlicher Merkmale eine hohe Glaubwürdigkeit aufweisen, zu einer besseren Argumentationsstruktur führt, wird durch IO moderiert in dem Sinne, als dass hohes IO diesen Effekt abschwächt. Da Personen mit hohem IO aufgrund ihrer heuristischen und oberflächlichen Verarbeitung (H. Hong & Kim, 2020) generell weniger Webseiten als glaubwürdig wahrnehmen und diese selektieren könnten und somit unzureichend als Idealbild nutzen könnten, könnte hieraus eine schlechtere Informationsgrundlage und somit Argumentationsstruktur folgen. Weiterhin verarbeiten Personen, die schnell IO empfinden, Information heuristisch, so dass sie die gefundenen Informationen möglicherweise nicht ausreichend gut durchdringen können. Wenn daraus ein mangelndes Verständnis resultiert, werden sie womöglich die gelesenen Informationen nicht gut strukturiert wieder gegeben können. Personen mit niedrigem IO verarbeiten Informationen systematisch, d. h., sie werden Webseiten mit einer höheren Glaubwürdigkeit selektieren, weshalb ihnen eine gute Grundlage für eine gute Argumentationsstruktur dargeboten wird, sodass in Summe der Effekt hierdurch verstärkt wird. Außerdem könnten Personen mit niedrigem IO aufgrund ihrer systematischen Informationsverarbeitung die Informationen besser und tiefergehend verstanden haben, sodass sie daraus möglicherweise die Informationen auch besser wiedergeben und für die begründete Argumentation verwenden könnten. Aus diesem Grund wird postuliert, dass

H14: in Abhängigkeit des IO der Effekt der wahrgenommenen Glaubwürdigkeit der selektierten Webseiten auf die Argumentationsstruktur variiert (Abbildung 15)

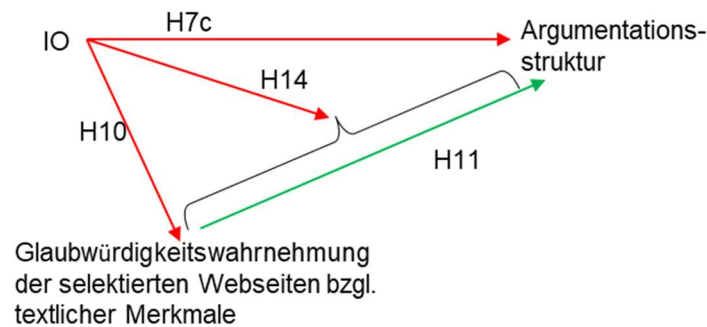


Abbildung 15: Darstellung der Hypothesen aus Modell VI-C.

4. Methodik

Die Erhebung und Analyse von Daten mit unterschiedlichen Datentypen in einer einzigen Studie kann als Mixed-Methods-Ansatz bezeichnet werden (Schreier & Odağ, 2020). Sie haben den Vorteil, dass sie es ermöglichen komplexere Forschungsfragen zu beantworten und ein vollständigeres Bild des zu untersuchenden Forschungsgegenstandes zu liefern (T. Lund, 2012). Da zum vorliegenden Forschungszweck Personenmerkmale (statische Daten, d. h. Trait) und Webseitenmerkmale (dynamische Daten, d. h. State) sowie Performanz- und Prozessfaktoren (ebenfalls dynamische Daten) kombiniert werden und diese aufgrund ihrer Datenstruktur unterschiedlich sind, ist die Wahl des Mixed-Methods-Ansatzes angebracht.

Abbildung 16 zeigt die Übersicht der Methoden sowie Datentypen in den jeweiligen Phasen. Die Erhebungsmethode ist eine Beobachtungsstudie mit Befragung und Erfassung von Verhaltensdaten (Logdaten). Zur Datenaufbereitung werden die dynamischen Daten der Webseitenmerkmale und Verhaltensdaten in ein jeweiliges numerisches Relativ überführt, sodass sie messbar gemacht werden können. In die Auswertung mithilfe der logistischen bzw. linearen Regression fließen die statischen Daten der Personenmerkmale sowie alle zwischenzeitlich aufbereiteten Daten (Webseiten- und Verhaltensdaten) ein.

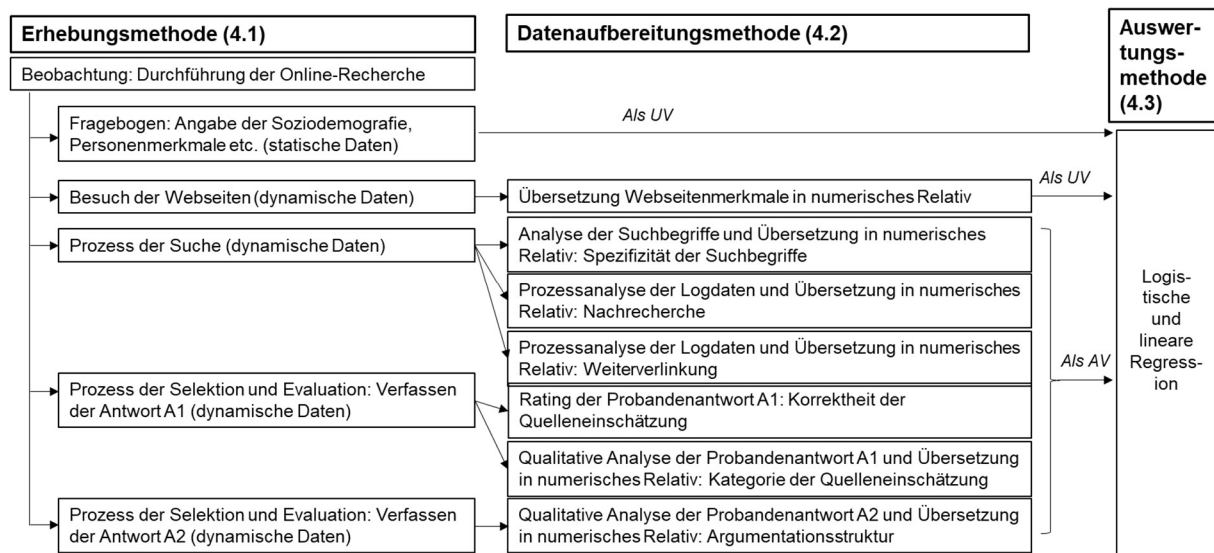


Abbildung 16: *Mixed-Methods-Ansatz im Forschungsvorhaben.*

4.1. Projektbeschreibung BRIDGE

2020 startete das Projekt BRIDGE. Die Studien aus dem Projekt wurden kooperativ von der Johannes Gutenberg Universität-Mainz und der Goethe Universität Frankfurt am Main im Zeitraum 07/2020 - 06/2023 durchgeführt.

Das allgemeine Studiendesign setzte sich aus einer Online-Studie im Bereich GEN-COR, einer aus DOM-COR, einem Interventionstraining sowie einer Post-Erhebung zusammen. Für vorliegende Forschungsarbeit wird die Online-Studie im Bereich des generischen COR (GEN-COR) zu Grunde gelegt, welche Ende 2020 durchgeführt wurde. Die Interventionstrainings wurden in verschiedenen Runden und für die entsprechenden Domänen in unterschiedlichen Zeitabständen in 2021 durchgeführt. Die Post-Erhebung fand Ende des Jahres 2021 statt. Die letzten Datenaufbereitungen und finalen Schritte (z. B. Rückversand der Laptops) wurden Anfang des Jahres 2022 getätigt. Verwendbare Daten belaufen sich unter anderem auf solche aus dem Bereich der Soziodemografie, Kognitionen, Informationsverarbeitung sowie Medienutzung (Inputfaktoren), welche durch die Fragebögen mithilfe der Software Unipark (*Online Umfrage erstellen*, o. J.) in der ersten Online-Studie erhoben wurden. Ein zweites Set der auswertbaren Daten bilden die Prozessdaten, welche beispielsweise die Anzahl der Suchanfragen, die verwendete Suchmaschine oder -begriffe abdeckt. Das dritte Datenset, wie Bearbeitungsdauer, Performanz, qualitative Bewertung der Argumente oder der verwendeten Webseiten stellen Output- oder Performanzfaktoren dar. Die folgenden Unterkapitel geben einen Einblick in den Erhebungsablauf, das Assessment sowie in die Zusammensetzung der Stichprobe.

4.1.1. Erhebungsablauf und -design

Die Studie umfasste drei Erhebungswellen, beginnend mit der ersten Welle im Dezember 2020. Die zweite Welle und somit Erhebung der zweiten Kohorte erfolgte ca. ein halbes Jahr später im Juni 2021. Den Abschluss fand die Erhebung mit der letzten Kohorte von Mitte bis Ende Dezember 2021.

Ablauf der Erhebung. Sobald sie sich verbindlich für die Studie angemeldet hatten, erhielten die Probanden entweder leihweise einen Laptop zugesandt oder konnten mit ihrem privaten Endgerät (Laptop, Rechner) daran teilnehmen. Ehe die Probanden mit der Teilnahme beginnen konnten, wurden sie darüber informiert, dass ihre Daten nur zu Studienzwecken verwendet und nicht weitergegeben werden. Alle, die angaben, dass sie das verstanden und akzeptiert haben, konnten an der Studie teilnehmen. Für die Erhebung arbeiteten die Probanden auf einer virtuellen Maschine in Azure Labs, für das sie zuvor eine Anleitung mit Instruktionen per E-Mail erhalten hatten. Auf der virtuellen Maschine wurde ein eigens entwickeltes Software-Plugin genutzt, um die Browser-Aktivitäten während der Erhebung zu tracken. So war es möglich, Seitenaufrufe, die Verweildauer auf den besuchten Webseiten und Mausbewegungen (Scrollen) während des Assessments zu dokumentieren. Das Plugin öffnete sich nach dem Start der virtuellen Maschine automatisch, ebenso wie die Erhebung im Webbrowser, welche mithilfe von Unipark erstellt und durchgeführt wurde. Um die Daten zu anonymisieren und somit keine Rückschlüsse zwischen den Antworten und Personen ziehen zu können (Ebel & Meyermann, 2015), bekam jeder Proband einen individuell verschlüsselten (hash-)Code bestehend aus Geburtstag, Geburtsname der Mutter etc. Sobald dieser auf der ersten Seite generiert wurde, konnten die Probanden mit der Umfrage und der anschließenden Aufgabenstellung beginnen.

Aufgabenstellung. Die Aufgabenstellung deckte GEN-COR ab, welches keinerlei Vorwissen erforderte (Zlatkin-Troitschanskaia, Brückner, et al., 2021). Dieses Problem stellt ein unklar definiertes, offenes Problem dar, denn es gibt keine weiteren Spezifikationen zur Lösung des Problems, wie beispielsweise eine klare Strategie oder Lösung (Eysenck, 2001; Jonassen, 1997). Diese Aufgabenart wurde unter anderem deshalb verwendet, da sie repräsentativ für die Informationsprobleme im digitalen Zeitalter ist. Denn viele Informationsprobleme, welche mithilfe des Internets gelöst werden, sind keine einfach zu beantwortenden Ja/Nein- oder Fakten-Fragen, sondern Abwägungsproblematiken. Alltägliche Informationsprobleme beinhalten beispielsweise gesundheitliche Fragestellungen oder Einkaufsentscheidungen. Hinzu kommt, dass im Zuge dieser Forschung Kognitionen und Kognitionsprozesse untersucht werden sollen. Ein derartiges, komplexes und schlecht-definiertes Informationsproblem erfordert ein Abwägen der Entscheidungen und nächsten Schritte. Hierdurch werden die Kognitionsprozesse

ausgelöst bzw. sichtbar gemacht und können miteinander verglichen und untersucht werden. Dieser Prozess und seine einzelnen Schritte besser zu verstehen, ist u. a. für das Ziel dieser Forschungsarbeit dienlich.

Die Probanden sollten sich vorstellen, ein E-Bike anzuschaffen und hierzu recherchieren, ob E-Bikes gesundheitsfördernd sind:

„Sie überlegen, sich ein E-Bike zur Förderung Ihrer Gesundheit anzuschaffen. Dafür recherchieren Sie online nach Informationen zur Förderung von E-Bikes zur Gesundheit. Fragestellung: Recherchieren Sie für die Beantwortung der Fragen im Internet. Prüfen Sie bitte dann die Zuverlässigkeit²⁵ der Informationen Ihrer Online-Recherche. Geben Sie bitte immer die verwendeten Internetquellen (URLs) an.

1. Fügen Sie bitte die von Ihnen verwendeten Quellen mit jeweiliger URL ein und geben Sie dahinter an, ob Sie die jeweilige Quelle für zuverlässig erachten und begründen Sie bitte kurz, warum (Sie haben hierfür 10 Minuten Zeit).

2. Verfassen Sie eine kurze Stellungnahme, in der Sie auf Basis Ihrer Recherche aus Aufgabenteil 1 ein begründetes Urteil darüber abgeben, ob E-Bikes zur Gesundheitsförderung beitragen. Verweisen Sie auch hier auf entsprechende Informationen Ihrer Recherche und geben Sie bitte die Quellen (URLs) an (Sie haben hierfür 10 Minuten Zeit).“

In Teilaufgabe 1 sollten die Probanden hierbei die verwendeten URLs angeben und aufzeigen, ob sie sie für zuverlässig (somit i. w. S. glaubwürdig) einschätzten oder nicht und eine Begründung für ihr Urteil abgeben. In Teilaufgabe 2 sollten die Probanden die ursprüngliche Fragestellung beantworten, in dem sie eine begründete Meinung oder Stellungnahme unter Zuhilfenahme und Angabe der in A1 angegebenen URLs schreiben. Für beide Aufgaben zusammen hatten die Probanden 20 Minuten Zeit. In Unipark wurde voreingestellt, dass die Seite nach exakt 20 Minuten im Fragebogen weiterspringt, sofern in beiden Freitextfeldern (A1 und A2) etwas angegeben war. Lediglich elf Probanden, welche nach 20 Minuten noch nichts in A2 eingetragen hatten, überschritten die Zeitgrenze von 20 Minuten. Die Durchschnittsdauer aller Probanden zur Bearbeitung der Aufgabenstellung betrug 17.85 Minuten.

Die Datenaufzeichnung ergab neben der Umfragedaten auch Logdaten im CSV-Format, welche die Events im Browser (Suchmaschine, Eingabe Suchbegriff, Scrollen, Klicken) und auf

²⁵ In der Aufgabenstellung wurde explizit nach Zuverlässigkeit gefragt, da dies ein Bestandteil („Trustworthiness“) der wahrgenommenen Glaubwürdigkeit darstellt. Es wurde sich dagegen entschieden, nach Glaubwürdigkeit zu fragen, da dies andernfalls zu einer zu subjektiven Einschätzung i. S. v. „Empfindung“ geführt hätte. Mit „Zuverlässigkeit“ sollten die Probanden zur Angabe einer vergleichbaren, wissenschaftlichen Begründung angeregt werden.

der Webseite (Scrollen, Klicken) erfassten. Mithilfe dieser Logdaten konnten somit die hier vorliegenden Analysen berechnet werden.

4.1.2. Beschreibung und Zusammensetzung der Stichprobe

Die Stichprobe setzte sich aus Berufsanfängern aus drei verschiedenen Domänen zusammen, wobei sich darunter Mediziner im praktischen Jahr sowie Referendare von Lehramt (Lehramtsanwärter) und Rechtswissenschaften (Rechtsreferendare) befanden. Die Auswahl für Berufsanfänger aus den drei Domänen, welche in ihrem ersten Jahr nach Abschluss ihres Masters waren, wurde bereits in der Einleitung kurz begründet. Die Wahl auf Berufsanfänger begründete sich mit der bereits abgeschlossenen Ausbildung eines (erfolgreichen) kritischen Online-Denkens und der Abfrage des Status-Quo über den Studienabschluss hinaus. Da der Fokus der Arbeit jedoch nicht explizit auf den unterschiedlichen Domänen liegt, wird die Verteilung der Studierenden in Bezug auf die Domänen hier nur kurz und zwecks Vollständigkeit thematisiert:

Die Stichprobe setzte sich aus insgesamt 144 Probanden zusammen, wovon 95 (65.5 %) weiblich, 40 (27.6 %) männlich und 9 (6.2 %) divers waren. Die Zugehörigkeiten zu den Domänen verteilte sich wie folgt: Der Großteil der Probanden ($n = 68$, 46.9 %) studierte Medizin, gefolgt von 46 (31.7 %) Probanden aus der Domäne Lehramt und abschließend 30 (20.7 %) Probanden aus Jura. Das Durchschnittsalter der Probanden lag bei 27.15 ($SD = 3.5$) und die Schulabschlussnote (Abiturnote) belief sich im Durchschnitt auf 1.76 ($SD = 0.64$).

4.2. Datenaufbereitung und -optimierung

Vorbereitende Schritte. Sämtliche Rohdaten aller Probanden wurden zunächst zentral auf dem Server zusammengetragen und gesammelt. Es handelte sich hierbei einmal um die von der PC-Software aufgenommenen Logdaten, welche als CSV-Format vorlagen. Im Weiteren wurden die Antworten der Probanden (Teilaufgabe 1 und 2) sowie die Umfragedaten mithilfe von Unipark erfasst. Letztere lagen ebenfalls im CSV-Format vor.

Für die Aufbereitung der Umfragedaten wurden die einzelnen Items zu einem Mittelwertindex verrechnet (im Folgenden abgekürzt mit M) ermittelt. Die Logdaten wurden als Excel-Tabelle formatiert und enthielten die individuelle Identifizierungsnummer des Probanden (= hash ID), einen Zeitstempel, die besuchte URL, die getätigte Aktion mit der URL (z. B. scrolled down/up, copy, paste) sowie eine Beschreibung der Worte, die in eine Suchmaschine eingetragen wurden. Die Daten wurden folgendermaßen aufbereitet: Leere Zeilen oder Spalten wurden entfernt, ebenso wie fehlerhafte Spalten oder Zeilen, welche beim Übertragen keine sinnstiftenden inhaltsbezogenen Daten enthielten. Die Beschreibung der gewählten Worte, die in die Suchmaschine eingetragen wurde, war nicht bei allen drei Kohorten enthalten. Aus diesem

Grund wurde dem Link (URL-Spalte) gefolgt, um die entsprechenden Suchbegriffe in die Spalte manuell nachzutragen.

Es wurde zudem eine weitere Spalte eingefügt, welche die URL, die teilweise sehr lang waren, in eine kürze und prägnantere Variante umwandelte. Zu diesem Zwecke wurden sämtliche URLs aufgerufen und gleiche URLs vereinheitlicht. Da Probanden teilweise über eine Suchmaschine, wie Google (*Google*, o. J.) oder Ecosia (*Ecosia – Die Suchmaschine, die Bäume pflanzt*, o. J.), die URL aufriefen, war der dazugehörige Link mit den entsprechenden Verlinkungen zu Google oder Ecosia verbunden (z. B. <https://search.google/quarks.co/xxx>). Diese waren allerdings inhaltsgleich mit der URL, wenn diese direkt aufgerufen wurde (wie z. B. www.quarks.co/xxx). Im Zuge dieses Angleichungsschrittes wurde somit eine neue Variable als eine einheitliche Bezeichnung für beide Wege geschaffen, beispielsweise *Quarks Artikel 1* (Schwandner, 2019). Wurden unterschiedliche Artikel der gleichen Webseite aufgerufen, wurden diese aufsteigend durchnummeriert (Beispiel *Quarks Artikel 1*, *Quarks Artikel 2* etc.). Es ergaben sich aus den unterschiedlichen vereinfachten Bezeichnungen diverse Kategorien, wie beispielsweise „Suchmaschine“, „Übersetzer“, „Startseiten“ oder „Über Uns“/„Impressum“ aber auch „Irrelevantes“ (wenn beispielsweise ein Proband nach der Umrechnung der Jura-Note gesucht hatte). Die Kategorien wurden für die weitere Aufbereitung der Logdaten verwendet.

Mithilfe der Logdaten konnten z. B. Suchaktivitäten (Eingabe von Suchbegriffen, Aufrufen von Webseiten) oder die Reihenfolge der Webseitenaufrufe rekonstruiert werden. Dies wurde in eine gesonderte Excel-Tabelle pro Probanden übertragen. Mithilfe der Antworten aus Unipark konnte zudem ermittelt werden, welche Webseiten für die Beantwortung der Aufgabe ausgewählt wurden. Dies wurde in einer weiteren Excel-Tabelle für jeden Probanden extrahiert.

Datenaufbereitung. Insgesamt 236 unique besuchte respektive aufgerufene Webseiten wurden in den Logdaten verzeichnet. Auch hier erfolgte eine Bereinigung im Sinne einer Löschung. So wurden Webseiten aus der inhaltlichen Analyse bezüglich der Webseitenmerkmale entfernt, wenn sie als eine Startseite, „Aktuelles“-Seite oder eine „Über Uns“/„Impressum-Seite“²⁶ erkennbar waren. Weiterhin entfernt wurden Webseitenaufrufe, die zur Überprüfung einer zuerst besuchten Webseite dienten, wie Wikipedia (*Wikipedia, die freie Enzyklopädie*, o. J.), Academic Accelerator (Accelerator, o. J.) oder Mediabiasfactcheck (*Media Bias/Fact Check News*, 2023). Nicht inhaltsrelevante Webseiten wurden ebenfalls ausgeschlossen, wie beispielsweise das Suchergebnis nach der Umrechnung der Juranote oder nach der in der Umfrage gestellten Denkaufgabe. Ebenso ausgeschlossen wurden Universitäts- oder Hochschuleseiten, die einen Login erforderten. Diese wurden von den Studierenden durch Eingabe

²⁶ „Über Uns“/„Impressums-Seiten wurden lediglich für die inhaltliche Analyse der Webseitenmerkmale entfernt. Für die Bildung der abhängigen Variable Nachrecherche wurden sie wieder herangezogen.

ihrer persönlichen Login-Daten geöffnet. Da hierzu aufgrund von datenschutzrechtlichen Regelungen keinerlei Informationen vorliegen konnten, konnten solche Ergebnisseiten auch nicht aufgerufen oder gar beurteilt werden. Alle Übersetzer-Online-Tools, wie DeepL (*DeepL Übersetzer*, o. J.) oder dict.cc (*dict.cc | Wörterbuch Englisch-Deutsch*, o. J.) wurden ebenfalls nicht berücksichtigt.

Nach erfolgter Bereinigung der Logdaten konnten insgesamt 147 unique, inhaltsrelevante und bewertbare Webseiten analysiert werden. Zwei Webseiten kamen im Zuge der Aufarbeitung hinzu, da diese zwar von den Probanden ausgewählt, aber aufgrund eines möglicherweise technischen Fehlers nicht in den Logdaten erschienen sind. Es ist daher davon auszugehen, dass diese beiden Webseiten eigentlich ebenso theoretisch in den Logdaten hätten enthalten sein müssen. Sie werden aus Gründen der Korrektheit allerdings nur zu den selektierten Webseiten dazugerechnet. Somit sind 149 Webseiten inhaltsrelevante und einzelne Webseiten mindestens aufgerufen worden, wovon 147 definitiv aufgerufen und 110 selektiert wurden. Nennenswert ist an dieser Stelle die Häufigkeit der unique aufgerufenen Webseiten. Denn es ist zu erkennen, dass gleiche Seiten sehr häufig aufgerufen wurden. So wurde beispielsweise ein Artikel der Webseite *Quarks* von 74 Probanden aufgerufen (Tabelle 4).

Tabelle 4: Meist aufgerufene unique Webseiten.

Aufgerufene Webseiten	URL	Probandenanzahl
Quarks Artikel 1 (Schwandner, 2019)	https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/	74
Zeitschrift Sportmedizin Artikel (Hutterer, 2019)	https://www.zeitschrift-sportmedizin.de/e-bikes-und-der-einfluss-auf-das-gesundheitsverhalten-von-risikogruppen	45
Emotion Technologies Artikel (Fördert e-Bike fahren die Gesundheit?, 2020)	https://emotion-technologies.de/e-bike-infos/vorteile/gesundheit/	44
Ärzteblatt Artikel 1 (E-Bikes, 2018)	https://www.aerzteblatt.de/nachrichten/96943/E-Bikes-gesund-und-gefaehrlich-zugleich	33
Envivas Artikel (Farin, 2023)	https://www.envivas.de/magazin/im-alltag/gesund-dank-e-bikes/	28
Zeit Artikel (Reidl, 2018)	https://www.zeit.de/mobilitaet/2018-08/pedelec-gesundheit-alter-bewegung-radfahren	19
Focus Arztsuche Artikel (Meyer, 2019)	https://focus-gesundheit.de/magazin/gesundheitsstipps/e-bikes-so-profi-tiert-ihre-gesundheit	13
Ebike Gesundheit Artikel (Wie gesund ist E-Bike fahren?, 2022)	https://www.ebike-gesundheit.de/	12
Elektrofahrrad24 Artikel (So gut ist E-Bike fahren für Deine Gesundheit!, 2019)	https://www.elektrofahrrad24.de/blog/e-biken-fuer-die-gesundheit	12
Stromerbike Artikel (E-Bike & Fitness, 2020)	https://www.stromerbike.com/de/e-bike-fitness-gesundheit	12
IJBPA Artikel 1 (Bourne et al., 2018)	https://ijbpa.biomedcentral.com/articles/10.1186/s12966-018-0751-8	11
Uni Basel Artikel (E-Bike fahren fördert Fitness und Gesundheit – schon nach vier Wochen, 2018)	https://www.unibas.ch/de/Aktuell/News/Uni-Research/E-Bike-fahren-foerdert-Fitness-und-Gesundheit-schon-nach-vier-Wochen.html	11

4.2.1. Vorgehen zur Operationalisierung der Variablen

Die Online-Studie hatte zum Ziel, das Informationssuch- und Selektionsverhalten der Probanden in einer realitätsnahen Umgebung zu untersuchen, was durch eine Beobachtungsstudie erzielt werden kann. Die unabhängigen Variablen sind zum einen die Personenmerkmale, welche in einem Eingangsfragebogen erhoben wurden. Im nachfolgenden Kapitel 4.2.2 werden die Personenmerkmale näher beschrieben.

Die Merkmale der angeschauten und selektierten Webseiten, welche anhand der Logdaten identifiziert wurden, bilden den zweiten Part der unabhängigen Variablen. Diese wurden mithilfe der aus der Literatur herausgearbeiteten Kriterien (s. Kapitel 2.6.1) bewertet und messbar gemacht. Es wurden hierzu nacheinander alle angeschauten und selektierten Webseiten geöffnet und anhand der entsprechenden Kriterien für das bestimmte Merkmal analysiert. Auf

Basis dieser Analyse wurden die jeweiligen Webseitenmerkmale in ein numerisches Relativ umgewandelt, was ausführlich für jedes Merkmal sowie für jede Variable in den Kapiteln 4.2.4 und 4.2.5 erläutert wird.

Die abhängigen Variablen wurden in Anlehnung an das IPS-I-Modell und des COR-Konstruktes aus dem Verhalten der Probanden bei der Suche, Selektion und dem Verfassen der Antworten abgeleitet und messbar gemacht. Im Detail bedeutet dies, dass aus dem Suchverhalten die Prozessfaktoren resultierten. Die Suchbegriffe wurden hierbei analysiert und bilden die erste AV. Des Weiteren erfolgte eine Prozessanalyse anhand der Logdaten, welche die beiden weiteren AVs aus dem Suchprozess darstellen. Das Selektions- und Evaluationsverhalten zeigte wiederum die Performanzfaktoren auf, welche sich aus den Probandenantworten ergaben. D. h. die Einschätzung der Zuverlässigkeit der Quelle (Begründung und Korrektheit) sowie die begründete Stellungnahme für oder gegen die Gesundheitsförderung von E-Bikes lieferten die Basis für die drei abhängigen Variablen der Evaluationsstrategie. Eine übersichtliche Darstellung der unabhängigen sowie abhängigen Variablen ist in folgender Tabelle 5 dargestellt:

Tabelle 5: Übersicht der verwendeten Input-, Prozess- und Performanzfaktoren.

Input-faktoren	Quelle	Prozessfaktoren	Quelle	Performanz- bzw. Outputfaktoren	Quelle
NfC	Ein-gangsfragebogen	Webseiteneigenschaften: kommerzielles Interesse, Quellen, Autor, Aktualität, Genre, Narrative, Argumentqualität, Bilder vorhanden, linguistische Merkmale etc.	Logdaten und Analyse der Webseiten	Analyse/Klassifikation der Begründungen der (Un-)Zuverlässigkeit → Systematische Informationsverarbeitung	Probandenantwort A1
IO		Suchbegriffe → Spezifität	Logdaten	Korrekte Quelleinschätzung	
NfCC		Vorgehen Suche: Tiefe und Breite des Prozesses → Weiterverlinkung und Nachrecherche		Struktur verwendeter Argumente	Probandenantwort A2 (Stellungnahme)

4.2.2. Operationalisierung der Personenmerkmale als unabhängige Variable

Die Personenmerkmale von Interesse wurden mithilfe des Umfrageprogramms Unipark und einem darin erstellten Fragebogen via Selbstauskunft erhoben. Nachdem die Probanden 20 Minuten Zeit hatten, die eigentliche Aufgabenstellung zu bearbeiten, wurden sie zu dem Fragebogen weitergeleitet. Dieser enthielt neben den Personenmerkmalen auch klassische soziodemografische Abfragen. Die Probanden wurden beispielsweise nach dem Alter, und zu welchem Geschlecht sie sich zuordnen würden, gefragt. Weiterhin wurde nach der Herkunft und dem Bildungsabschluss der und Kommunikationssprache mit den Eltern gefragt. Der bisher

berufliche Werdegang wurde durch die Nachfrage nach der Hochschulzugangsberechtigungsnote, einer möglichen Ausbildung, einem Praktikum und oder einem (medizinischen) Nebenjob abgedeckt. Zuletzt wurde erhoben, was die Probanden studierten.

NfC wurde in Anlehnung an die Originalskala von Petty und Cacioppo (1982) als kurze deutschsprachige Version mithilfe von vier Items erhoben (Beißert et al., 2014)²⁷. Diese Skala erfasste die Items „Es genügt mir einfach die Antwort zu kennen, ohne die Gründe für die Antwort des Problems zu verstehen (i)“, „Ich habe es gern, wenn mein Leben voller kniffliger Aufgaben ist, die ich lösen muss“, „Ich würde kompliziertere Probleme einfacheren Problemen vorziehen“ sowie „In erster Linie denke ich, weil ich muss (i)“. Die Probanden hatten dabei eine 6-Likert-Skala zur Auswahl, aus der sie zwischen „1 = *trifft gar nicht zu*“ und „6 = *trifft voll zu*“ auswählen konnten.

NfCC wurde in der Post-Erhebung in Analogie zur Ursprungsskala mithilfe der deutschen Skala bestehend aus 16 Items von Schlink und Walther (2007)²⁸ abgefragt. Beispielhafte Items waren: „Ich mag Aufgaben, bei denen noch unklar ist, wie der genaue Lösungsweg aussieht (i)“, „Ich bevorzuge die Gesellschaft guter Freunde, weil ich weiß, was ich von ihnen zu erwarten habe“, „Ich mag keine unvorhersehbaren Situationen“ oder „Ich finde es spannend nicht zu wissen, was das Leben einem bringt (i)“. Die Probanden hatten dabei eine 6-Likert-Skala zur Auswahl, aus der sie zwischen „1 = *stimme überhaupt nicht zu*“ und „4 = *stimme voll und ganz zu*“ auswählen konnten.

Die drei Items der englischen Original-Skala IO von Karr-Wisniewski & Lu (2010)²⁹ wurden für die Studie ins Deutsche übersetzt und lauten: „Ich bin oftmals abgelenkt von der übermäßigen Fülle an Informationen in den Medien“, „Ich fühle mich überfordert von der Menge an Informationen aus den Medien, mit denen ich täglich konfrontiert werde“ sowie „Meistens ist mein Problem, dass ich zu viele Informationen verarbeiten muss und nicht, dass ich zu wenig Informationen habe, um mir eine Meinung zu bilden“. Die Probanden konnten dabei aus einer 6-Likert-Skala mit „1 = *trifft gar nicht zu*“ und „6 = *trifft voll zu*“ auswählen.

4.2.3. Problematische Webseitenmerkmale

Im Theorieteil wurden insgesamt 21 Merkmale von Webseiten thematisiert. 12 konnten im Untersuchungsmaterial nicht angewendet werden und spielen daher im Folgenden auch keine Rolle mehr. Die Gründe für den Ausschluss beliefen sich dabei auf konzeptuell-kontextuelle

²⁷ Die Originalskala weist ein Cronbach's α von .51 und .54 auf (Beißert et al., 2014, S. 11). Jedoch verweisen die Autoren darauf, dass bei einer Kurzskala aufgrund der geringen Testlänge an sich ein niedrigeres Cronbach's α zu erwarten ist (Cortina, 1993, S. 103). Sie verwendeten zum Beweis einer ausreichenden Reliabilität ihrer Skala die Retest-Reliabilität, welche Werte von $r_{tt} = .58$ und $.78$ aufweist (Beißert et al., 2014, S. 11).

²⁸ Die Items weisen im Original eine interne Konsistenz von Cronbach's $\alpha = .78$ auf.

²⁹ Die interne Konsistenz der Originalskala mit den drei Items liegt bei $\alpha = 0.72$.

oder analysebedingte Probleme. Ersteres meint, dass ausgehend vom E-Bike-Kontext versucht wurde, zunächst die Merkmale für die Analyse bzw. Operationalisierung vorzubereiten. Dabei wurde jedoch festgestellt, dass einige Merkmale nicht für den Kontext anwendbar waren bzw. sich konzeptionell mit weiteren Webseitenmerkmalen zu stark überschneiden würden. Analysebedingte Probleme traten bei dem Versuch auf, die Merkmale auf den von den Probanden aufgerufenen Webseiten zu analysieren. Nachstehend werden die entsprechenden Merkmale mit den Begründungen für den Ausschluss dargestellt.

Kontextuelle Probleme ergaben sich für folgende Merkmale:

Spezialisierungsgrad. Die Beurteilung der Glaubwürdigkeitswahrnehmung in diesem Kontext erschloss sich aufgrund der gering vorhandenen Literatur nicht eindeutig: Webseiten, die sich auf E-Bikes und Gesundheit spezialisiert haben, könnten eine größere Expertise suggerieren, und deshalb als glaubwürdiger wahrgenommen werden. Oder aber diesen Webseiten könnte eine mangelnde Objektivität bzw. ein Vorhandensein von kommerziellem Interesse vorgeworfen werden, was wiederum einer positiven Glaubwürdigkeitsbeurteilung schadet. Umgekehrt könnten allgemeine, nicht-spezifische Webseiten als glaubwürdiger wahrgenommen werden, weil sie Objektivität vermitteln oder aber als unglaubwürdiger, weil sie aufgrund ihrer Nicht-Spezialisierung eine geringe Expertise aufweisen. Aufgrund der Uneindeutigkeit dieser Wirkungsrichtung und zudem der Überschneidung mit kommerziellem Interesse wurde das Merkmal des Spezialisierungsgrades zwar gewürdigt, aber nicht für vorliegenden Forschungskontext angewandt.

Das Webseitenmerkmal der *Statistiken* wurde für vorliegende Forschung in das Merkmal der Angabe von Quellen integriert und als separates somit ausgeschlossen. Grund hierfür ist, dass eine Unterscheidung hinsichtlich numerischer oder textlicher Angaben zu einer zu detaillierten Differenzierung führen und zudem dieses Merkmal gleichzeitig mit der Quellenangabe korrelieren würde.

Design wurde aus der Forschung ausgeschlossen, weil die Unterscheidungen innerhalb des Merkmals literaturbedingt an sich schon wenig trennscharf und zu subjektiv waren und diese bei einem Versuch der Operationalisierung für den E-Bike-Kontext noch unschärfer wurden. Es ist beispielsweise nicht näher präzisiert, welche Art von Layout oder Farbgestaltung mehr oder weniger zur Glaubwürdigkeitswahrnehmung beitragen. Da die Beurteilung des Designs bzw. der Ästhetik stark mit dem beurteilenden Individuum zusammenhängt konnte nicht von einer allgemeingültigen Glaubwürdigkeitsbeurteilung in Bezug auf E-Bike-Webseiten ausgegangen werden.

Vollständigkeit konnte nicht in vorliegende Forschung integriert werden. In der Realität und insbesondere aufgrund der Informationsmenge gibt es nie die absolute Sicherheit, dass die

Information vollständig ist. Insbesondere bei einer Aufgabe, die nicht konkret nach einem Fakt fragt (z. B. Name oder Datum einer historischen Persönlichkeit), wie in vorliegendem Beispiel, kann zu keinem Zeitpunkt gemessen werden, ob alle auf einer Webseite befindlichen Informationen als vollständig angesehen werden können. Insofern findet dieses Merkmal zwar hier eine Erwähnung, wurde aber aufgrund der soeben beschriebenen kontextuellen Einschränkungen nicht berücksichtigt.

Reputation. Ähnlich wie bei dem Merkmal Bekanntheit ist die Reputation eine subjektive Wahrnehmung eines Individuums. Ob eine Quelle oder ein Medium bei einem Probanden mit einer gewissen Reputation und somit einem Ruf wahrgenommen wurde, kann in einem realen Szenario ohne vorherige Kontrolle der aufzurufenden Webseiten nicht erfasst werden. Anders als beim Merkmal Bekanntheit steht hierbei keine Applikation oder Webseite zur Verfügung (Ausführungen zur Operationalisierung der Bekanntheit folgen im Anschluss), die eine Annäherung eines Durchschnittswertes für die empfundene Reputation der besuchten Webseiten wagen kann. Zweitens setzt das Wissen um die Reputation einer Webseite oder eines Autors voraus, dass man zumindest schon einmal etwas von jemanden/etwas gehört hat und derjenige/diese vielleicht sogar bekannt ist. Somit würde das Kennen einer Webseite und somit die Bekanntheit eine notwendige, vorgelagerte Bedingung darstellen, dass die Reputation überhaupt erst beurteilt werden könnte. Dies führte zu einer konzeptionellen Überschneidung. In Verbindung mit der erst genannten Problematik wurde daher entschieden, die Reputation nicht zu untersuchen.

Analysebedingte Probleme ergaben sich für folgende Merkmale:

Foto des Autors: Alle $n = 97$ Webseiten, die einen Autor angegeben hatten, wurden nochmals auf Fotos überprüft. Es wurde festgestellt, dass lediglich vier Webseiten ein Foto des Autors enthielten. Aufgrund des geringen Vorkommens und der daraus folgenden geringen Varianz musste dieses Webseitenmerkmal somit aus der Analyse ausgeschlossen werden.

Im Kapitel 2.6.1.6 noch zusammen thematisiert, wurde das Merkmal der Struktur für die Operationalisierung aufgeteilt, da sich die Struktur entweder auf die Webseite oder auf die Informationen beziehen kann und sich daraus unterschiedliche Bezugsobjekte ergeben. Mit der Aufteilung wird somit präziser und korrekter gehandelt. Da die Literatur keinerlei Hinweise darauf gibt, wie welche Aspekte einer Struktur die Glaubwürdigkeit erhöhen und welche sie vermindern, wurde mit weiteren Beteiligten im Projekt BRIDGE ein Workshop zur Erarbeitung einer möglichen Operationalisierung für beide Merkmale (Information und Webseite) durchgeführt. Die Ergebnisse und hieraus sich ergebende Anwendungsmöglichkeiten werden für die *Struktur der Information* an dieser Stelle thematisiert. Der Part *Struktur der Webseite* folgt anschließend.

Das Merkmal *Struktur der Information* bildet ab, wie gut strukturiert ein Text ist. Zur Operationalisierung dieses Merkmals kann sich der Linguistik bedient werden, wobei hierin die inhaltliche Strukturierung über die Abfolge von Ideen bzw. Gedankengänge im Text (d. h. Textkohärenz) oder die grammatikalische und syntaktische³⁰ Struktur an der Textoberfläche (d. h. Kohäsion) fallen. Die Kohäsion kann rein formell über gewisse Marker identifiziert werden. Kohäsionsmarker sind beispielsweise logische Satzverbindungen für Abfolge (erstens, zweitens, drittens etc.), Begründung (da, weil) oder Einschränkung (obwohl; trotz dessen, dass). Diese wecken beim Lesen die Erwartung an einen Zusammenhang der nachfolgenden Anteile, und suggerieren dem Lesenden, dass die ausgedrückten Aussagen tatsächlich entsprechend verbunden sind. Für die inhaltliche Struktur müssen wiederum die inhaltlichen Aussagen des Textes analysiert werden, denn „Textkohärenz ist eine Frage von Interpretationen“ (Busse, 2015, S. 182) und somit immer vom Kontext und vom Verständnis der Leser abhängig (Busse, 1992). Sowohl die Analyse der oberflächlichen Struktur (Kohäsion) über Marker als auch die inhaltliche Analyse der Ideen im Text (Kohärenz) erfordert eine ausführliche linguistische Analyse, die mithilfe von Computerlinguistik durchgeführt werden könnte. Das Feld der linguistischen Marker ist allerdings sehr umfangreich und eine laienhafte Analyse des Merkmals Informationsstruktur würde diesem nicht gerecht werden. Außerdem würde eine professionellere Analyse mithilfe von Computerlinguistik den Fokus der Arbeit verlagern. Aus diesem Grund wurde das Merkmal für die vorliegende Forschung ausgeschlossen.

Struktur der Webseite. Die Kriterien als Resultat des Workshops wurden nach mehreren Diskussionsrunden übereinstimmend festgelegt und beliefen sich auf folgende: Eine gut strukturierte Webseite respektive Nachricht arbeitet mit Absätzen, Überschriften und visuellen Highlights, wie beispielsweise Fettmarkierungen, Unterstreichungen oder Kursiv-Text. Hierbei ist es notwendig, nicht die textlichen Aspekte zu betrachten, sondern die Webseite als das „große Ganze“ in einer größeren Perspektive zu beleuchten. Dabei zeigte sich, dass alle Webseiten gut strukturiert waren und zu wenige Unterschiede bestanden, die Unterschiede in Glaubwürdigkeit und folgenden Prozessen verursachen können. Eine qualitative oder gewichtete Unterscheidung zwischen den einzelnen Kriterien war nicht möglich, was zur Folge hatte, dass die Webseiten in diesem Merkmal nicht variierten. Aus diesem Grund wurde das Merkmal der Struktur der Nachricht aus der weiteren Analyse ausgeschlossen.

Faktische Korrektheit. Vorüberlegungen zur Operationalisierung der faktischen Korrektheit ergaben, dass dieses Merkmal für diesen Forschungskontext nicht problemlos anwendbar war. Ob E-Bikes gesundheitsfördernd sind, ist nicht per se mit einer eindeutigen Antwort zu klären und stellt somit keinen einfach zu verifizierenden Fakt dar, wie beispielsweise Name

³⁰ Syntax ist ein Teilgebiet der Linguistik und beschreibt die auf Regeln basierenden Strukturen von Sätzen und ihren Elementen (Hentschel, 2023, S. 147).

oder das Geburtsdatum einer bekannten historischen Persönlichkeit. Die „richtigere“ Antwort für die Frage nach der gesundheitsfördernden Wirkung von E-Bikes ist daher differenziert zu betrachten. Es kommt auf Umstände, wie Alter, bisherige Bewegung, allgemeine Gesundheit etc. an. Insofern waren bei dem Merkmal die faktische Korrektheit der zusätzlichen Informationen und Inhalte auf den jeweiligen Webseiten zu überprüfen. Nach der ersten Analyserunde wurde allerdings festgestellt, dass die angegebenen Informationen zwar alle faktisch korrekt sind, sich aber immer wieder auf die gleichen Ursprungsstudien bezogen. Es lag somit eine geringe Variabilität zwischen den Webseiten respektive den Probanden vor, weil die Querverweise sich inhaltlich deckten, und somit auch die faktische Korrektheit. Nach einer erneuten Validierung der Vorgehensweise wurde eine neue und vereinfachte Operationalisierung der faktischen Korrektheit angewandt: Aus den Vorüberlegungen wurde angenommen, dass immer dann von einer faktischen Korrektheit auszugehen ist, wenn einmal der Urheber selbst die Daten erhoben hat (Primärquelle) und weiterhin im Genre mit der höchsten Kategorie und somit „wissenschaftliche Fachzeitschrift“ zu finden ist. Grund hierfür ist, dass, wenn Daten selbst experimentell erhoben wurden, es angenommen werden kann, dass sie laut den Kriterien zur Durchführung von Experimenten (Kubbe, 2016) objektiv, reliabel und valide sind. Es kann nicht immer von einer höchsten Validitätsstufe ausgegangen werden, dennoch wird Forschenden viel daran gelegen sein, ein hohes Maß an Validität und folglich faktischer Korrektheit nachzuweisen, wenn sie in einer wissenschaftlichen Fachzeitschrift publizieren. Es kommt die weitere Annahme hinzu, dass alle diejenigen Quellen sekundärer Art, die die zuvor genannte Schnittmenge in irgendeiner Art zitiert, ebenfalls als faktisch korrekt angesehen werden kann. Mithilfe dieser Annahmen wäre eine Operationalisierung der faktischen Korrektheit gegeben, gleichzeitig erweckt es die Problematik, dass zwei bereits analysierte Merkmale damit korrelieren würden. Aus diesem Grund und weil zudem die beiden anderen Merkmale zueinander eine höhere Trennschärfe aufweisen, wurde das Merkmal der faktischen Korrektheit an dieser Stelle ausgeschlossen.

Die folgenden vier Merkmale unterscheiden sich hinsichtlich des Ausschlussgrundes von den bereits zuvor analysebedingt ausgeschlossenen Merkmalen. Sie alle könnten zwar analysiert werden, jedoch erfordert die Analyse eine tiefergehende, fachspezifische Analyse, die erstens den Fokus der Arbeit stark verändern würde und zweitens einen Experten auf dem Gebiet bedarf. Beides ist nicht zweckdienlich für den eigentlichen Fokus der Arbeit, weshalb diese zwar erwähnt, jedoch nicht weiter ausgearbeitet werden.

Sprachstil. Die Analyse des Sprachstils könnte unter Zuhilfenahme von Computerlinguistik und linguistischer Marker durchgeführt werden. Dies erforderte jedoch eine Experten-geleitete, tiefergehende linguistische Analyse, welche nicht der Fokus der Arbeit ist.

Wissenschaftlichkeit. Wissenschaftlichkeit kann einen Einfluss auf die wahrgenommene Glaubwürdigkeit haben. Jedoch müssen hierzu die linguistischen Marker bzw. semantischen Bestandteile genauestens herausgearbeitet und die Texte analysiert werden, was im Zuge einer computerlinguistischen Analyse erfolgen könnte. Diese würde den Fokus der Arbeit zu sehr verändern, weshalb dies zwar thematisch erwähnt, aber letztlich nicht durchgeführt wurde.

Argumentqualität. Die bereits beschriebenen Merkmale des Sprachstils und der Wissenschaftlichkeit, welche via linguistische Analyse untersucht werden könnten, könnten ebenfalls mit der Sprachintensität und folglich Argumentqualität zusammen hängen (Craig & Blankenship, 2011; Hamilton & Hunter, 1998). Dies erfordert eine Trennschärfe schaffende Analyse, um die Merkmale ausreichend voneinander abgrenzen zu können. Dies wäre lediglich mithilfe linguistischer Analyse möglich, welche jedoch nicht im Fokus dieser Forschungsarbeit liegt und den Grund zum Ausschluss des Merkmals Argumentqualität lieferte.

Narrative. Da die narrative Analysis ein eigenes Forschungsgebiet darstellt und ähnlich zur Computerlinguistik einen Experten benötigen würde, wurde sich für vorliegende Forschungsarbeit gegen die Integration dieses Merkmals entschieden. Außerdem würde auch hier der Fokus der Arbeit verschoben und dem Thema der Narrative nicht gänzlich gerecht werden.

4.2.4. Operationalisierung der Webseitenmerkmale als Zwischenergebnis der vorbereiteten Analyse

Da die Webseiten überwiegend theoretisch-fundiert und somit deduktiver und im Geringeren auch datengeleitet und somit induktiv³¹ operationalisiert wurden, resultierten klare Muster zur Operationalisierung der Merkmale und Analyse der Webseiten. Die Webseiten wurden daher ausschließlich von der forschenden Person, nicht jedoch einer weiteren beurteilt. Obwohl es gängig ist eine zweite Person für eine Analyse abzustellen, um u. a. die Vergleichbarkeit und Verallgemeinerung im Sinne einer hohen Inter-coder-Reliabilität (O'Connor & Joffe, 2020) zu erhöhen, wird dies nicht für vorliegende Forschungsarbeit berücksichtigt. Denn die Operationalisierung wurde bereits theoretisch fundiert hergeleitet, sodass eine Anleitung einer weiteren Person mit derselben Vorgehensweise und Beurteilungskriterien auch zu den gleichen Ergebnissen geführt hätte.³²

Insofern wurden alle besuchten und selektierten Webseiten einschlägig auf die in der Theorie vorgestellten (siehe Kapitel 2.6.1) verbliebenen Merkmale untersucht. Für die bewertbaren Merkmale finden sich im

³¹ Anpassungen der Operationalisierungen werden je Webseitenmerkmal im weiteren Verlauf des Kapitels thematisiert.

³² Aufgrund dieser Vorgehensweise werden lediglich systematische Fehler ausgeschlossen, nicht jedoch zufallsbedingte Fehler.

Anhang A 1 je Webseitenmerkmal empirische Ergebnisse in Form von deskriptiven Statistiken, die sich ausschließlich auf die Webseiten und ihre Merkmale beziehen. Sie sind lediglich ein Zwischenprodukt und für die eigentliche Analyse und Überprüfung der Hypothesen nicht von Interesse, sollen der Vollständigkeit wegen aber aufgeführt werden. Die Webseitenmerkmale werden im weiteren Verlauf mit den Individuen und somit auf Personenebene verknüpft. Es wird an entsprechender Stelle nochmals auf die Unterscheidung und die weiteren deskriptiven Beschreibungen hingewiesen.

Nachstehend werden die Operationalisierungsschritte der verbliebenen neun Webseitenmerkmale beschrieben.

4.2.4.1. Aktualität

Aktuellere Webseiten werden, je nach Kontext bzw. Thema, als glaubwürdiger erachtet im Vergleich zu weniger aktuellen Webseiten. Für das hier vorliegende Thema – E-Bikes – ist davon auszugehen, dass diese Aussage so getroffen werden kann. Denn die Technologiebranche, zu der das Produkt E-Bike ebenfalls gehört, entwickelte sich gerade in den letzten Jahren sehr schnell und insbesondere innovativ. Weil sich zwischenzeitlich neuere Erkenntnisse aus der schnelllebigen Technologieentwicklung ergeben haben könnten, könnten daher veraltete Informationen als weniger glaubwürdig empfunden werden. Für die Operationalisierung der Aktualität wurde sich daher stark an dem Kontext bzw. Thema orientiert.

Da sich die Studie auf drei verschiedene Kohorten verteilt und sich die genaue Bearbeitung der Probanden nicht uneindeutig auf einen Tag terminieren lässt, wurde aufgrund der Einfachheit die Aktualität wie folgt ermittelt: In einer ersten Analyserunde wurde von allen Webseiten ermittelt, ob überhaupt ein Datum vorhanden ist (0 = nein, 1 = ja). Die eigentliche Angabe des Datums trägt maßgeblich dazu bei, dass die Probanden die Glaubwürdigkeit überhaupt beurteilen können. In einer zweiten Runde wurden alle Webseiten, welche ein Datum enthielten, weiter analysiert. Da die erste Erhebungswelle im November 2020 stattfand, wurde sich zur Operationalisierung an diesem Datum orientiert. Alle Webseiten, die noch das gesamte Jahr 2020 (d. h. ab 01.01.2020) hinsichtlich der Datumsanzeige betrafen, wurden als „aktuell“ (= 2) codiert. Alle Artikel, die vor dem 01.01.2020 geschrieben wurden, wurden als weniger aktuell (= 1) festgehalten. Die Festsetzung von 2020 als Trennzeitpunkt wurde deshalb gewählt, weil in den Jahren zuvor (2018 und 2019) ein ansteigendes Interesse deutlich wurde, erkennbar u. a. an einer enormen Steigung der Produktionszahlen sowie Angaben über den Besitz eines E-Bikes in Deutschland: Abgesehen von dem ohnehin konstanten Anstieg seit 2017 hinsichtlich der genannten Größen, wurde beispielsweise von 2017 auf 2018 mit 46.81 % die größte Steigerung an produzierten E-Bikes verzeichnet (2017: 690 Mio., 2018: 960 Mio.) (ZIV Zweirad-Industrie-Verband. Marktdaten Fahrräder und E-Bikes 2022, 2023, S.

10). Auch die Bestände an E-Bikes wuchs von 2018 mit 5.4 Mio. E-Bikes um 35 % auf 2019 7.1 Mio. E-Bikes (ZIV Zweirad-Industrie-Verband. Marktdaten Fahrräder und E-Bikes 2022, 2023, S. 8). Es ist also erkennbar, dass sich insbesondere in den Jahren 2018 und 2019 ein Boom andeutete. Mit der Festsetzung auf das Jahr 2020 als Trennwert wurden daher zwei Problematiken entgegengewirkt: Die zu dem Zeitpunkt der anfänglichen Boom-Phase (2018 - 2019) getätigten Studien könnten veraltet sein, da sich die Technologie-Branche und somit auch E-Bikes schnell weiterentwickeln. Mit neuen, optimierten Technologien könnten daher auch Änderungen in Bezug auf einen möglichen gesundheitsfördernden Effekt einhergehen, wenn z. B. der Antrieb oder die Akkuleistung verbessert wird, was somit eine bessere Laufleistung und einen veränderten Effekt bedeuten würde. Zweitens könnten die zu diesem Zeitpunkt (2018 - 2019) veröffentlichten Studien eine geringere Vergleichbarkeit und somit Validität aufweisen, denn Gesundheitseffekte zeigen sich u. a. auch als Langzeiteffekte. Wenn direkt zu Beginn der Boomphase Studien mit einer noch nicht ganz so großen Stichprobe durchgeführt wurden, sind sie ggf. weniger aussagekräftig als Studien, die mit einer längeren Dauer und einer größeren Stichprobe liefen, was zu einem späteren Zeitpunkt, und somit 2020, eher gewährleistet sein könnte. Mit einem gewissen Abstand zu dieser Hochphase wird es daher zunehmend realistisch, dass sowohl genügend Datenmaterial (i. S. v. aktiven Nutzern) vorhanden ist als auch erste Erkenntnisse aus der anfänglichen Nutzung gewonnen werden konnten, um die Technologie ausreichend weiterentwickelt zu haben und weiterhin valide Aussagen über einen Zusammenhang zum Gesundheitszustand tätigen zu können und diese auch in aussagekräftigen Studien zu publizieren. Aus diesen genannten Gründen wurde die Aktualität mit Trennwert November 2020 operationalisiert.

4.2.4.2. Verständlichkeit

Gut verständliche Inhalte gehen mit einer höher empfundenen Glaubwürdigkeit einher. Dies wurde eindringlich an theoretischer Stelle in Kapitel 2.2.2.1 dargestellt. Aus diesem Grund wurden die Webseiten bzw. die darauf befindlichen textlichen Inhalte einer Verständlichkeitsanalyse unterzogen. Es ist jedoch notwendig zu beachten, dass es bei der Analyse der Verständlichkeit nicht darum geht, wie gut eine Person den Text verstanden hat. Es wird also nicht die Rezipienten-Seite analysiert, sondern die Seite des Senders einer Information (hier: Text). Die Analyse der Verständlichkeit betrachtet also die formale, textliche Ebene und sieht vor zu ermitteln, welche Indizien der Form nach für oder gegen eine Verständlichkeit sprechen.

Alle inhaltsbezogenen Seiten, die zur Beantwortung der Aufgabe dienlich sein konnten, wurden für die Analyse herangezogen. Ausgeschlossen wurden beispielsweise Startseiten oder Impressums-Seiten, da diese keinen inhaltlichen Mehrwert lieferten. Weiterhin mussten für die Berechnungen deutsche von englischen Texten unterschieden werden, da diese der Natur der

Sprache unterschiedlich sind, so auch in Satz- oder Wortlänge (Amstad, 1978; Beck & Gergel, 2014). Mit der Wahl unterschiedlicher, auf die Sprache bezogenen Formeln wurde diesem Unterschied Rechnung getragen. Basis für die Berechnung war bei englischen Texten der Reading-Ease nach Flesch (1952) und bei deutschen Texten eine angepasste Formel von Amstad (1978), die beide auf der durchschnittlichen Satzlänge (ASL³³) sowie durchschnittlichen Wortlänge (ASW³⁴) beruhen:

$$\text{Index für Deutsch (nach Amstad): } 180 - \text{ASL} - (58.5 \times \text{ASW})$$

$$\text{Index für Englisch (nach Flesch): } 206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$$

Die Satzlängen, Anzahl der Silben, Anzahl der Wörter etc. wurden mithilfe eines Online-Tools ermittelt (gezählt) und für jede Webseite notiert. Mithilfe der unterschiedlichen Formeln konnten deutsche und englische Webseiten den entsprechenden Lesbarkeitsindex erhalten und in die Kategorien eingeteilt werden. Für die Kategorisierung wurde sich hierbei auf die deutsche Variante gestützt, da diese eine siebenstufige Einteilung (Schöll, 2023) und somit eine präzisere und trennschärfere Einteilung erlaubt. Die englischen Texte wurden daher ebenfalls gemäß den deutschen Kategorien eingeteilt. Ein weiteres Argument dafür, die unterschiedlichsprachigen Texte in ein und dieselbe Kategorienskala einzuteilen, ist die Tatsache, dass die ursprüngliche, vierstufige, englische Kategorieneinteilung ähnliche Minimal- und Maximalpunkten (0 und 100) sowie Schwellenpunkten (neue Stufe ab 30 oder 65) aufwies, was sie somit annähernd vergleichbar macht und die Übertragung auf die deutsche Kategorieneinteilung weiter bestärkt. Nach der Kategorieneinteilung der deutschen Formel wurden daher schwer verständliche Texte anschließend mit 0, verständliche Texte mit 1 usw. bis hin zu sehr leicht verständlichen Texten mit 6 festgesetzt.

4.2.4.3. Transparenz

Transparente Informationen führen zu einer hohen Glaubwürdigkeitsbeurteilung, so die theoretische Annahme. Transparenz im Sinne der „Disclosure Transparency“ von Karlsson (2010) bedeutet daher, dass Informationen bereitgestellt werden, die Aufschluss darüber geben, wie die Daten erhoben oder Informationen zusammengetragen wurden. Für die vorliegende Studie wurde daher ermittelt, wie transparent die Informationen auf den Webseiten dargestellt sind. Hierzu wurden die Webseiten in einer ersten Runde untersucht und diagnostiziert, inwieweit es Unterschiede zwischen den Webseiten hinsichtlich einer Transparenz geben könnte. Nach dem ersten Eindruck konnten folgende Aspekte festgesetzt werden: Transparent als solche

³³ ASL = Average Sentence Length. Zur Berechnung wird die Anzahl der Wörter durch die Anzahl der Sätze aus einem Text dividiert.

³⁴ ASW = Average Number of Syllables per Word. Zur Berechnung wird die Anzahl der Silben durch die Anzahl der Wörter aus einem Text dividiert.

waren alle diejenigen Webseiten bzw. Informationen, die Informationen dazu bereitstellen, wie die zu berichtenden Daten zu Stande kommen oder was mit ihnen passiert ist, ehe sie veröffentlicht wurden. Im Detail bedeutet dies, dass jegliche Hinweise zu „redaktionell überprüft“, „journalistisch überprüft“ oder „peer-reviewed“ sowie bei Primärdaten die Beschreibung der Methodik und Datenerhebung dazu führten, dass sie als transparent angesehen wurden. Denn die Autoren und/oder Forschenden geben in dem Moment Einsicht darüber, wie die Daten zu Stande gekommen sind und warum sie als valide gehalten werden können. Sie sind nicht verpflichtet dazu, diese Informationen mit anzuführen. Wenn sie die Informationen trotzdem zusätzlich freiwillig preisgaben, so wurden sie in dieser Kategorie mit „1“ (= transparent) versehen. Alle Webseiten oder Informationen, die keines der oben genannten Anzeichen aufwies, wurden mit „0“ und somit „nicht transparent“ angesehen.

4.2.4.4. Quellenherkunft (primär, sekundär)

Primärquellen wird grundsätzlich eine größere Glaubwürdigkeit zugeschrieben als Sekundärquellen. Aus diesem Grund wurden alle Webseiten bzw. Informationen dahingehend untersucht, ob es sich bei den berichteten Daten um Informationen erster oder zweiter Hand handelte. Primärquellen wurden als solche kategorisiert, wenn erkennbar war, dass die Daten eigenständig erhoben und ausgewertet wurden. Sekundärquellen wurden als solche diagnostiziert, wenn sie über die Datenerhebung anderer berichteten. So waren beispielsweise Artikel auf Nachrichtenseiten tendenziell sekundärer (und somit „0“) und Artikel aus Fachzeitschriften tendenziell primärer Art (und somit „1“).

4.2.4.5. Angabe von Quellen wird zu Nachweisqualität

Wenn Seiten Quellen in Form von Referenzen oder Evidenzen aufweisen, werden diese als glaubwürdiger erachtet im Vergleich dazu, wenn nichts Dergleichen ersichtlich ist (s. Kapitel 2.6.1.5). Aus diesem Grund wurden die Webseiten bezüglich dieses Merkmals untersucht. In einer ersten Runde wurde festgehalten, ob die Webseiten weitere Quellen/Evidenzen/Zitationen enthielten. Eine erste Kategorisierung ergab 0 = nein und 1 = ja, wobei dies zunächst als „Angabe von Quellen“ betitelt wurde. In einer zweiten Runde wurden die Quellen/Evidenzen/Zitationen weiterhin spezifiziert und voneinander abgegrenzt, da sie eine unterschiedliche Qualität aufwiesen, was zum finalen Merkmal „Nachweisqualität“ führte. Hierbei ist anzumerken, dass es für alle Kategorien gleichermaßen egal war, ob die Quellen/Evidenzen/Zitationen im Text eingebaut, im Text verlinkt oder unterhalb des Textes als weiterführende Quellen/Referenzen zu finden waren. Alle Webseiten bzw. Informationen wurden daher nochmals erneut begutachtet, um Unterschiede zu finden. Es kristallisierten sich folgende Unterschiede heraus: Webseiten enthielten keinerlei Angaben (= 0). Webseiten enthielten zwar Quellen oder weiterführende Links z. B. innerhalb eines Textes, diese führten allerdings zu allgemeinen und

zusätzlichen Informationen, die für die Grundaussage der Information nicht entscheidend waren. Ein Beispiel hierfür ist, dass in einem Text erklärt wird, welche positiven Wirkungen Bewegung auf das Herz hat. Hierbei ist eine Verlinkung hinter dem Wort „Herz“ zu finden, was zu einem weiteren Artikel führt, der den Aufbau des Herzes erklärt³⁵. Insofern sind für diese Kategorie zwar Verlinkungen oder Evidenzen enthalten, diese sind aber für die eigentliche Beantwortung der Aufgabe nicht relevant und qualitativ schlechter zu betrachten. Die entsprechende Kategorie wurde mit „1“ festgesetzt. Die dritte Kategorie „2“ enthält ebenfalls Quellen, die sich allerdings auf Meinungen oder Interviews beziehen. Es werden Interviewpartner (z. B. ein Professor bei *Deutsches Gesundheitsportal*³⁶) vorgestellt, die ihre Expertenmeinung zu E-Bikes mitteilen. Diese sind ebenfalls als Evidenzen zu betrachten, stützen sich aber vermehrt auf (subjektive) Meinungen von Personen, als dass sie wissenschaftlich mithilfe eines Experimentes oder Daten belegt wurden. Die letzte Kategorie „3“ sind Evidenzen bzw. Quellenangaben, die zu wissenschaftlichen Studien führen. Diese sind also experimentell mithilfe einer Datenerhebung bzw. -analyse belegt und stützen sich nicht nur auf Expertenmeinungen. Darunter fallen auch die bereits thematisierten Statistiken aus Kapitel 2.6.1.8 (Henke et al., 2019), wobei eine erneute Unterteilung in numerische oder textliche Angabe zu weit geführt hätte und somit nicht erfolgt ist. Artikel aus dieser Kategorie erhielten die höchsten Glaubwürdigkeitszuschreibungen.

4.2.4.6. Kommerzielles Interesse

Wie in der Theorie bereits beschrieben, werden Webseiten, die ein kommerzielles Interesse darstellen, in Form von z. B. Werbeanzeigen, eine geringere Glaubwürdigkeit beigemessen. Ausgehend von der Datenlage (gesamte besuchte Webseiten) wurden die einzelnen Webseiten auf dieses Merkmal wie folgt untersucht: Es wurde in einer ersten Runde festgehalten, ob auf den ersten Blick eine Werbeanzeige ersichtlich ist. Dies wurde dichotom notiert (0 = nein, 1 = ja). In einer zweiten Runde wurde dies weiter spezifiziert. Da Werbeanzeigen auch von unabhängigen Dritten geschaltet werden können, ohne dass die entsprechende Webseite darauf Einfluss nehmen kann, musste die erste Annahme überarbeitet werden: „However, just because a site includes advertising does not necessarily mean that the information contained on the site is not objective³⁷“ (Tate, 2019, S. 50). Aus diesem Grund wurde eine dreiteilige Einteilung verwendet. Dabei besagt die Ausprägung „0“, dass ein kommerzielles Interesse unterstellt werden kann, da es sich entweder um einen E-Bike-Hersteller handelt oder eine E-Bike-relevante Werbung in Form einer E-Bike-Anzeige geschaltet ist. Es ist somit von hohem

³⁵ Zum Zeitpunkt der Erstellung dieser Forschungsarbeit war kein Zugriff mehr auf die Webseite möglich. Ein entsprechender Screenshot als Beleg kann daher nicht angeführt werden.

³⁶ Link zur Webseite: <https://www.deutschesgesundheitsportal.de/2021/05/11/mit-fahrradfahren-das-herz-kraeftigen/>.

³⁷ Übersetzung ins Deutsche: „Die Tatsache, dass eine Website Werbung enthält, bedeutet jedoch nicht zwangsläufig, dass die auf der Website enthaltenen Informationen nicht objektiv sind.“ (Tate, 2019, S. 50).

kommerziellem Interesse auszugehen. Das andere Extrem („2“) betitelt alle Webseiten oder Informationen, die explizit einen Interessenskonflikt (Englisch: Conflict of Interest) ausgeschlossen haben oder eindeutig als Peer-Review-Artikel kategorisiert werden konnten. Dies war insbesondere bei Artikeln von Fachzeitschriften der Fall. Der explizit ausgeschlossene Interessenskonflikt wird von allen Autoren eines Papers unterzeichnet, weshalb von keinem kommerziellen Interesse ausgegangen werden kann. Dass ein Artikel peer-reviewed ist, trägt ebenso dazu bei, dass ein möglich vorhandener Interessenskonflikt in erster Linie im Review erkannt und verbessert werden könnte. Aus diesem Grund ist bei Peer-Review-Artikeln ebenfalls davon auszugehen, dass kein kommerzielles Interesse vorliegt. Alle Webseiten, die weder der Kategorie 0 noch 2 uneindeutig zuzuordnen waren, erhielten die Ausprägung „1“. Für sie gilt somit, dass kein Interessenskonflikt bzw. kommerzielles Interesse explizit ausgeschlossen werden kann, es aber auch keine Anzeichen für ein kommerzielles Interesse gibt, da beispielsweise keine inhaltsbezogenen Werbeanzeigen geschaltet werden. Sie sind somit als neutral zu betrachten.

4.2.4.7. Genre

Informationen oder Medien eines gewissen Genres, wie von Fachzeitschriften oder wissenschaftlichen Datenbanken, werden eine höhere Glaubwürdigkeit zugewiesen als Informationen von Nachrichtenzeitungen oder populärwissenschaftlichen Zeitungen/Zeitschriften. Die gefundenen Medien wurden daher entsprechend einer Medienkategorisierung in unterschiedliche Genres eingeteilt. Die Medienkategorisierung basiert dabei auf einer wissenschaftlichen³⁸ Arbeit (Lauterbach, 2020), die sich ausgiebig mit den unterschiedlichen Genres und deren beigemessener Glaubwürdigkeit auseinandergesetzt hatte. Die Einteilung dieser Medien wurde genutzt, um analoge Einteilungen mit den hierin befindlichen Erhebungsdaten vorzunehmen. Die ursprüngliche Einteilung aus den Vorarbeiten aus BRIDGE belief sich auf 1 bis 5. Im Zuge der Angleichung des Merkmals und Vergleichbarkeit mit anderen Merkmalen wurde sich dafür entschieden, immer mit 0 zu beginnen. Angewandt auf die Einteilung wurde daher diese um 1 reduziert, sodass die Spanne von 0 bis 4 reicht. So wurden beispielsweise Online-Shops mit „0“ klassifiziert. Im Mittelfeld mit beispielsweise dem Wert „2“ befinden sich Nachrichtenseiten. Die höchste Glaubwürdigkeit mit „4“ erhalten Online-Kataloge von Universitäten oder wissenschaftlichen Forschungsinstituten. Die aufgerufenen und selektierten Webseiten aus der E-Bike-Studie wurden auf diese Weise den entsprechenden Medienkategorisierungen zugeordnet, wobei ein höherer Wert in dieser Arbeit mit einer höheren Glaubwürdigkeit assoziiert wird.

³⁸ Die Masterarbeit wurde von Mitarbeitern des Projektteams BRIDGE betreut, wurde jedoch zum Zeitpunkt des hier Verfassten nicht veröffentlicht.

4.2.4.8. Bekanntheit

Bekanntere Webseiten werden mit positiveren Attributen und möglicherweise auch mit einer erhöhten Glaubwürdigkeitsbeurteilung assoziiert. Die Webseiten wurden aus diesem Grund hinsichtlich ihrer Bekanntheit analysiert. Es sei anzumerken, dass nicht ermittelt werden konnte, ob die Probanden die Webseiten tatsächlich kannten. Es wurde stattdessen mithilfe einer Webseite (www.similarweb.com) die in der deutschen Bundesbevölkerung vorherrschende durchschnittliche Bekanntheit ermittelt. Hierzu wurde lediglich die Second-Level-Domain³⁹ in das Suchfeld eingegeben. Der voreingestellte Zeitraum, der aufgrund der kostenlosen Version verwendet werden musste, war April bis Juni 2021. Da dies aber über alle Webseiten hinweg gleich war, spielte der Zeitraum der Beurteilung keine entscheidende Rolle. Die herangezogene Kennzahl war daher der „Durchschnitt der einmaligen Besuche zwischen April und Juni 2021 in Deutschland“. Es wurden für alle Webseiten die Kennzahlen zusammengetragen. Letztlich wurde erneut ein Mittelwert für alle zu analysierenden Webseiten gebildet. Auf Basis dieses Mittelwertes wurden die Webseiten anschließend in zwei Kategorien, nämlich unterhalb des Mittelwertes als weniger bekannt (= 0), und oberhalb des Mittelwertes als bekannter (= 1), eingeteilt.

4.2.4.9. Angabe des Autors

Wenn ein Leser die Herkunft der Informationen mit einem Autor assoziieren kann, wird dem Inhalt bzw. dem Autor eine höhere Glaubwürdigkeit zugerechnet. Aus diesem Grund wurden sämtliche Webseiten dahingehend untersucht, ob vor oder nach dem Artikel bzw. der Information ersichtlich wird, wer der Autor des entsprechenden Artikels ist. War dies erkennbar, als dass zumindest ein Nachname enthalten war, so wurde der Webseite die Kategorie „1 = Autor angegeben“ zugeordnet. War kein Autor angegeben oder mit „unbekannt“ indiziert, so wurde „0 = kein Autor angegeben“ zugeordnet.

4.2.4.10. Blockweise praxisbezogene Zusammenfassung der Webseitenmerkmale

Die theoretische Zusammenfassung der Webseitenmerkmale erzwang an dieser Stelle zur Operationalisierung eine Entscheidung. Denn, wie bereits ausgeführt, sind zum einen auf Basis der Vorüberlegungen zur Operationalisierung einige Webseitenmerkmale nicht mehr anwendbar und zum anderen befinden sich immer noch manche in den Schnittmengen, die in der praktischen Umsetzung zu Problemen führen würden; Probleme im Sinne davon, dass die Zurechnung der sich doppelnden Webseitenmerkmale zu mehreren Kategorien zu starken Korrelationen zwischen den Kategorien führen würden. Betreffend den ersteren Punkt wurde

³⁹ Die Second-Level-Domain bezeichnet den Kern einer Webseite, d. h. zwischen „www“ und „com“. Bei <https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/> wäre „quarks“ die entsprechende Second-Level-Domain.

die Übersicht bezüglich der Problematik aktualisiert, d. h. die nicht mehr zu analysierbaren Webseitenmerkmale (z. B. faktische Korrektheit) wurden entsprechend gekennzeichnet. Bezüglich des zweiten Punktes musste entschieden werden, wie die sich überschneidenden Webseitenmerkmale zu handhaben sind. Trotz des Ausschlusses aus den Modellberechnungen werden aus Gründen der Vollständigkeit sowohl die Struktur der Information sowie Webseiten – als mittlerweile separiert betrachtet – als auch die Reputation in die für sie thematisch passenden Merkmalskategorien eingeordnet.

Die *Struktur der Webseite* wird demnach in die Kategorie der Quellen- bzw. Medienmerkmale eingeteilt, da sie kennzeichnet, wie gut ein Medium Inhalte darstellt. Die Darstellung kann hierbei zum Beispiel mithilfe von Absätzen, Überschriften, oder Variation in visuellen Highlights (z. B. Fettmarkierungen) erfolgen. Hierbei ist weder der Autor daran beteiligt, noch sind dies Informationsmerkmale, da die Betreiber einer Webseite oder eines Mediums ihre Medienplattform so gestalten können, wie sie es für ansprechend halten. Die *Struktur einer Information* hingegen wird in die Kategorie der Nachrichten- bzw. Informationsmerkmale eingeordnet, da sie die textliche bzw. inhaltliche Struktur einer Information wiedergibt.

Die *Reputation* bzw. der Ruf kann sich einerseits auf das Medium und andererseits auf den Verfasser eines Textes, also Autor, beziehen. Es wurde sich für die Einteilung in die Kategorie der Medienmerkmale entschieden, da im Durchschnitt eher noch Rückschlüsse auf die Reputation von Medien im Vergleich zu einzelnen Autoren möglich wären. Dass jemand zufällig genau einen Autor eines Artikels erstens kennt und zweitens die Reputation beurteilen kann, ist tendenziell unrealistisch. Aus diesem Grund ist es konzeptionell sinnvoller, die Reputation kategorisch zu den Medien zuzurechnen.

Kommerzielles Interesse und *Bekanntheit* befinden sich ebenfalls als Überschneidung zwischen den Quellen- bzw. Medienmerkmalen und Autorenmerkmalen in der bisherigen Übersicht. Diese beiden Merkmale werden zu erstgenannter Kategorie zugeordnet, was folgendermaßen begründet wird: Die Bekanntheit eines Autors oder sogar der Information (d. h. des Themas) können zwar einen Einfluss auf die wahrgenommene Glaubwürdigkeit haben, jedoch konnte anhand der Daten und Operationalisierungsmöglichkeiten nur annäherungsweise (auch nur unter der getätigten Annahme, dass es ein Durchschnittswert ist) die Bekanntheit des Mediums gemessen werden. Eine derartige Webseite für Autoren ist nicht verfügbar und es wurde auch nicht ermittelt, ob die Probanden einerseits den Autor oder weiterhin sogar das Thema kennen.

Das kommerzielle Interesse eines Autors kann ebenfalls eine Rolle bei der Glaubwürdigkeitsbeurteilung spielen, jedoch sind auch hier die Operationalisierungsmöglichkeiten begrenzt. Aus der Analyse hinsichtlich des Merkmals sind die Unterteilungen eruiert, die sich lediglich

auf das Medium beziehen. Zwar kann ein Autor ebenfalls explizit ein kommerzielles Interesse ausschließen. Dies war jedoch an keiner Stelle auf den Webseiten explizit enthalten. Wenn auf einer Webseite der Interessenskonflikt explizit ausgeschlossen war, konnte dies keinerlei Rückschlüsse erlauben, ob dies für den Autor oder das gesamte Medium der Fall war. Da nicht alle Webseiten Autoren angaben, aber auch in der Tat Interessenskonflikte erkennbar waren (da z. B. E-Bike Shop), wurde dieses Merkmal in die Kategorie Quellen- bzw. Medienmerkmal gezogen.

Die aktualisierte Einteilung der Webseitenmerkmale ist in Abbildung 17 dargestellt. Sie werden nachstehend für zukünftige Berechnungen blockweise zusammen betrachtet.

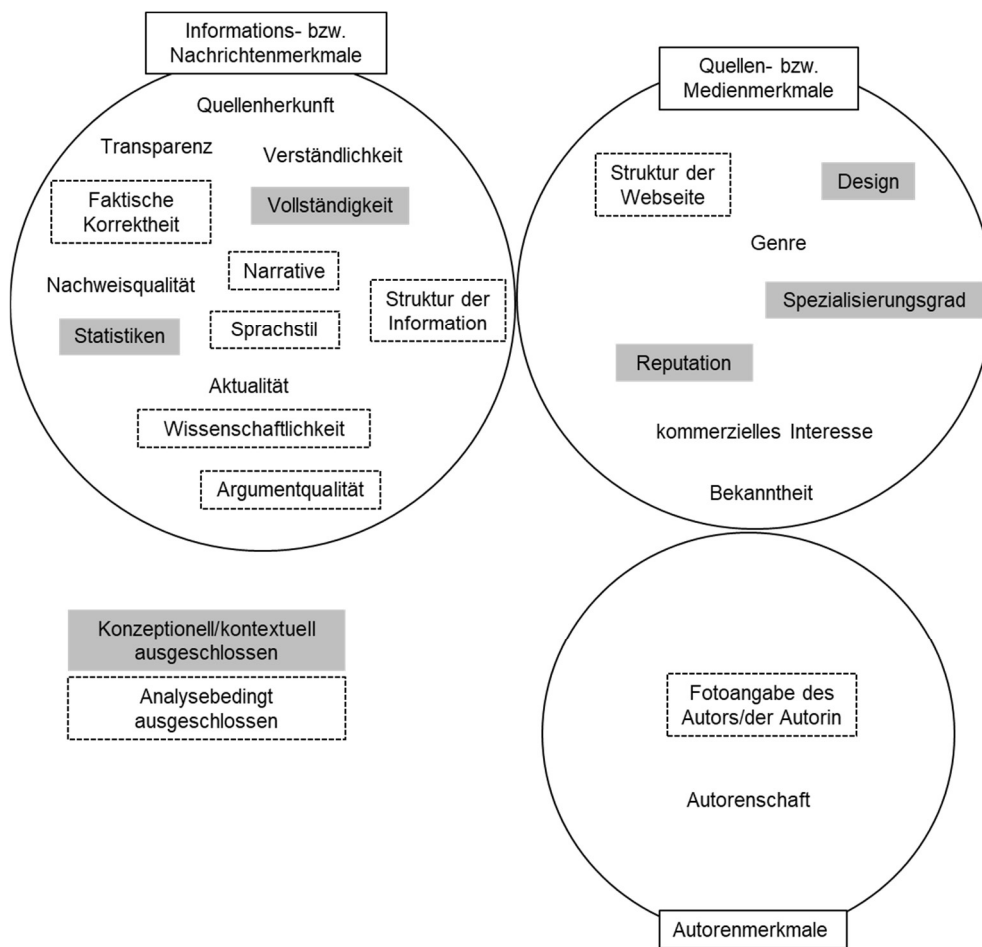


Abbildung 17: *Finale Einteilung der Webseitenmerkmale bezüglich Kategorien.*

4.2.4.11. Zwischenfazit zu den verbliebenen Webseitenmerkmalen

Vorliegende Forschungsarbeit startete mit dem Ziel, ein holistisches Modell abzubilden, in der sowohl Personenmerkmale als auch möglichst viele Webseitenmerkmale integriert sind. Aus der Literatur theoretisch herausgearbeitete Webseitenmerkmale ergaben 22 Merkmale, wovon fünf kontextuell-konzeptionell und acht analysebedingt nicht anwendbar und nun final noch neun übrig sind. In Anbetracht der anfänglich herausgearbeiteten 22 Merkmale ist es absolut

betrachtet erst einmal unbefriedigend. Unter Betrachtung der genauen Gründe für den Ausschluss ist zu sagen, dass das Ziel eines holistischen Modells zu Beginn an zu hoch angesetzt war, das Erreichen eines annähernd holistischen Modells dennoch nach wie vor möglich ist. Bestehende Forschende haben nicht ohne Grund nur ein oder zwei Merkmale gleichzeitig untersucht. Denn es zeigt sich, dass die Merkmale stets vom Kontext (z. B. Thema, Stichprobe) abhängig sind und nicht eins zu eins verglichen werden können. Funktioniert die *faktische Korrektheit* beispielsweise bei einem uneindeutigen Problem, z. B. bei der Suche nach einer bestimmten historischen Persönlichkeit, wird dies jedoch problematisch, wenn eine Abwägung einer Fragestellung (Sind E-Bikes gesundheitsfördernd?) gefordert ist. Verwenden Probanden immer die gleichen Webseiten, die an sich schon die gleichen Webseitenmerkmale (z. B. *kein Foto des Autors gegeben*) beinhalten, kann der Einfluss dieser Merkmale nicht untersucht werden, weil sie nicht variieren. Dennoch werden die verbliebenen neun Merkmale genutzt, um die möglichen Einflüsse, auch im Zusammenhang mit den Personenmerkmalen, auf das Verhalten bei der Lösung eines Informationsproblems online zu ermitteln. Dem aktualisierten Ziel, einem annähernd holistischen Modell abzubilden, soll damit nähergekommen werden.

4.2.4.12. Personenbezogene Glaubwürdigkeitsindices auf Basis der Webseitenmerkmale

Um einen Zusammenhang zwischen Personen- und Webseitenmerkmalen feststellen zu können, müssen die Webseitenmerkmale exakt den Personen zugeordnet werden können. Denn bis zum jetzigen Stand sind die Webseiten und Personen zwei unterschiedliche Betrachtungsaspekte, die aber in einem Modell zusammen die abhängigen Variablen erklären sollen. Folglich wurden die Webseitenmerkmale aus dem Webseitendatensatz in den Personendatensatz integriert, indem Merkmale einer Webseite, die eine Person nutzte, dieser Person zugewiesen wurde. Nachstehend wird anhand eines Beispiels erläutert, wie dieser Umwandlungsprozess von Statten geht.

Wahrgenommene Glaubwürdigkeit der besuchten Webseiten (PIXb). Anhand der Logdaten wurde ersichtlich, welche Webseiten aufgerufen wurden. Die Logdaten enthielten auch die Eintragung „scrolled up/down“, die anzeigte, ob die Probanden auf der Webseite scrollten. Zunächst wurden nur solche Webseiten betrachtet, bei denen ein Eintrag dieser Art vorzufinden war. Nach Sichtung aller Webseiten wurde allerdings ersichtlich, dass manche Webseiten bereits auf einen Blick alle Informationen enthielten, sodass es keine Notwendigkeit gab zu scrollen. Aus diesem Grund wurde diese Betrachtungsweise wieder verworfen und es wurden alle Webseiten zu Rate gezogen, was als Limitation kritisch diskutiert werden wird. Wie bereits erwähnt, fielen in die Betrachtung jedoch nur solche, die einen konkreten Inhaltsbezug zur Aufgabenstellung hatten (s. Kapitel 4.2).

Mithilfe der so erhobenen Prozessdaten konnte ermittelt werden, welche Webseiten und in welcher Reihenfolge die Probanden diese aufrufen. Im dargestellten Beispiel hat Proband *ID70* insgesamt drei Webseiten aufgerufen, die mit den Kurztiteln *IKAÖ* Artikel (Spalte B), *Quarks* Artikel (Spalte C) und *Envivas* (Spalte D) Artikel bezeichnet werden (Tabelle 6). Anschließend wurde die literaturbasierte Operationalisierung der Webseitenmerkmale verwendet, um einen Probanden pro Merkmal einer Webseite einen Wert zuzuordnen. Im Beispiel werden zur Vereinfachung lediglich drei Webseitenmerkmale (Genre, kommerzielles Interesse und Bekanntheit) dargestellt. Im Beispiel erhält der Proband *ID70* für die Webseite *IKAÖ* und das Merkmal Genre den Wert 5, da es sich bei dieser Webseite um ein wissenschaftliches Forschungsinstitut handelt (Zelle 2B). *Quarks* und *Envivas* erhalten für das Merkmal Genre jeweils den Wert 3 (Zellen 2C, 2D), da es sich bei ihnen um eine Nachrichtenseite und einen Verband handelt. Werden diese Werte addiert, ergibt sich für den Probanden die Summe 11 für das Merkmal Genre (Zelle E2). Da die Probanden nicht allesamt immer drei Webseiten besucht haben und dadurch kein interindividueller Vergleich möglich gewesen wäre, wurde die Summe durch die jeweilige, probandenspezifische Anzahl der aufgerufenen⁴⁰ Webseiten dividiert. Dies resultierte bei Proband *ID70* in dem Index 3.66 (Zelle 2F). Im Zuge der gleichen Gewichtung musste der Index anschließend durch das Maximum aus der jeweiligen Merkmalskategorie, bei Genre 5 (Zelle 2G), geteilt werden, sodass alle Werte zwischen 0 und 1 und somit gleichgewichtig werden. Es resultiert für Proband *ID70* somit ein Wert von 0.732 (2H). Dieser *PIXbGenre* sagt aus, dass der Proband solche Webseiten aufgerufen hat, die bei der Bewertung der wahrgenommenen Glaubwürdigkeit hinsichtlich des Merkmals Genre im Durchschnitt den Wert 0.732 aufweisen. Für die beiden anderen Merkmale ergeben sich nach analoger Berechnung die Werte 0.5 für kommerzielles Interesse *PIXbkommerziellesInteresse* (3H) und 0.33 für Bekanntheit *PIXbBekanntheit* (4H).

Tabelle 6: Berechnungsbeispiel des personenbezogenen Glaubwürdigkeitsindices.

	A	B	C	D	E	F	G	H
	Proband <i>ID70</i>	1. Link: <i>IKAÖ</i> Artikel	2. Link: <i>Quarks</i>	3. Link: <i>Envivas</i>	Summe der Be- wertun- gen	Index (di- vidiert durch Anzahl Websei- ten, d. h. 3)	Maxima der jeweiligen Ka- tegorie der Webseiten- merkmale	Dividiert durch die Maxima der jeweiligen Kategorie
Websei- tenmerk- 1	Genre	5	3	3	11 (5+3+3)	3.666 (11/3)	5	0.732 (3.666/5)
2	Kommerzielles Interesse	1	1	1	3	1	2	0,5
3	Bekanntheit	0	1	0	1	0.333	1	0.33
4								

⁴⁰ Manche aufgerufenen Webseiten wurden aus der Summierung ausgeschlossen, da sie nicht mehr abrufbar waren oder die Probanden keinen Zugriff darauf hatten. Um den Index somit fair zu halten, wurden diese nicht mitberechnet.

Eine letzte Datenoptimierung wurde vorgenommen, da die Webseitenmerkmale, wie bereits in Kapitel 4.2.4.10 erklärt, nicht einzeln, sondern blockweise in die Modelle aufgenommen wurden. Am Beispiel des Probanden *ID70* liegen alle drei Webseitenmerkmale im Block Medium.

Die Werte der personenbezogenen Glaubwürdigkeitsindices pro Merkmal wurden addiert und zum Zwecke der Vergleichbarkeit mit den anderen Blöcken durch die Anzahl geteilt:

$$\frac{\text{PIXbGenre} + \text{PIXbkommerziellesInteresse} + \text{PIXbBekanntheit}}{\text{Anzahl der Merkmale}}$$

$$\frac{0.732 + 0.5 + 0.333}{3} \sim 0.522$$

Die hieraus resultierende Zahl gibt an, wie hoch der wahrgenommene Glaubwürdigkeitswert der mediumsbezogenen Merkmale der *besuchten* Webseiten von Proband *ID70* ist (*PIXbMedium*). Der Wert von *PIXbMedium* bedeutet, dass der Proband *ID70* Webseiten aufgerufen hat, die bezüglich mediumsbezogener Merkmale eine wahrscheinlich, subjektiv wahrgenommene Glaubwürdigkeit von 0.522 aufweisen. Im Vergleich mit anderen Probanden mit unterschiedlichen Merkmalen und im Zusammenhang mit den Prozess- und Performanzindikatoren kann dem Wert eine Bedeutung beigemessen werden.

Beispielsweise könnte eine Aussage eines vereinfachten Personenvergleichs wie folgt aussehen: Ein Proband *ID64* existiert mit einem *PIXbMedium* von 0.21 in den Daten. Somit hat Proband *ID64* im Vergleich zum Probanden *ID70* durchschnittlich Webseiten besucht, die weniger von den subjektiv wahrnehmbaren und hier mediumsbezogenen Merkmalen für eine positive Glaubwürdigkeitsbeurteilung bereitstellen, und somit auch als weniger glaubwürdig wahrgenommen werden könnten. Anders ausgedrückt besuchte der Proband *ID64* Webseiten, die weniger aufgrund des Genres, kommerzielles Interesse oder Bekanntheit für eine positive Glaubwürdigkeitsbeurteilung sprachen im Vergleich zum Probanden *ID70*.

Wahrgenommene Glaubwürdigkeit der selektierten Webseiten (PIXs). Im Unterschied zu den besuchten Webseiten wurden die selektierten Webseiten im Antwortteil A1 kenntlich gemacht. Anhand der Angaben aus A1 wurde eine Excel-Tabelle erstellt, die aufzeigt, welche Webseiten und in welcher Reihenfolge diese für die Aufgabenstellung ausgewählt wurden. Für die anschließende Berechnung der wahrgenommenen Glaubwürdigkeitsindices der einzelnen Merkmale (z. B. *PIXsVerständlichkeit*) oder die blockweise Zusammenfassung der Merkmale (*PIXsText*) wurde die Berechnung analog zu oben durchgeführt.

PIXs bedeutet, wie hoch der wahrgenommene Glaubwürdigkeitswert der textlichen Merkmale der *selektierten* Webseiten von einem Probanden ist (*PIXsText*). Z. B. sagt ein Wert von 0.6 aus, dass ein Proband Webseiten selektiert (also in Teilaufgabe 1 angegeben) hat, die bezüglich textlicher Merkmale eine wahrscheinlich, subjektiv wahrgenommene Glaubwürdigkeit von

0.6 aufweisen. Auch hier wird dem Wert erst eine sinnvolle Bedeutung beigemessen, wenn dieser im Vergleich zu anderen Probanden, Personenmerkmalen oder Performanz- und Prozessindikatoren betrachtet wird.

4.2.5. Generierung der Prozess- und Performanzindikatoren als abhängige Variablen

Zur Generierung der abhängigen Variablen *Spezifizität der Suchbegriffseingabe* wurden die Logdaten als Datenbasis zu Grunde gelegt. Für jeden Probanden wurde ausgewertet, welche Wörter und in welcher Reihenfolge diese in die Suchmaske eingegeben wurden. Alle Begriffe wurden anschließend qualitativ analysiert und in zwei Gruppen eingeteilt, welche sich induktiv aus den Daten ergaben. Die abhängige Variable liegt somit in den Ausprägungen *allgemein* und *spezifisch* vor. Die Daten aus der Sucheingabe wurden ohne jegliche Erwartungen oder Theorien gesichtet, um Voreingenommenheit i. S. eines Bestätigungs-Bias zu reduzieren (Gioia et al., 2013)⁴¹. Welche Kategorien sich aus den Daten ergaben, wurde erst im Verlauf der Analyse ersichtlich und somit induktiv ermittelt. Dabei wurden alle Suchanfragen und dementsprechende Suchbegriffseingaben zusammengefasst, die sich thematisch am ähnlichsten und zu anderen Kategorien am unähnlichsten waren. Irrelevante Kategorien, wie das Suchen einer allgemeinen Frage aus dem Fragebogen („Ein Ball und ein Schläger kosten...“⁴² oder „die Umrechnung der Juranote“)⁴³ wurden für die Analyse nicht beachtet, da sie nicht aufgabenrelevant waren. Suchanfragen, die dabei die Wiederholung, Kopie der Fragestellung „Sind E-Bikes gesundheitsfördernd“ oder die Schlagwörter (E-Bikes, Gesundheitsförderung) enthielten, wurden der allgemeinen Kategorie zugeordnet. Die Probanden dieser Kategorie kopierten lediglich die Fragestellung oder formulierten sie leicht um, ohne einen höheren Aufwand zu betreiben. Die zweite Kategorie der spezifischen Suchanfragen umfasst Anfragen, welche über die Fragestellung oder Schlagwörter hinausgingen. Dies ist z. B. durch die Zusätze „Studie“, „Effekt“ oder „Outcome“ der Fall. In die gleiche Kategorie wurden auch Anfragen geordnet, welche eine konkrete Quelle enthielten, wie z. B. *PubMed*⁴⁴. Im weiteren Schritt wurden die gefundenen Kategorien auf die Probanden und den Suchprozess transferiert, sodass nachvollzogen werden konnte, ob die Probanden beispielsweise mit allgemeinen Anfragen beginnen, um am Ende zu spezifischeren Anfragen über zu gehen. War dies der Fall, so erhielten die Probanden die Bewertung „spezifisch“. Wurde direkt spezifisch begonnen und im Weiteren

⁴¹ Der Bestätigungs-Bias (Englisch: Confirmation Bias) (Wason, 1968, S. 280) ist eine kognitive Verzerrung, bei der in Aussagen gezielt nach einer Bestätigung der eigenen, zuvor getätigten Annahmen gesucht wird und Informationen somit nicht objektiv beurteilt werden (M. Oswald & Grosjean, 2004, S. 79 ff.).

⁴² Die Probanden sollten den Cognitive Reflection Test nach Frederick (2005) lösen. Manche Probanden suchten im Internet nach der Lösung, weshalb dieser Eintrag in den Logdaten erscheint.

⁴³ Die Probanden der Domäne Jura sollten im Rahmen des Fragebogens angeben, welche Note sie im ersten Staatsexamen hatten. Aus diesem Grund suchten manche Probanden nach einem Umrechner der erreichten Punkte zur Note.

⁴⁴ *PubMed* (*PubMed*, o. J.) ist eine fachspezifische, medizinische Datenbank, über die wissenschaftliche Artikel gesucht werden können.

nochmal allgemein oder von vorneherein nur spezifisch gesucht, waren die Probanden ebenfalls dieser Kategorie zuzuordnen. Wenn die Probanden lediglich auf einer oberflächlichen Ebene suchten und somit allgemeine Suchanfragen stellten, ohne spezifischer zu werden, wurden sie der Kategorie „allgemein“ zugeordnet. Wichtig zu erwähnen ist, dass nicht bewertet wurde, ob oberflächlich oder spezifisch gesucht wurde. Probanden könnten auch mit einer allgemeinen, spezifischen Anfrage ausreichend gute, glaubwürdige Informationen gefunden haben und deshalb nicht nochmals erneut gesucht oder die Suchbegriffe spezifiziert haben. Was allerdings einen bewertenden Charakter hat ist die Tatsache der Effizienz. Mit einer oberflächlichen Suche finden sich ggf. die gleichen, guten Informationen, es dauert aber möglicherweise länger. Insofern ist im Zuge einer idealen und effizienten Informations-Recherche eine spezifische Suchanfrage wünschenswert.

Nachrecherche. Anhand der Logdaten und dem Eintrag „Active Tab changed“ sowie der Zeitstempeln konnte erkennbar werden, ob ein Proband eine vorher aufgerufene Webseite oder Information auf dieser nochmals in der Suchmaschine gesucht hatte. Hatte der Proband genau die zuvor gefundene Webseite in der Suchmaschine eingegeben und somit gesucht, wurde der Versuch einer Überprüfung der Quelle, also laterales Lesen (Breakstone, Smith, Connors, et al., 2021; Wineburg & McGrew, 2019) unterstellt. Dabei ist es unerheblich, wie häufig das passiert. Hatte ein Proband dies mindestens einmal für mindestens eine Webseite gemacht, zählte dies bereits zu einem unterstellten Versuch der Nachrecherche, was mit 1 annotiert wurde. Für die noch in der Theorie (Brand-Gruwel et al., 2009) ermittelte Annahme, dass auch gefundene Informationen auf einer Webseite zur Evaluation erneut gesucht werden würden, konnten keine entsprechenden Logdaten verzeichnet werden. Insofern bezeichnet der Teil dieser abhängigen Variablen nur die unterstellte Überprüfung der Quellenglaubwürdigkeit, und nicht der gefundenen Informationen. Der zweite Teil dieser abhängigen Variablen repräsentiert das Anklicken des Impressums und/oder der „Über Uns“-Webseite als Nachrecherche-Verhalten. Wenn ein solcher Klick stattgefunden hatte, wurde dies mit einem neuen Eintrag in den Logdaten erkennbar. Es wird daher unterstellt, dass der Proband auch bei solch einem Verhalten eine Nachrecherche über die Quelle getätigt hat, was ihm die Zuteilung zu 1 erbrachte. Ein Proband erhielt folglich für diese abhängige Variable die Annotation 1, wenn er entweder die Quelle erneut suchte oder die „Über-Uns“-Impressums-Seite anklickte oder beides tat.

Weiterverlinkung anklicken. Wie zuvor sind die Logdaten die Basis zur Generierung dieser abhängigen Variablen. Dank der Zeitstempel innerhalb der Logdaten sowie der Einträge in den Logdaten selbst konnte nachvollzogen werden, welche Webseiten die Probanden in welcher chronologischen Reihenfolge aufrufen. Der Eintrag „Active Tab changed“ zeigte an, dass die Probanden von einer Webseite, welche sie zuerst aufrufen und offen hatten, direkt auf eine weitere Webseite weitergeleitet wurden. Diese Übergänge wurden mithilfe der ersten und

zweiten besuchten Webseite analysiert. Dabei wurde festgestellt, dass sich auf der ersten Webseite Verlinkung zu den entsprechenden zweiten Webseiten befanden. Insofern ist davon auszugehen, dass die Probanden, welche solche Übergänge in den Logdaten hatten, diese Links angeklickt haben mit der Intention, sie zu verfolgen und zu überprüfen, sei es zum Zweck der technischen oder inhaltlichen Überprüfung. Wenn ein Proband solch ein Verhalten mindestens einmal in den Logdaten aufwies, wurde 1 zugeteilt. Dies bedeutet, dass der Proband die Weiterverlinkung anklickte. 0 hingegen bedeutet, dass kein einziges Mal aus den Logdaten ersichtlich wird, dass ein solcher Übergang vorhanden ist und somit auch keine Weiterverlinkungen angeklickt wurde.

Korrektheit der Quelleneinschätzung. Das bereits erwähnte Bewertungsschema des BRIDGE-Projektes beinhaltet eine Teilbewertung, welche diese abhängige Variable darstellt. Alle von den Probanden verwendeten Quellen wurden im Vorfeld von drei unabhängigen und ausgebildeten Ratern auf einer ordinalskalierten Skala (0 bis 4) bewertet. Diese erhielten zur korrekten Durchführung des Ratings eine Schulung und konnten mithilfe dessen die verwendeten Quellen anhand von Aktualität, Qualität und inhaltliche Eignung bewerten. Der Kendall'sche Konkordanzkoeffizient (Kendall & Smith, 1939) wurde aufgrund der Ordinalskalierung und der Anzahl der Rater ($n = 3$) herangezogen, um die Qualität der Raterübereinstimmung zu ermitteln. Die Teststatistik war statistisch signifikant ($p < 0.001$) mit einem Kendall $W = .445$, was einer moderaten⁴⁵ Übereinstimmung entspricht (Landis & Koch, 1977).

Die Raterbewertung wurde mit der Bewertung der Probanden verglichen und auf einer Bewertungsskala von 0 bis 4 Punkten erfasst. 0 bedeutet, dass überhaupt keine Einschätzung der Quelle stattgefunden hat, wohingegen die korrekte Einschätzung der Quelle (und somit Übereinstimmung mit der Bewertung der Rater) mit 4 Punkten bewertet wurde⁴⁶. Korrekte Quelleneinschätzung bedeutet, dass eine glaubwürdige Quelle korrekt als glaubwürdig bzw. eine unglaubwürdige Quelle korrekt als unglaubwürdig identifiziert wurde. Je größer somit der Wert, desto „richtiger“ die Einschätzung der Quelle.

Systematische Kategorie zur Begründung der Quelleneinschätzung. Die zweite abhängige Variable aus der Evaluationsstrategie wurde aus den Antworten der Probanden zu Teilaufgabe 1 (Angabe der URL und Beurteilung der Glaubwürdigkeit mit Begründung) abgeleitet. Alle Begründungen für oder gegen die Glaubwürdigkeit einer Quelle wurden qualitativ analysiert und in der ersten Runde nah am Wortlaut der verfassten Probandenantwort beschrieben (s. Tabelle 7 Abschnitte „Beispielsätze“ und „Codes erster Ebene“) (Murphy et al., 2017). In der

⁴⁵ Komplette Einteilung: $< .001$ = schlecht; $.01 - .20$ = leicht; $.21 - .40$ = angemessen; $.41 - .60$ = moderat; $.61 - .80$ = erheblich; $.81 - 1.00$ = fast perfekt.

⁴⁶ Weitere Skalenpunkte: 1 = (Teilweise) falsche Einschätzung der Quellen; 2 = Teilweise richtige Einschätzung der Quellen; 3 = Größtenteils richtige Einschätzung der Quellen.

zweiten Runde wurde versucht, Gemeinsamkeiten und Unterschiede zwischen den Codes erster Ebene und übergeordnete Kategorien zu finden. Dieser Prozess verlief iterativ, denn die Begründungen der Probanden waren nicht immer uneindeutig voneinander zu trennen, weshalb die Kategorien und -zuordnungen über mehrere Runden verändert und so optimiert wurden, was, gemäß dem Vorteil einer qualitativen Analyse, in einer ausreichenden Tiefenschärfe resultieren kann (Moschner & Anschütz, 2010). Es ergab sich induktiv eine Einteilung in final drei Kategorien (Text, Autor, Medium), die als Kategorien auf der zweiten Ebene festgesetzt wurden und zu dem vorher herausgearbeiteten Modell auf Basis des 3-S-Modells passten (s. Tabelle 7 Abschnitt „Kategorien zweiter Ebene“). Um eine möglichst große Reliabilität zu erhalten (Mayring & Fenzl, 2019), wurden in einer letzten Runde mit einem Zeitabstand von einer Woche und ohne vorherige Begutachtung der Kategorien alle Codes erneut auf sinnhafte Zugehörigkeit überprüft.

Tabelle 7: Ergebnisse der qualitativen Analyse der Probandenantwort zur Glaubwürdigkeit.

Beispielsätze der Probanden aus A1					
„Differenzierte, faktenbasierte Darstellung“ (204, Pos. 1)		„zuverlässig da die Quellen im Text hinterlegt werden“ (143, Pos. 1)	„Zudem wird der Autor der Studie genannt“ (35, Pos. 1)	„Ich halte diese Quelle, auch wenn sie eher populärwissenschaftlich ist, für zuverlässig“ (37, Pos. 1)	
„Jedoch nicht nur positiv geschrieben, sondern auch die Risiken gelistet“ (127, Pos. 1)		„im Text selbst wird auf verschiedene Quellen verwiesen“ (173, Pos. 1)	„Außerdem wird der Autor des Artikels“ (69, Pos. 1)	„...das Ärzteblatt ist eine etablierte Informationsquelle für medizinische und gesundheitsbezogene Informationen, jedoch sehr breit gefächert“ (146, Pos. 1)	
„auch negative Faktoren werden aufgezählt“ (168, Pos. 1)			„Sie gibt weiterhin an, von welcher Medizin-Journalistin sie geschrieben und von welchem Redakteur sie überprüft wurde. Beide Personen können nachrecherchiert werden.“ (99, Pos. 1)		
Codes erster Ebene:					
„beidseitige Darstellung“		„Verlinkung der Quellen“	„Autor angegeben/erkennbar“	„Populärwissenschaftlich / Genre“	
Kategorien der zweiten Ebene					
Text		Autor		Medium	
Einordnung der Begründungen in angepasstes Modell des Forschungsvorhabens					
Systematisch	Heuristisch (auch Oberflächenmerkmale)	Systematisch	Heuristisch (auch Oberflächenmerkmale)	Systematisch	Heuristisch (auch Oberflächenmerkmale)

Anmerkung.

Die Nummerierung inklusive Positionsangabe hinter den jeweiligen Probandenangaben bezieht sich auf den in MaxQDA Plus (Release 22.3.0 vom 18.10.2022) hinterlegten Datensatz, in welchem die Probandenantworten analysiert wurden.

Wie bereits in Kapitel 2.5.2 (s. Tabelle 1) eruiert, wurde in Anlehnung an das 3-S-Modell von Lucassen und Schraagen (2011) im Vorfeld eine eigene, angepasste Kategorisierung vorgenommen. Diese wurde auf die drei induktiv gefundenen Kategorien Semantik/Text, Quelle/Autor und Quelle/Medium angewandt und im Anschluss in jeweils „heuristisch“ oder „systematisch“ eingeordnet (s. Tabelle 7 Abschnitt „Einordnung und Begründung“).

Resultierend hieraus ergab sich für die einzelnen Kategorien die in Tabelle 8 dargestellte Verteilung. Es ist zu erkennen, dass der Großteil (> 53 %) der Codes bezüglich des Mediums als

Quelle und heuristisch getätigt wurde. Eine personenbezogene Häufigkeitsverteilung für die einzelnen Codes wurde an dieser Stelle ausgelassen, da dies lediglich ein Zwischenschritt zu Generierung der eigentlichen abhängigen Variablen darstellte.

Tabelle 8: *Häufigkeitsverteilung der Codes*

Verteilung der Codierungen	<i>n</i>
Codes aus dem Bereich Text systematisch	204
Codes aus dem Bereich Text heuristisch	129
Codes aus dem Bereich Autor als Quelle systematisch	0
Codes aus dem Bereich Autor als Quelle heuristisch	11
Codes aus dem Bereich Medium als Quelle systematisch	0
Codes aus dem Bereich Medium als Quelle heuristisch	394
Gesamtanzahl der Codes	738

Anschließend wurden die Codierungen gemäß der heuristischen und systematischen Verarbeitung pro Probanden erfasst. Da sich die Anzahl der gesamten Begründungen der jeweiligen Probanden unterschied ($Min = 0$, $Max = 14$, $M = 5.11$, $SD = 3.014$), hätte eine Auszählung der absoluten, einzelnen Begründungen aus den korrespondierenden Kategorien zu keinem vergleichbaren Ergebnis geführt. Die Überlegung, ein Verhältnis der einzelnen Kategorien zu der Gesamtsumme der angegebenen Begründungen zu bilden, wurde verworfen und sich stattdessen für die Dichotomisierung der Variablen entschieden. Dies wird entsprechend in den Limitationen nochmals thematisiert. Folglich bildet diese abhängige Variable eine nominalskalierte Variable mit den Ausprägungen 0 und 1. Ersteres steht ausschließlich für die Angabe von Begründungen aus der heuristischen Kategorie zur Beurteilung der Quellenglaubwürdigkeit und letzteres für mindestens eine systematisch-kategorisierte Begründung.

Argumentationsstruktur. Die Basis zur Generierung der abhängigen Variablen Argumentationsstruktur bilden die Antworten der Probanden zur Lösung der Teilaufgabe 2 (begründetes Argument, ob E-Bikes gesundheitsfördernd sind). Mithilfe eines induktiv-deduktiven Ansatzes wurden die Textdaten qualitativ analysiert. Ausgangspunkt für die deduktive Analyse stellt das bereits validierte und in der Praxis angewandte Bewertungsraster des Projektes BRIDGE (Zlatkin-Troitschanskaia, Brückner, et al., 2021) dar, welches die grundlegenden Bestandteile einer guten Argumentation in Anlehnung an Toulmin (2003) anführt. Das daraus resultierende Bewertungskriterium ist die Vollständigkeit, welche bei einer einseitigen Argumentation mit folgenden Punkten erfüllt ist: These, Argument/Begründung und Beleg. Da die Aufgabenstellung nicht explizit eine ein- oder zweiseitige, sondern lediglich eine begründete Stellungnahme forderte, reichte eine einseitige Begründung mit den entsprechenden Punkten zur Erfüllung der

Vollständigkeitskomponente aus. Wählte ein Proband in der Antwort jedoch eine zweiseitige Argumentation, so musste ein Fazit, Zusammenfassung oder Schlusssatz hinzugefügt sein, damit die Positionierung des Probanden im Sinne der Aufgabenerfüllung (begründete Stellungnahme) auch klar transportiert wurde. Das zweite Bewertungskriterium resultierte induktiv aus den Daten, ohne vorher eine Theorie oder mögliche Einflussfaktoren zu Grunde gelegt zu haben. Es ging aus den Daten hervor, dass sich die Probanden in zwei Gruppen aufteilten, wobei das Unterscheidungskriterium der beiden Gruppen mit „Gesundheitsbezogenheit“ treffend beschrieben werden kann. Der inhaltliche Bezug zur Aufgabenstellung – in diesem Fall der Gesundheitsbezug – war bei einigen Probanden nicht gegeben. Sie argumentierten zu allgemein (z. B. mit der Nutzungshäufigkeit) ohne eine logische Verbindung zur Gesundheit zu ziehen (Welche Konsequenzen ergeben sich aus der erhöhten Nutzung eines E-Bikes?). Andere schafften die Verbindung und stützten ihr Argument mit gesundheitsrelevanten Belegen (z. B. die Konsequenz aus der Nutzungshäufigkeit). Folgende, in Tabelle 9 beispielhaft dargestellte Textausschnitte aus den Probandenantworten verdeutlichen den Unterschied:

Tabelle 9: *Beispielhafte Ausschnitte aus den Stellungnahmen der Probanden.*

<u>Einteilung</u>	<u>Beispielsätze</u>	<u>Erklärung für Einteilung</u>
Argumentation ohne direkten Gesundheitsbezug	Zwar ist das Fahren mit einem E-Bike deutlich weniger anstrengend als mit einem normalen Fahrrad, trotzdem fördert es die Gesundheit. Denn Studien haben nachgewiesen, dass E-Bike-Fahrer ihr Rad deutlich häufiger nutzen, als Nutzer eines einfachen Fahrrads (Quelle: https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/ ; https://www.egms.de/static/de/meetings/dkou2015/15dkou230.shtml). Viele E-Bike Fahrer würden, hätten sie kein E-Bike, wohl eher gar nicht Rad fahren. Außerdem trauen sich Menschen mit dem E-Bike eher, auch Berge hoch zu fahren. Dies fördert die Gesundheit besonders (201, Pos. 2) Durch die Verwendung von E-Bikes fahren mehr Menschen Fahrrad. Sie trauen sich zudem längere Strecken und mehr Höhenmeter zu, als sie es mit einem Fahrrad ohne Motor tun würden. Insbesondere für ältere Menschen ist das E-Bike eine Alternative zum Auto, die das Fahrrad ohne Motor nicht wäre. Verschiedene Studien haben dies belegt (33, Pos. 2)	Der Proband argumentiert mit der häufigeren Nutzung und erwähnt als Konsequenz hieraus das Fördern der Gesundheit („denn), jedoch zieht er keinen konkreteren Rückschluss aus der Nutzungshäufigkeit und den daraus folgenden Konsequenzen für die Gesundheit (z. B., dass es zu mehr Bewegung führt und diese wiederum gut für das Herz, den Stoffwechsel, die Muskulatur etc. hat) Der Proband argumentiert mit der Nutzungshäufigkeit, zieht aber keine Rückschlüsse daraus, ob, warum oder inwieweit eine Nutzungshäufigkeit eine Auswirkung auf die Gesundheit hat.
Argumentation mit Gesundheitsbezug	Meiner Meinung nach können E-Bikes zur Gesundheitsförderung beitragen. Durch den integrierten Motor bieten sie eine Unterstützung beim Fahren um statt einer hohen körperlichen Belastung eine stetige moderate zu ermöglichen, ohne darauf zu verzichten eine beliebige Strecke auch mit Steigungen zu wählen. Älteren und weniger körperlich fitten Menschen ermöglichen sie sportlich aktiv zu sein. Durch eine moderate Bewegung und mittlere Belastung bei etwa 75% laut Studienangaben in dem Artikel von Quarks und aus einer sportzeitschrift (https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/; https://www.zeitschrift-sportmedizin.de/e-bikes-und-der-einfluss-auf-das-gesundheitsverhalten-von-risikogruppen/) wird das Herzkreislaufsystem gestärkt. Auch die Lungenfunktion wird gestärkt und es ist gelenkschonend (120, Pos. 2) Insgesamt denke ich, dass E-Bikes zur Gesundheitsförderung beitragen. Zwar ist die physische Aktivität beim Benutzen eines E-Bikes geringer als bei einem konventionellen Fahrrad, dennoch ist sie im Vergleich höher als beim Laufen. [...] Weiterhin ist der sportliche Effekt mit dem E-Bike geringer, allerdings wird es auch häufiger benutzt und trainiert entsprechend die Muskulatur. Darüber hinaus unterstützt das regelmäßige Fahren mit dem E-Bike den Körperfettabbau und reduziert das Stresslevel. (118, Pos. 2)	Der Proband argumentiert damit, dass die Belastung zwar moderat ist, die Bewegung insgesamt aber eine positive Auswirkung auf das Herzkreislaufsystem und die Lungenfunktion hat. Der Proband argumentiert zunächst mit der Nutzungshäufigkeit und leitet aus der Nutzungshäufigkeit gezielte positiven Konsequenzen für die Gesundheit (Muskeltraining, Reduktion Körperfett und Stress) ab.

Anmerkung.

Die Nummerierung inklusive Positionsangabe hinter den jeweiligen Probandenangaben bezieht sich auf den in MaxQDA Plus (Release 22.3.0 vom 18.10.2022) hinterlegten Datensatz, in welchem die Probandenantworten analysiert wurden.

Anand dieser beiden Kriterien wurden die Probandenantworten analysiert und in die Kategorien „gute“ oder „schlechte“ Argumentationsstruktur (i. S. v. vollständig und inhaltsbezogen)

eingeteilt. Alle diejenigen, welche vollständig (egal ob ein- oder zweiseitig) und gesundheitsbezogen argumentieren, erhielten die Einteilung 1 – gute Argumentationsstruktur. Alle diejenigen, bei denen entweder der Aspekt der Vollständigkeit oder des Gesundheitsbezuges oder sogar beides fehlten, erhielten die Einteilung in die schlechte Argumentationsstruktur 0. Weiterhin fielen in die gleiche Kategorie Probanden, welche die Aufgabe nicht beantworteten oder verfehlten.

4.3. Datenanalyse

Für die finale Analyse der Modelle wurden die zuvor analysierten und bereits operationalisierten Daten in die entsprechenden Modelle integriert. Für jedes der nachstehenden Modelle wurden die Korrelationen und Regressionen berechnet, die für die korrespondierenden Skalenniveaus (metrisch oder dichotom) notwendig waren. Da in der Arbeit eine Beobachtungsstudie zu Grunde liegt, können neben den zu untersuchenden Variablen auch damit verbundene Variablen (Störvariablen) auftreten. Diese gilt es zu kontrollieren, damit zumindest erste Hinweise auf die Kausaleffekte der interessierenden Variablen geliefert werden könnten (Döring & Bortz, 2016).

Gemäß diesen Anforderungen werden also Störvariablen, die einen möglichen Einfluss auf die unabhängigen Variablen haben könnten, als Kontrollvariablen in das Analysedesign mit aufgenommen; diese sind das Alter als metrische Variable sowie das Geschlecht als trichotome Variable (divers, männlich, weiblich) mit einer Referenzkategorie. Zudem wurde die Hochschulzugangsberechtigungsnote (HZB-Note) als Indikator für allgemeine Vorbildung aufgenommen und als Kontrollvariable eingeführt. Es können sich weiterhin auch Unterschiede im Studiengang zeigen. Dies resultiert insbesondere daher, dass die Studienordnungen und Curricula von Jura, Wirtschaftspädagogik und Medizin einen unterschiedlichen Aufbau, Inhalt sowie Fokus aufweisen. Es ist belegt, dass das kritische Denken über verschiedene Studiengänge hinweg kaum als kohärentes Konzept gelehrt wird (Valavičienė et al., 2020), weshalb es zu Unterschieden in der Fähigkeit des kritischen Denkens kommen könnte. Außerdem zeigen Forschende, dass sich die Mediennutzung zwischen verschiedenen Studiengängen unterscheidet (le Roux & Parry, 2017), weshalb sich auch für vorliegende Studiengänge bzw. Domänen Unterschiede ergeben könnten. Es ist daher anzunehmen, dass sich im Laufe der Studiendauer die Recherche- und Beurteilungsfähigkeiten sowie Mediennutzung der Berufsanfänger je nach Domäne voneinander differenzieren lassen, was die Betrachtung der Domäne als Kontrollvariable rechtfertigt. Diese wurden analog zum Geschlecht als trichotome Variable mit einer Referenzkategorie für die jeweiligen Studiengänge ins Erhebungsdesign inkludiert.

Modelle und ihre Analysen in der Übersicht. In Summe werden 15 verschiedene Modelle berechnet, welche sich grundsätzlich in die Modelle der Suchphase (Modelle I bis III) und die der Evaluationsphase (IV bis VI) unterteilen lassen (Tabelle 10). In beiden Phasen sind je drei Variablen als abhängige Variablen festgesetzt. In den Modellen I, IV und V können NfC und IO gemeinsam in einem Modell mittels multipler logistischer bzw. linearer Regression berechnet werden. Die Modelle II und III beinhalten Interaktionseffekte (Moderation) sowie die Einbindung der Webseitenmerkmale als weitere Prädiktorvariablen. Modell VI untersucht zusätzlich die Webseitenmerkmale als Mediatorvariable.

Tabelle 10: *Übersicht aller Modelle.*

Phase	Modell	AV	UV Personenmerkmale	UV Webseitenmerkmale	Analyse
Suchphase	I	Spezifität der Suchbegriffseingabe	I-A: NfC, IO I-B: NfCC	/	(Multiple) Logistische Regression
	II	Nachrecherche	II-A: NfC (+ Moderator) II-B: IO (+ Moderator) II-C: NfCC (+ Moderator)	Textglaubwürdigkeit besuchter Webseiten (linear sowie non-linear) + Medienglaubwürdigkeit besuchter Webseiten (linear sowie non-linear) + Autorglaubwürdigkeit besuchter Webseiten (linear sowie non-linear)	Logistische Regression mit Interaktionstermen
	III	Weiterverlinkung anklicken	III-A: NfC (+ Moderator) III-B: IO (+ Moderator) III-C: NfCC (+ Moderator)	Textglaubwürdigkeit besuchter Webseiten (linear sowie non-linear) + Medienglaubwürdigkeit besuchter Webseiten (linear sowie non-linear) + Autorglaubwürdigkeit besuchter Webseiten (linear sowie non-linear)	
Evaluationsphase	IV	Korrekte Quelleneinschätzung	IV-A: NfC, IO IV-B: NfCC	/	(Multiple) Lineare Regression
	V	Systematische Informationsverarbeitung	V-A: NfC, IO V-B: NfCC	/	(Multiple) Logistische Regression
	VI	Argumentationsstruktur	VI-A: NfC (+ Moderator) VI-B: IO (+ Moderator) VI-C: NfCC (+ Moderator)	Textglaubwürdigkeit selektierter Webseiten (als Mediator)	Logistische Regression mit moderierter Mediation via Process
I bis VI		Soziodemografie als Kontrollvariablen			

5. Ergebnisse

Die folgenden Ergebniskapitel berichten zum einen die deskriptiven Statistiken (Dispersionsmaße und Häufigkeiten) der unabhängigen Variablen. Es werden an dieser Stelle auch die Variablenlängen dargestellt, da diese u. a. mithilfe der deskriptiven Statistiken beurteilt werden. Die kombinierte Darstellung aus Beurteilung der Variablenlänge sowie deskriptive Statistik sollen den Lesefluss vereinfachen und Wiederholungen verhindern. Weiterhin werden die abhängigen Variablen deskriptiv untersucht und hier dargestellt. Diese geben einen Einblick in die Verteilung der Variablen sowie die Häufigkeit des untersuchten Verhaltens, was für die daran anschließende Diskussion der Ergebnisse notwendig ist. Korrelative Ergebnisse werden zwecks Vollständigkeit und Nachvollziehbarkeit nur kurz präsentiert, da sie für die Beantwortung der Forschungsfrage bzw. Überprüfung der Hypothesen, welche als gerichtete Hypothesen formuliert sind, nicht notwendig sind. Zuletzt werden die inferenzstatistischen Ergebnisse dargestellt, welche zur Beantwortung der Hypothesen herangezogen werden.

5.1. Deskriptive Ergebnisse

5.1.1. Beurteilung der Variablenlänge der Personen- und Webseitenmerkmale

Personenmerkmale. Für alle unabhängigen Variablen, welche mittels einer Skala durch Selbstauskunft erhoben wurden, werden die Reliabilität, die Streuung und die Normalverteilung zur Aussage der Variablenlänge herangezogen. Dies wird im Anhang A 2 dargestellt. Die Reliabilität misst die Verzerrung eines Testes durch einen Messfehler (Döring & Bortz, 2016). Für die Überprüfung der Reliabilität wird sich der gängigsten Methode, die der internen Konsistenz mithilfe von Cronbach's α , bedient (Döring & Bortz, 2016, S. 443). Für die Beurteilung, ob die Variable in der Stichprobe gut gestreut hat, wird die gefundene Streuung der gemittelten Variablen, welche in der finalen Analyse eingeflossen ist, mit der einer als vergleichbar gut geltenden Streuung bei einem 6-Einheitsintervall und einer Standardabweichung von 1 abgeglichen. Der Shapiro-Wilk-Test zur Überprüfung der Normalverteilung wird für jede UV durchgeführt.

Die Personenmerkmale lassen sich in der Stichprobe, wie in Tabelle 11 dargestellt, beschreiben. Es ist anzumerken, dass NfCC eine geringere Stichprobe aufweist, da dieses Merkmal in der Post-Erhebung erhoben wurde.

Tabelle 11: *Deskriptive Statistiken der Personenmerkmale.*

	<i>N</i>	<i>Min / Max</i>	<i>M (SD)</i>
NfC	144	2 / 6	4.154 (0.833)
NfCC	64	1.688 / 3.125	2.410 (0.319)
IO	144	1 / 6	3.322 (1.257)

Need for Cognition. Der Reliabilitätstest mittels Cronbach's α ⁴⁷ bestätigt eine akzeptable interne Konsistenz mit Cronbach's $\alpha = .672$ in den vorliegenden Daten (Anhang A 2). Obwohl sich eine Verbesserung der Reliabilität durch Löschen zweier Items hätte erzielen lassen können, wurde sich aufgrund der ohnehin schon geringen Anzahl an Items dagegen entschieden. Der Shapiro-Wilk-Test kann eine Normalverteilung bestätigen, da $p > .05$, was zudem durch eine Inspektion des Histogramms und des Q-Q-Plots bestätigt werden kann. Die Standardabweichung SD beläuft sich auf 0.833 mit $Min = 2$ und $Max = 6$. Die für 4-Einheitsintervalle als gut geltende Standardabweichung liegt bei 0.66⁴⁸. Für NfC liegt aufgrund des $Min = 2$ und $Max = 6$ ebenfalls ein 4-Einheitsintervall vor. NfC hat somit in der Stichprobe gut gestreut (da $0.833 > 0.66$). Es besteht in der Stichprobe zudem eine leichte Tendenz, dass vermehrt Menschen mit hohem NfC enthalten sind, was einerseits anhand des Mittelwerts von 4.154, weiterhin auch durch das hier vorliegende Minimum von 2 (statt laut Skala 1) zu erkennen ist.

Need for Cognitive Closure. Der Reliabilitätstest mittels Cronbach's α bestätigt eine gute interne Konsistenz mit Cronbach's $\alpha = .752$ (Anhang A 2). Obwohl sich eine Verbesserung der Reliabilität durch Löschen zweier Items hätte erzielen lassen können, wurde sich dagegen entschieden. Die Verbesserung wäre maximal um lediglich 0.007 Punkte wirksam geworden, was keine tatsächliche Verbesserung darstellt. Die Untersuchung auf Normalverteilung mit dem Shapiro-Wilk-Test ergibt $p > .05$, weshalb von Normalverteilung ausgegangen werden kann. Die Standardabweichung SD beläuft sich auf 0.319 mit $Min = 1.688$ und $Max = 3.125$. Die für dieses Intervall als gut geltende Standardabweichung liegt bei 0.240. NfCC liegt somit in der Stichprobe gut gestreut vor (da $0.319 > 0.240$). Der Mittelwert liegt mit einem Wert von 2.410 bei einer Skala von 1 bis 4 ebenfalls tendenziell zentriert.

Information Overload. Der Reliabilitätstest mittels Cronbach's α bestätigt eine sehr gute interne Konsistenz mit Cronbach's $\alpha = .855$ (Anhang A 2). Obwohl sich eine Verbesserung der Reliabilität durch Löschen eines Items hätte erzielen lassen können, wurde sich aufgrund der bereits ohnehin schon geringen Anzahl an Items (4) dagegen entschieden. Der Test auf Normalverteilung zeigt ein signifikantes Ergebnis ($p = .009$), was folglich gegen eine Normalverteilung spricht. Die Nicht-Normalverteilung kann sich in der späteren Modellberechnung auch in einer Nicht-Normalverteilung der Residuen niederschlagen, was weitere Probleme mit entsprechenden Prämissen mit sich ziehen kann. An der entsprechenden Stelle wird nochmals darauf hingewiesen und zudem in der Diskussion kritisch beurteilt. Die Standardabweichung SD beläuft

⁴⁷ Zur Interpretation von Cronbach's α sind folgende Angaben dienlich: größer als .9 = exzellent, größer als .8 = gut, größer als .7 = akzeptabel, größer als .6 = fragwürdig, größer als .5 = schlecht und kleiner als .5 = inakzeptabel (Blaž, 2015; Cronbach, 1951).

⁴⁸ Die Zahlen ergeben sich durch proportionale Umrechnung der Größen „Einheitsintervall = 6“ sowie „Standardabweichung = 1“. Um auf ein „Einheitsintervall = 4“ zu kommen, muss durch 1.5 dividiert werden, was eine Standardabweichung = 0.66 ergibt.

sich auf 1.257 mit $Min = 1$ und $Max = 6$. Die für 5-Einheitsintervalle als gut geltende Standardabweichung liegt bei 0.833. IO liegt somit in der Stichprobe gut gestreut vor (da $1.257 > 0.833$). Der Mittelwert liegt bei 3.322, was bei einer Skala von 1 bis 6 fast die Mitte repräsentiert.

Hinsichtlich der soziodemografischen Daten (Tabelle 12) ist die Stichprobe bezüglich einzelner Merkmale heterogener im Vergleich zu der Grundgesamtheit in Deutschland (*Bevölkerung nach Nationalität und Geschlecht (Quartalszahlen)*, o. J.). So befinden sich hauptsächlich Frauen in der Stichprobe ($n = 95$). Auch sind angehende Mediziner mit 68 am meisten vertreten, gefolgt von 46 Studierenden aus der Domäne Lehramt und zuletzt 30 aus Jura. Das Durchschnittsalter der Stichprobe beläuft sich auf 27.146 Jahren mit einem Minimum von 23 und Maximum von 46. In Anbetracht dessen, dass sich die Stichprobe aus Berufsanfängern im ersten praktischen Jahr zusammensetzt und somit automatisch einen Masterabschluss voraussetzt, sind diese Altersdaten zu erwarten. Auch ist die durchschnittliche Hochschulzugangsberechtigung mit 1.76 angemessen und erwartbar, da zumindest für Jura und Medizin ein Numerus Clausus gilt.

Tabelle 12: *Deskriptive Statistiken der Soziodemografie.*

	<i>N</i>	<i>Min / Max</i>	<i>M (SD)</i>
HZB-Note	144	1 / 3.6	1.760 (0.641)
Alter	144	23 / 46	27.146 (3.500)
Geschlecht	Männlich	Weiblich	Divers
Häufigkeit	40	95	9
Studiengang	Lehramt	Medizin	Jura
Häufigkeit	46	68	30

Webseitenmerkmale. Auch die blockweise zusammengefassten Webseitenmerkmale werden nachstehend auf zumindest Normalverteilung und Streuung geprüft, da diese sowohl als unabhängige Variablen als auch als Mediator- bzw. Moderatorvariablen eine gewisse Güte aufweisen sollten. Die deskriptiven Statistiken der blockweise zusammengefassten Webseitenmerkmale sind in Tabelle 13 dargestellt. Die deskriptiven Statistiken der separaten Webseitenmerkmale befinden sich, wie bereits erwähnt, im Anhang A 1. Denn sie stellen nur ein Zwischenprodukt dar und werden deshalb hier nicht ausführlich thematisiert.

Tabelle 13: *Deskriptive Statistiken der Merkmale der besuchten und selektierten Webseiten.*

	<i>n</i>	<i>Min / Max</i>	<i>M (SD)</i>
Textglaubwürdigkeit <i>aufge- rufener</i> Webseiten	142	0.16 / 0.73	0.535 (0.117)
Medienglaubwürdigkeit <i>auf- gerufener</i> Webseiten	142	0.00 / 0.75	0.481 (0.156)
Autorenglaubwürdigkeit <i>auf- gerufener</i> Webseiten	142	0.00 / 1.00	0.699 (0.242)
Textglaubwürdigkeit <i>selek- tierter</i> Webseiten	143	0.06 / 0.72	0.515 (0.123)

Die Untersuchung auf Normalverteilung mit dem Shapiro-Wilk-Test für alle Glaubwürdigkeitsindices der sowohl besuchten als auch selektierten Webseiten zeigt $p < .05$, weshalb nicht von Normalverteilung ausgegangen werden kann.

Anhand des Mittelwertes zeigt sich, dass ungefähr die Hälfte der Probanden Webseiten *besuchen*, die subjektiv wahrnehmbare textliche sowie mediumsbezogene Glaubwürdigkeitsmerkmale bereitstellen und deshalb auch als glaubwürdiger wahrgenommen werden könnten. Das bedeutet, dass circa 50 % der Probanden kumulativ⁴⁹ betrachtet Webseiten besuchen, die z. B. verständlich oder aktuell sind oder eine hohe Nachweisqualität haben. Die visuelle Begutachtung des Plots der beiden Variablen zeigt, dass sich sehr viele Probanden auf die jeweiligen Extrempunkte verteilen, was sich bereits durch die Nicht-Normalverteilung andeutete. Es ist daher von einer eingeschränkten Streuung auszugehen. Bezüglich subjektiv wahrnehmbarer autorenbezogener Glaubwürdigkeitsmerkmale *besuchen* circa 70 % der Probanden solche Webseiten. D. h., 70 % der Probanden besuchen Webseiten, die die Autoren angegeben haben und deshalb als glaubwürdiger wahrgenommen werden könnten. Weiterhin wird ersichtlich, dass etwa die Hälfte der Probanden Webseiten *selektieren*, die subjektiv wahrnehmbare textliche Merkmale bereitstellen und deshalb als glaubwürdiger wahrgenommen werden könnten. Die visuelle Inspektion der Häufigkeitsverteilung zeigt, dass beide Variablen zwar gut verteilt sind, es jedoch viele Probanden am äußeren, rechten Extrempunkt gibt. Es existieren folglich viele Probanden, die Webseiten *besuchen* bzw. *selektieren*, die bezüglich der Text- und Autorenmerkmale als sehr glaubwürdig bewertet werden können. Diese Problematiken der Nicht-Normalverteilung und damit einhergehend schlechten Häufigkeitsverteilung werden in der Diskussion aufgegriffen.

⁴⁹ Kumulativ deshalb, weil ein relativer Einfluss der einzelnen Merkmale (z. B. Verständlichkeit) auf die blockweise zusammengefassten Variablen (z. B. PIXbText) nicht gemessen wird und sie lediglich als Äquivalente in die Berechnung der blockweisen Variablen aufgenommen wurden.

5.1.2. Beschreibung der Häufigkeiten der abhängigen Variablen

Die deskriptiven Statistiken bezüglich der abhängigen Variablen der Suchstrategie (Tabelle 14) zeigen, dass genau die Hälfte der Probanden mit einem Mittelwert von 0.5 spezifische und oberflächliche Suchbegriffe eingeben. Anders sieht es bei den zwei weiteren abhängigen Variablen aus: Während noch 32.6 % (44 von 135) Probanden eine Nachrecherche betreiben, klicken lediglich ca. 26 % (35 von 144) Probanden die Weiterverlinkungen an.

Tabelle 14: *Häufigkeitsverteilung der dichotomen abhängigen Variablen.*

	<i>n</i>	<i>M (SD)</i>	<i>0</i>	<i>1</i>
Spezifität der Suchbegriffseingabe	134	0.5 (0.502)	67	67
Nachrecherche	135	0.326 (0.470)	91	44
Weiterverlinkung	135	0.259 (0.440)	100	35
Systematische Begründung	135	0.69 (0.465)	42	93
Argumentationsstruktur	143	0.615 (0.488)	55	88

Hinsichtlich der Weiterverlinkung ist ein Nebenfund zu berichten, welches in Tabelle 15 dargestellt ist und für die anschließende Diskussion notwendig wird. Die Tabelle zeigt in den Spalten 1 - 4, dass auf den am meisten aufgerufenen, unique Webseiten auch anklickbare Links zur Verfügung stehen⁵⁰. Die 35 Probanden, welche die Links weiterverfolgen, haben dies überwiegend von den meist aufgerufenen Webseiten in Summe getan haben. Beispielsweise rufen 74 Probanden mindestens einmal *Quarks* auf. Lediglich 11 von den 74 Probanden (ca. 15 %) davon verfolgen die Weiterverlinkung von *Quarks*. Wie die Spalten 1 - 4 verdeutlichen sollen, ist die Möglichkeit zum Anklicken der Weiterverlinkung von Seiten der Webseiten gegeben und es kann die Verteilung dieser AV zumindest nicht auf nicht-vorhandene Möglichkeit zur Linkverfolgung zurück zu führen sein. Spalte 5 und 6 geben zusätzliche Informationen über die Medien- bzw. Glaubwürdigkeit der Top aufgerufenen Webseiten. Beide Informationen werden an entsprechender Diskussionsstelle nochmals aufgegriffen.

⁵⁰ In gesamter Betrachtung stellen 100 von 147 aufgerufene Webseiten anklickbare Links zur Verfügung.

Tabelle 15: *Aufgerufene Webseitenverteilung und Weiterverlinkung.*

Aufgerufene Webseiten	Probandenanzahl	Link verfügbar?	Weiterverlinkung angeklickt	Mediumsbezogene GW	Textliche GW
<i>Quarks</i> Artikel	74	ja	Ja, von 11	Hoch	Hoch
<i>Zeitschrift Sportmedizin</i> Artikel	45	ja	Ja, von 6	Niedrig	Mittel
<i>Emotion Technologies</i> Artikel	44	Nein	-	Niedrig	Niedrig
<i>Ärzteblatt</i> Artikel 1	33	ja	Ja, von 6	Hoch	Mittel
<i>Envivas</i> Artikel	28	ja	Ja, von 1	Niedrig	Hoch
<i>Zeit</i> Artikel	19	ja		Hoch	Hoch
<i>Focus</i> Arzt-suche Artikel	13	ja		Hoch	Hoch
<i>Ebike</i> Gesundheit Artikel	12	ja	Ja, von 1	Niedrig	Niedrig
<i>Elektrofahrrad24</i> Artikel	12	ja		Niedrig	Niedrig
<i>Stromerbike</i> Artikel	12	ja	Ja, von 3	Niedrig	Hoch
<i>IJBPA</i> Artikel 1	11	ja	Ja, von 2	Hoch	Hoch
<i>Uni Basel</i> Artikel	11	ja	Ja, von 3	Mittel	Hoch

Anmerkung.

Legende: hoch = im oberen Dittel im Vergleich zu den restlichen bewerteten aufgerufenen Webseiten, mittel = im mittleren Dittel im Vergleich zu den restlichen bewerteten aufgerufenen Webseiten, niedrig = im unteren Dittel im Vergleich zu den restlichen bewerteten aufgerufenen Webseiten

Die abhängigen Variablen der Evaluationsstrategien befinden sich, sofern dichotom, ebenfalls in Tabelle 14 und, sofern metrisch, in Tabelle 16. Sowohl der Mittelwert als auch die Verteilung der Ausprägung zeigen, dass der überwiegende Teil, sprich 62 % der Probanden, eine gute Argumentationsstruktur aufweisen. Hinsichtlich der Kategorie der Quelleneinschätzung wird deutlich, dass 93 Probanden mindestens einmal eine Begründung aus der systematischen Kategorie angeben. Anders ausgedrückt geben circa 70 % der Probanden mindestens einmal eine Begründung für die Glaubwürdigkeit einer Quelle an, die auf eine systematische Informationsverarbeitung schließen lässt. Die Qualität der Quellen wird im Durchschnitt mit 2.95 korrekt eingeschätzt, was bei einer Skala von 1 bis 4 somit eine gute Performanz in dem Bereich darstellt. Das bedeutet, dass Probanden eher überdurchschnittlich glaubwürdige Quellen als korrekt glaubwürdig und unglaubwürdige Quellen als korrekt unglaubwürdig einschätzen.

Tabelle 16: *Deskriptive Statistiken der metrischen abhängigen Variablen.*

	<i>n</i>	<i>Min / Max</i>	<i>M (SD)</i>
Korrektheit der Quelleneinschätzung	143	0 / 4	2.947 (0.798)

5.2. Korrelative Ergebnisse

Nachstehend werden die korrelativen Ergebnisse zwecks Vollständigkeit kurz berichtet. Sie können erste Hinweise für Zusammenhänge und die spätere Bestätigung der Hypothesen liefern. Es sei anzumerken, dass lediglich diejenigen bivariaten Zusammenhänge betrachtet werden, welche später auch zusammen in den Modellen integriert werden.

Generelles Vorgehen. Da es sich bei den Variablen um verschiedene Typen (kategorial vs. metrisch) handelt, werden die Korrelationen gemäß der Kombination der unterschiedlichen Datentypen berechnet. Bei metrischen Variablen, wie dem Alter oder NfC, wird die Korrelation mithilfe von Pearson berechnet. Hierdurch können die Stärke und die Richtung der signifikanten Zusammenhänge direkt abgelesen werden. Bei einer Kombination aus dichotomen und metrischen Variablen (z. B. Alter und Geschlecht) wird mithilfe einer Kreuztabelle zunächst der Eta-Koeffizient (η) ermittelt (Benninghaus, 2007, S. 228 ff.). Dieser gibt jedoch nur die Stärke des Zusammenhangs an⁵¹. Um die tatsächliche Signifikanz des Zusammenhangs zu prüfen, wird anschließend ein *t*-Test für unabhängige Stichproben⁵² gerechnet. Bei signifikanten Ergebnissen werden dann die entsprechenden Mittelwerte zu Rate gezogen, um die Richtung des Zusammenhangs festzustellen. Bei rein dichotomen Variablenpaaren wird ebenfalls die Kreuztabelle verwendet, welche allerdings mit Cramers V und dem χ^2 (χ^2) kombiniert wird. Cramers V ist ein Maß für die Assoziationsstärke eines Zusammenhangs (Backhaus et al., 2021, S. 392)⁵³. Mithilfe des χ^2 -Tests wird auch hier die Signifikanz separat berechnet. Letztlich kann durch Anschauen der tatsächlichen und erwarteten Häufigkeiten die Richtung der signifikanten Effekte erkannt werden.

Es ist noch anzumerken, dass sowohl das Geschlecht als auch die Domäne als trichotome Variablen im Datensatz enthalten sind. Es wurden Dummyvariablen für die jeweiligen drei Ausprägungen gebildet, da so zielgerichtetere Aussagen zu den einzelnen Domänen bzw. Geschlechtern möglich sind. Aus Folge daraus muss beachtet werden, dass korrelative

⁵¹ η^2 nimmt Werte zwischen 0 und 1 an, wobei höhere Werte für einen stärkeren Zusammenhang zwischen den Variablen sprechen (Benninghaus, 2007).

⁵² Eine detaillierte Beschreibung zu den theoretischen Hintergründen und Vorgehen beim *t*-Test ist in Kapitel 5.3.2 erst dargestellt. Grund hierfür ist, dass das Korrelationskapitel mit den entsprechenden Berechnungsmethoden hier nur die notwendigen Informationen zum Verständnis der Zusammenhänge enthält und deshalb keine theoretischen Hintergründe und Vorgehensbeschreibungen zu den Methoden geliefert werden.

⁵³ Einteilung: V hat eine kleine Effektstärke mit .1, eine mittlere Effektstärke mit .3 und eine große Effektstärke mit .5 (J. Cohen, 1988).

Ergebnisse zum Geschlecht bzw. zur Domäne stets im Vergleich zu den beiden anderen, zusammengenommenen Gruppen berichtet werden. Tabelle 17 zeigt die korrelativen Ergebnisse. Die farblich hinterlegten Felder geben Aufschluss über die entsprechende Methode zur Überprüfung der Zusammenhänge. Die Korrelationen werden nachstehend in thematisch passender Reihenfolge berichtet, d. h. beginnend mit den soziodemografischen Merkmalen und den unabhängigen Variablen, anschließend die abhängigen Variablen etc.

Soziodemografische Merkmale, Personenmerkmale sowie Webseitenmerkmale. Zwischen Geschlecht und den Personenmerkmalen ist lediglich eine leicht signifikante und tendenziell schwache Korrelation zu finden ($t(142) = -1.989$, $p = .049$; $\eta = .165$): Männer ($M = 4.375$; $n = 40$) haben im Vergleich zu den restlichen Probanden ($M = 4.07$; $n = 102$) ein höheres NfC.

Hinsichtlich des Geschlechts und der Webseitenmerkmale ist zu erkennen, dass Männer ($M = 0.532$ für Medienglaubwürdigkeit, $M = 0.763$ für Autorenglaubwürdigkeit; $n = 40$) im Vergleich zu Frauen bzw. Personen diversen Geschlechts ($M = 0.461$ für Medienglaubwürdigkeit, $M = 0.673$ für Autorenglaubwürdigkeit; $n = 102$) Webseiten besuchen, die sowohl eine höhere Medienglaubwürdigkeit ($t(140) = -2.511$, $p = .013$; $\eta = .208$) als auch eine höhere Autorenglaubwürdigkeit aufweisen ($t(140) = -1.992$, $p = .048$; $\eta = .166$). Beide Effekte sind leicht signifikant, jedoch eher schwach.

Innerhalb der Domänen zeigen sich ebenfalls Zusammenhänge mit den Webseitenmerkmalen. Angehende Lehrer ($M = 0.491$ für Textglaubwürdigkeit; $M = 0.407$ für Medienglaubwürdigkeit; $n = 44$) besuchen im Vergleich zu angehenden Juristen und Medizinern ($M = 0.5548$ für Textglaubwürdigkeit; $M = 0.5139$ für Medienglaubwürdigkeit; $n = 98$) Webseiten mit sowohl geringer Textglaubwürdigkeit ($t(140) = 3.093$, $p = .002$; $\eta = .253$) als auch geringer Medienglaubwürdigkeit ($t(140) = 3.968$, $p < .001$; $\eta = .318$). Beide Effekte sind signifikant und zeigen einen – vergleichsweise zu den bisher berichteten – stärkeren Zusammenhang. Daran anschließend ist ein weiterer Effekt statistisch signifikant und stark: Angehende Lehrer ($M = 0.464$; $n = 45$) selektieren im Vergleich zu Medizinern und Juristen ($M = 0.539$, $n = 98$) Webseiten, die eine geringe Textglaubwürdigkeit aufweisen ($t(141) = 3.495$, $p < .001$; $\eta = .282$).

Im Gegensatz dazu besuchen die angehenden Mediziner ($M = 0.575$ für Textglaubwürdigkeit; $M = 0.5322$ für Medienglaubwürdigkeit, $n = 68$) im Vergleich zu angehenden Lehrern und Juristen ($M = 0.4985$ für Textglaubwürdigkeit; $M = 0.434$ für Medienglaubwürdigkeit, $n = 74$) Webseiten mit einer höheren Textglaubwürdigkeit ($t(140) = -4.107$, $p < .001$; $\eta = .328$) und Medienglaubwürdigkeit ($t(140) = -3.969$, $p < .001$; $\eta = .318$). Weiterhin selektieren sie ($M = 0.561$; $n = 68$) Webseiten mit einer höheren Textglaubwürdigkeit ($t(141) = -4.514$, $p < .001$; $\eta = .355$) im Vergleich zu den anderen beiden Gruppen ($M = 0.4739$; $n = 74$). Alle Zusammenhänge sind statistisch signifikant.

Es gibt zudem einen leicht signifikant negativen Zusammenhang zwischen Alter und der Textglaubwürdigkeit, d. h. jüngere Probanden selektieren Webseiten, die eine höhere Textglaubwürdigkeit aufweisen ($r = -.196, p = .019$). Außerdem gibt es signifikant negative Zusammenhänge zwischen der HZB-Note und den Webseitenmerkmalen: Eine geringere HZB-Note, d. h. ein besseres Abitur, spricht für den Besuch von Webseiten mit einer höheren Textglaubwürdigkeit ($r = -.276, p < .001$) und einer höheren Medienglaubwürdigkeit ($r = -.275, p < .001$). Zudem selektieren bessere Abiturienten Webseiten, die eine höhere Textglaubwürdigkeit aufweisen ($r = -.320, p < .001$).

NfCC steht in einem leicht signifikant positiven Zusammenhang mit der Textglaubwürdigkeit besuchter Webseiten ($r = .256, p = .043$). Das bedeutet, dass Probanden mit hohem NfCC Webseiten besuchen, die eine höhere Textglaubwürdigkeit aufweisen im Vergleich zu Probanden mit niedrigem NfCC.

Letztlich ist für diesen Abschnitt noch zu erwähnen, dass die Merkmale der besuchten Webseiten auch miteinander stark korrelieren: Medienglaubwürdigkeit mit Textglaubwürdigkeit ($r = .644, p < .001$), Textglaubwürdigkeit mit Autorenglaubwürdigkeit ($r = .666, p < .001$) und Medienglaubwürdigkeit mit Autorenglaubwürdigkeit ($r = .399, p < .001$).

Soziodemografische Merkmale mit den abhängigen Variablen der Suchphase. Es zeigt sich ein signifikanter und fast mittlerer Effekt ($\chi^2(1) = 6.466, p = .011, V = .219$) dahingehend, dass weniger Mediziner (d. h. 13) im Vergleich zu angehenden Lehrern und Juristen nachrecherchieren, als es laut den Randhäufigkeiten eigentlich zu erwarten wäre (ca. 20). Umgekehrt zeigt sich ebenfalls als signifikanter und fast mittlerer Effekt, dass mehr Lehrer (d. h. 22) als eigentlich erwartet (d. h. ca. 14) im Vergleich zu angehenden Juristen und Mediziner nachrecherchieren.

Hinsichtlich der Weiterverlinkung sind ebenfalls Erkenntnisse zwischen den Domänen festzustellen. Weniger Lehrer (d. h. 6) klicken im Vergleich zu angehenden Juristen und Mediziner die Weiterverlinkung an, obwohl laut Randhäufigkeiten mehr (d. h. ca. 12) erwartet wären. Dieser Effekt ist jedoch statistisch nur leicht signifikant und von kleiner Effektstärke ($\chi^2(1) = 5.134, p = .023, V = .195$). Mehr angehende Mediziner (d. h. 23) als eigentlich erwartet (d. h. ca. 16) klicken im Vergleich zu den anderen Probanden die Weiterverlinkung an, wobei dieser Effekt ebenfalls statistisch signifikant, aber von mittlerer Größe ist ($\chi^2(1) = 8.040, p = .005, V = .244$).

Zuletzt ist zu berichten, dass die Probanden, die die Weiterverlinkung anklicken, eine HZB-Note von durchschnittlich 1.477 aufweisen. Probanden, die nicht die Links verfolgen, haben ein durchschnittlich schlechteres Abitur ($M = 1.866$). Das bedeutet, dass Probanden mit einem

besseren Abitur eher die Weiterverlinkung anklicken und dass dies ebenfalls statistisch signifikant ist ($t(133) = 3.155, p = .002; \eta = .264$).

Soziodemografische Merkmale mit den abhängigen Variablen der Evaluationsphase. Angehende Lehrer ($M = 2.652; n = 45$) schneiden im Vergleich zu angehenden Juristen und Mediziner ($M = 3.082; n = 98$) schlechter bei der korrekten Einschätzung der Quellenglaubwürdigkeit ab, d. h. sie schätzen häufiger eine glaubwürdige Quelle fälschlicherweise als unglaubwürdig bzw. eine unglaubwürdige Quelle fälschlicherweise als glaubwürdig ein ($t(141) = 3.086, p = .002; \eta = .252$).

Mediziner ($M = 3.094; n = 68$) hingegen sind im Vergleich zu den beiden anderen Gruppen ($M = 2.813; n = 75$) besser bei der korrekten Quelleneinschätzung. Das heißt sie können eher glaubwürdige Quellen als korrekt glaubwürdig und unglaubwürdige Quellen als korrekt unglaubwürdig identifizieren. Dieser Effekt ist statistisch leicht signifikant, zeigt aber eine geringere Effektstärke im Vergleich zu dem soeben berichteten Effekt auf ($t(141) = -2.126, p = .035; \eta = .176$).

Weiterhin gibt es einen statistisch signifikanten und mittleren Zusammenhang zwischen der Kategorie der Quelleneinschätzung und den angehenden Lehrern ($\chi^2(1) = 13.227, p < .001, V = .313$). Es geben weniger angehende Lehrer (d. h. 18) im Vergleich zu den angehenden Mediziner und Juristen mindestens eine Begründung aus der Semantikkategorie an, als es eigentlich erwartet wäre (d.h. ca. 27).

Letztlich ist zu berichten, dass es einen signifikant negativen Zusammenhang zwischen Alter und der Korrektheit der Quelleneinschätzung gibt ($r = -.392, p < .001$). Jüngere Probanden schneiden besser bei der Korrektheit der Quelleneinschätzung ab. Das bedeutet, dass eher jüngere Probanden im Vergleich zu älteren entweder unglaubwürdige Quellen als korrekt unglaubwürdig oder glaubwürdige Quellen als korrekt glaubwürdig identifizieren.

Webseitenmerkmale mit den abhängigen Variablen der Suchphase. Der Mittelwert für die Textglaubwürdigkeit von den Probanden, die die Links nicht weiterverfolgt haben ($n = 100$), liegt bei $M = 0.513$, wohingegen er bei den Probanden, die weiterverfolgt haben ($n = 35$), bei $M = 0.597$ liegt. Der Aufruf von Webseiten mit einer hohen Textglaubwürdigkeit könnte daher vermehrt zum Anklicken des Links führen ($t(133) = -3.813, p < .001; \eta = .314$).

Ähnlich sieht es bei der Medienglaubwürdigkeit aus: Der Mittelwert für die Medienlaubwürdigkeit von den Probanden, die die Links nicht weiterverfolgt haben ($n = 100$), liegt bei $M = 0.444$, wohingegen er bei den Probanden, die weiterverfolgt haben ($n = 35$), bei $M = 0.57$ liegt. Das heißt, dass der Aufruf von Webseiten mit einer hohen Medienglaubwürdigkeit daher ebenfalls eher zum Anklicken des Links führt ($t(133) = -4.476, p < .001; \eta = .362$).

Tabelle 17: Korrelationen.

	Personenmerk-			Webseitenmerkmale			Suchstrategie			Evaluationsstrategie				
	male	NfC	NfCC	IO	TextGW	Medien- GW	Auto- renGW	TextGWs	Spezifizi- tät	Nach- recherche	Weiterver- linkung	Kor- rektheit	Katego- rie	Argumen- tationsstruktur
männ- lich		.165*	.085	.002	.118	.208*	.166*	.104	.099	.092	.043	.038	.077	.065
weiblich		.109	.085	.025	.074	.151	.125	.043	.105	.056	.042	.008	.147	.077
divers		.091	-	.051	.073	.089	.063	.108	.030	.059	.158	.054	.141	.032
Lehramt		.051	.010	.022	.253**	.318**	.079	.282**	.048	.258**	.195*	.252**	.313**	.102
Medizin		.105	.114	.069	.328**	.318**	.162	.355**	.150	.219*	.244**	.176*	.145	.082
Jura		.038	.121	.059	.115	.029	.109	.114	.127	.030	.072	.071	.167	.016
Alter		.053	.044	.007	-.124	-.106	-.058	-.196*	.071	.086	.038	-.392**	.081	.072
HZB- Note		-.027	-.008	.080	-.276**	-.275**	-.144	-.320**	.106	.062	.264**	-.156	.117	.039
NfC				-.144	-.117	-.020	-.142	-.060	.018	.122	.038	-.085	.034	.120
NfCC					.256*	.210	.069	.193	.005	.179	.162	.007	.004	.077
IO					-.023	.019	-.081	-.098	.022	.110	.048	-.078	.085	.027
Text- GW						.644**	.666**			.114	.314**			
Medi- enGW							.399*			.038	.362**			
Auto- renGW										.026	.119			
Text- GWs														.095

Anmerkung.

** die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant

* die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant

GW = Glaubwürdigkeit

GWS = Glaubwürdigkeit selektierter Webseiten

Legende: Rot = ausschließlich metrische Variable, d. h. Berechnung durch Pearson; Gelb = Metrische und kategoriale Variablen, d. h. Berechnung mithilfe von Eta und t-Test; Grün = ausschließlich kategoriale Variablen, d. h. Berechnung mithilfe von Cramers V und Kreuztabelle

Grau = signifikantes Ergebnis

5.3. Ergebnisse der Inferenzstatistik

Für die Regressionsanalysen werden die entsprechenden Variablen zentriert (im Folgenden abgekürzt mit *zentr*) und anschließend standardisiert (im Folgenden abgekürzt mit *std*). Eine Zentrierung erfolgt durch Subtraktion des jeweiligen Mittelwertes, sodass eine bessere Interpretation ermöglicht werden kann (Eid et al., 2017). Die Standardisierung erfolgt durch die Division durch die jeweilige Standardabweichung, was die Interpretierbarkeit optimiert (Eid et al., 2017). Die dichotomen respektive trichotomen soziodemografischen Merkmale (Domäne und Geschlecht) werden als Dummyvariable gespeichert und jeweils zu zweit in die Modelle aufgenommen. Dadurch wird nur noch eine Referenzkategorie gebildet, im Gegensatz zu den zwei Referenzkategorien bei den zuvor berichteten korrelativen Ergebnissen. Für das Geschlecht ist die Referenzkategorie immer männlich, d. h. in den Modellen befinden sich die Dummyvariablen für Frauen und Divers, da davon ausgegangen wird, dass sich insbesondere Männer und Frauen sowie Männer und Personen diversen Geschlechts am ehesten unterscheiden. Für die Domäne ist die Referenzkategorie Medizin. Auch hier wird angenommen, dass sich die Domäne Lehramt und Medizin sowie Jura und Medizin am stärksten unterscheiden. Da es sich bei Modell II, III und VI um Moderationen bzw. moderierte Mediation handelt, werden Interaktionsterme aus der Multiplikation der jeweiligen Moderatorvariablen und unabhängigen Variablen gebildet (Eid et al., 2017).

Ehe die eigentlichen Ergebnisse vorgestellt werden, folgt anschließend zunächst die Beschreibung der Prämissenüberprüfung.

5.3.1. Prämissenüberprüfung

Es wurde sich dafür entschieden, die theoretische Beschreibung zur Überprüfung der modellunabhängigen, modellspezifischen Prämissen sowie die tatsächliche Überprüfung der modellunabhängigen Prämissen zusammen und vorgelagert zur Ergebnisdarstellung zu berichten. Grund hierfür ist, dass dadurch einerseits eine ausreichende Grundlage zum allgemeinen Vorgehen geschaffen wird. Andererseits wird der Lesefluss bei der Ergebnisdarstellung nur geringfügig durch die Überprüfung der modellspezifischen Prämissen unterbrochen.

Unabhängig davon, ob die Daten das vorhergesagte Modell auch in der Realität widerspiegeln, gelten für logistische als auch lineare Regressionen gewisse Prämissen. Diese sind übersichtsartig in nachstehende Tabelle 18 dargestellt. Einige Prämissen sind modellübergreifend. Sie werden auf Basis der Datensätze unabhängig der spezifischen Modelle geprüft. Diese werden zunächst nachfolgend für alle Modelle beschrieben und überprüft. Andere Prämissen hingegen sind modellspezifisch. Sie können nur im Rahmen mit den individuell für das Modell geltenden Daten und Variablen überprüft werden. Das Vorgehen zur Überprüfung dieser modellspezifischen Prämissen wird hier ebenfalls allgemein erläutert. Die tatsächliche

Überprüfung der modellspezifischen Prämissen erfolgt jedoch im Anschluss an jede Modellberechnung, was laut Literatur (Backhaus et al., 2021) dem grundsätzlichen Vorgehen bei einer Regressionsanalyse entspricht.

Tabelle 18: Prämissenübersicht der Modelle

Prämisse	Modell	Gilt für	Quelle
AV ist nominalskaliert mit 2 Ausprägungen	Übergreifend, d. h. im Basisdatensatz	Logistische Regression	Eid et al. (2017)
UVs sind alle oder teilweise metrischer Natur ⁵⁴			
Unabhängigkeit der Beobachtungen		Grundsätzlich immer	z. B. für Logreg: Backhaus et al. (2021, S. 379)
Fehlerfreie Messung der UVs		Lineare Regression	Backhaus et al. (2021, S. 101)
Gewisse Fallanzahl pro Prädiktor		Grundsätzlich immer	z. B. für Logreg: Backhaus et al. (2021, S. 379)
Keine Autokorrelation		Logistische Regression	Logreg: Stoltzfuß (2011, S. 1101)
		Lineare Regression	Linreg: Backhaus et al., (2021, S. 117)
Linearer Zusammenhang metrischen UVs und ihrer Logit-Transformation	Spezifisch: I-A; I-III, VI	Logistische Regression	Stoltzfuß (2011, S. 1101)
Linearität in den Parametern	IV, V	Lineare Regression	Backhaus et al. (2021, S. 101)
Kein Underfitting	IV, V		
Normalverteilung der Residuen	IV, V		
Homoskedastizität	IV, V		
Keine Multikollinearität der UVs	Spezifisch: I-A; I-III, VI IV, V	Logistische Regression	Logreg: Backhaus et al., (2021, S. 379), Stoltzfuß (2011, S. 1101)
		Lineare Regression	Linreg: Backhaus et al. (2021, S. 121)
keine Ausreißer	Univariat: übergreifend Spezifisch: I-A; -III, VI IV, V	Logistische Regression	Logreg: Stoltzfuß (2011, S. 1101)
		Lineare Regression	Linreg: Backhaus et al. (2021, S. 125)

Modellübergreifende Prämissen. Für alle Modelle, die mit einer logistischen Regression berechnet werden, müssen die abhängigen Variablen nominalskaliert sein und genau zwei

⁵⁴ Wenn die UVs ausschließlich kategorial sind, eignen sich zur Bestimmung der ungerichteten Zusammenhänge log-lineare Modelle.

Ausprägungen haben (Eid et al., 2017). Dies kann nicht mit einem Test per se überprüft werden, sondern muss anhand der Überlegungen zum Untersuchungsdesign bestimmt werden. Alle drei abhängigen Variablen der Suchstrategie und die abhängige Variable der Evaluationsstrategie sind nominalskaliert und haben zwei Ausprägungen. Insofern ist für alle entsprechenden Modelle, die mit einer logistischen Regression ermittelt werden, diese Prämisse erfüllt.

Weiterhin gilt, dass die unabhängigen Variablen teilweise oder ganz metrisch sind (Eid et al., 2017). Auch an dieser Stelle müssen die Forschenden dies vielmehr am Untersuchungsdesign festmachen und sich die Daten hinsichtlich ihrer Variablenskalierung anschauen. In Anbetracht der Personen- als auch Webseitenmerkmale, die mit Ausnahme des Geschlechts und der Domäne somit fast alle metrisch skaliert sind, kann diese Prämisse ebenfalls bestätigt werden.

Eine weitere modellübergreifende Prämisse für sowohl die lineare als auch logistische Regression ist, dass die Beobachtungen der abhängigen Variablen unabhängig voneinander sein müssen (Backhaus et al., 2021; Eid et al., 2017). Voneinander abhängige Daten könnten vorliegen, wenn Messwiederholungen oder gematchte Daten vorliegen. Dies wird nicht per se mithilfe von SPSS gemessen, sondern liegt bereits bei der Konstruktion des Untersuchungsdesigns vor oder nicht. Da das Untersuchungsdesign nicht aus wiederholten Messungen oder gematchten Daten bestand, kann auch diese Prämisse erfüllt werden.

Eine Prämisse, welche für alle linearen Regressionen gilt, ist die fehlerfreie Messung der unabhängigen Variablen (Backhaus et al., 2021). Zufällige Fehler in der abhängigen Variable werden durch Fehlerterme in der Regressionsgleichung berücksichtigt, jedoch können auch zufällige Fehler bei der Erhebung der UVs auftreten (Backhaus et al., 2021). Dies ist jedoch in der Praxis schwierig zu messen, da die Fehlervarianzen meistens unbekannt sind (Backhaus et al., 2021). Es sei an dieser Stelle jedoch der Vollständigkeit wegen erwähnt.

Eine ausreichende Fallzahl stellt ebenfalls eine gemeinsame, modellübergreifende Prämisse für die logistische und lineare Regression dar, die ebenfalls vielmehr mit dem Untersuchungsdesign und der eigentlichen Erhebung erreicht werden kann. Für die logistische Regression gilt, dass die Fallzahl pro Kategorie der AV nicht kleiner als 25, also in Summe nicht kleiner als $n = 50$, sein darf (Backhaus et al., 2021; Eid et al., 2017). Dies ist bei den Modellen mit NfC und IO mit einem $n = 146$ und bei dem Modell mit NfCC mit einem $n = 64$ erfüllt. Außerdem wird empfohlen, dass bei einer größeren Anzahl an UVs auch größere Fallzahlen, d. h. 10 Fälle pro zu schätzenden Parameter, existieren sollten (Backhaus et al., 2021). Für die linearen Regressionen werden ebenfalls etwa 10 Beobachtungen pro Parameter als grober Richtwert angenommen (Backhaus et al., 2021). Es finden sich zwischen 6 und 19 Parameter in den Modellen, was in den meisten Modellen bei $n = 146$ und $n = 64$ eine ausreichend große

Fallzahl darstellt. Diejenigen wenigen Modelle, bei denen das nicht zutrifft, werden in der Diskussion dahingehend kritisch beurteilt.

Autokorrelation bedeutet, dass die Residuen mit sich selbst korrelieren würden, was zur Folge hätte, dass die Präzision der Schätzung abnimmt (Backhaus et al., 2021). Da Autokorrelation hauptsächlich in Zeitreihen auftritt und dies in vorliegendem Datensatz keine Längsschnittdaten sind, könnte Autokorrelation von vornherein ausgeschlossen werden. Dennoch wird testweise eine lineare Regression durchgeführt, bei der alle unabhängigen Variablen integriert werden. Die Ergebnisse können dann analog auf alle zu berechnenden Modelle übertragen werden. Zur Erkennung, ob Autokorrelation vorliegt, wird sich des Durbin-Watson-Tests bedient, wobei Werte nahe 4 oder 0 auf eine Autokorrelation hinweisen (Backhaus et al., 2021). Optimalerweise liegt der Wert bei 2 und weist somit auf die Abwesenheit von Autokorrelation hin. Der Durbin-Watson-Test für vorliegenden Datensatz ergab $DW = 1.476$, was somit gegen Autokorrelation und eine Prämissenverletzung spricht.

Zuletzt dürfen sich keine Ausreißer in den Daten befinden. Zur Überprüfung der Datenqualität wurden die einzelnen Variablen auf Ausreißer analysiert. Hierzu wurde sich der univariaten Methode des Boxplots bedient. Eine Ausreißer-Analyse in Bezug auf zwei oder mehr Variablen wird modellspezifisch für die entsprechenden Variablen des Modells innerhalb der Berechnungen durch Betrachtung von Scatterplots durchgeführt. Da der Fokus an dieser Stelle jedoch nur auf die Überprüfung einer grundlegenden Datenqualität, unabhängig der angedachten Modelle und Analysen, liegt, reicht die Betrachtung eines Boxplots. Mithilfe von SPSS und den Schritten *Analysieren – deskriptive Statistik – explorative Datenanalyse* wird die Analyse durchgeführt. Alle unabhängigen sowie abhängigen metrischen Variablen wurden in das Bedienfeld „abhängige Variable“ gezogen und mithilfe der Option „Statistik – Ausreißer“ auf Ausreißer analysiert. Die Outputs *Tabelle Extremwerte*, *Stamm-Blatt-Diagramm* sowie der *Boxplot* lassen somit erkennen, ob und welche Fälle als Ausreißer identifiziert werden.

Mithilfe der Boxplot-Funktion werden alle Fälle grafisch in einem Diagramm dargestellt, wobei die Mittellinie des Boxplot den Median darstellt, d. h. 50 % der Werte liegen unterhalb bzw. oberhalb dieser Linie (Döring & Bortz, 2016). Die Boxhöhe umfasst 50 % der Werte (Backhaus et al., 2021, S. 51 f.). Der obere Rand kennzeichnet das dritte Quartil (75 %-Perzentil) und der untere das erste Quartil (25 %-Perzentil). Die Länge des Boxplots ist die Differenz des dritten und ersten Quartils, d. h. repräsentiert den Interquartilsabstand (IQA). Die Striche werden als Antennen bezeichnet. Mithilfe des IQA lassen sich die Ausreißer berechnen und grafisch darstellen: der 1.5-fache IQA addiert auf das obere Ende des Boxplots oder subtrahiert ab dem unteren Ende des Boxplots werden mithilfe der Antennen dargestellt und kennzeichnen das äußere Ende eines Toleranzbereiches (Backhaus et al., 2021, S. 52). Alle Werte außerhalb

dieses Bereiches werden als *einfache* Ausreißer angesehen. Fälle, die sich beim 2.5-fachen IQA befinden, werden als *extreme* Ausreißer definiert (Abbildung 18).

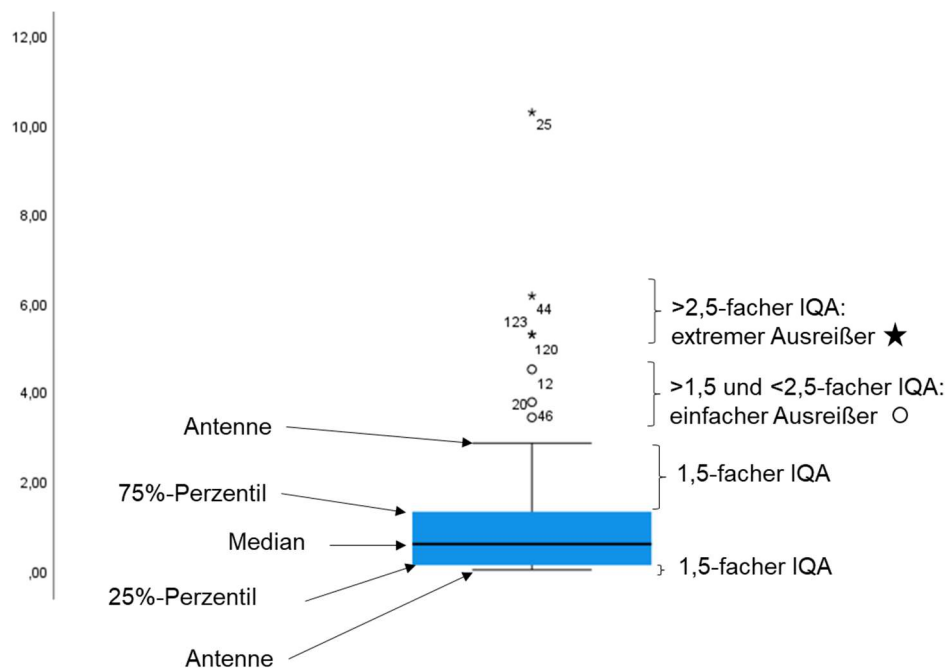


Abbildung 18: *Beispielhafter Boxplot.*

Die Analyse der Ausreißer zeigt, dass drei einfache Ausreißer vorlagen, insbesondere jeweils zwei Mal bei den Probanden *ID27* und *ID94* (Tabelle 19).

Tabelle 19: *Modellübergreifende Ausreißer-Analyse*

Variablen	Fälle als Ausreißer erkannt (einfache, <i>extreme</i>)
NfC	keine
NfCC	keine
IO	keine
Textglaubwürdigkeit besuchter Webseiten	<i>ID27</i>
Medienglaubwürdigkeit besuchter Webseiten	Keine
Autorglaubwürdigkeit besuchter Webseiten	Keine
Textglaubwürdigkeit selektierter Webseiten	<i>ID27</i> <i>ID94</i>
Systematische Begründung	keine
Korrekte Quelleneinschätzung	<i>ID104</i> <i>ID94</i>

Aktuelle Literatur empfiehlt, den Grund für den Ausreißer zu finden (Backhaus et al., 2021). Nur wenn es einen Grund für den Ausreißer gibt, kann er ausgeschlossen werden. Ausschlussgründe sind beispielsweise Mess- oder Eingabefehler oder ein ungewöhnliches Ereignis außerhalb des Forschungskontextes (z. B. Stromausfall) (Backhaus et al., 2021, S. 53)

Eine individuelle Überprüfung der einfachen Ausreißer mithilfe des Datensatzes zeigte, dass keine Messfehler (z. B. undefinierte fehlende Werte oder Zeichenfehler) vorlagen. Ein

ungewöhnliches Ereignis, wie ein Stromausfall, wäre anhand der Logdaten zu erkennen, sodass dies ebenfalls kein Ausschlussgrund liefert. Da angenommen wird, dass die Ausreißer somit keinen Fehler repräsentieren oder durch ein außergewöhnliches Ereignis verursacht wurden, werden sie nicht aus der Stichprobe entfernt.

Modellspezifische Prämissen. Bei einer logistischen Regression muss es einen linearen Zusammenhang zwischen den metrischen unabhängigen Variablen (z. B. NfC, IO) und dem Logit der AV geben (Tabachnick & Fidell, 2007). Ist dies nicht so, müsste eine Transformation der UV vorgenommen werden, da ansonsten die Schätzungen die wahren Abhängigkeiten nicht widerspiegeln und die Messungen verzerren. Zur Überprüfung dieser Prämisse kann das Verfahren von Box und Tidwell (1962) herangezogen werden. Hierzu werden die metrischen, unstandardisierten, unzentrierten Variablen mithilfe des Logarithmus $\ln$ transformiert. Anschließend werden Interaktionsterme zwischen den metrischen Variablen und ihrer logarithmierten Form gebildet. Zuletzt wird die logistische Regression unter Einbezug aller Variablen und der neu gebildeten Interaktionsterme berechnet. Ist die Interaktion nicht signifikant, ist die Prämisse der Linearität des Logits erfüllt. In der Praxis ist es jedoch üblich, diese implizit anzunehmen und sie erst dann zu überprüfen, wenn theoretisch unerwartete oder nicht interpretierbare Schätzungen eingehen. Dieser Praxis wird sich angeschlossen, sodass die entsprechenden Prämissen erst bei Aufkommen von insignifikanten Modellen oder Koeffizienten geprüft und diskutiert werden.

Auch bei einer linearen oder multiplen Regression gilt die Prämisse der Linearität, d. h. die Parameter im Modell sollten linear sein (Backhaus et al., 2021), da ansonsten die geschätzten Parameter verzerrt sind (Backhaus et al., 2021). Zur Entdeckung einer möglichen Nichtlinearität kann ein Streudiagramm erstellt und untersucht werden. Bei linearen bzw. multiplen Regressionen werden dazu jede unabhängige Variable mit der abhängigen Variablen in den Plot aufgetragen (Backhaus et al., 2021). Wünschenswert ist eine Streuung um den linearen Graph. Alternativ und auf effizientere Weise kann auch der Tukey-Anscombe-Plot verwendet werden, welcher die Residuen gegen die geschätzten y-Werte auf der x-Achse aufträgt. Die geschätzten Y-Werte sind Linearkombinationen der x-Werte, weshalb bei multipler Regression lediglich eine Grafik statt zwei oder mehreren (je nach Anzahl der UVs) erstellt werden muss. Hierbei ist das Ziel, dass die Residuen ohne Struktur oder Muster und somit zufällig streuen, dass die Linearität erfüllt werden kann. Backhaus empfehlen zum Visualisieren (auch Plotten genannt) der Werte bzw. Residuen die standardisierten (nach Backhaus et al., 2021, S. 143).

Weiterhin darf kein Underfitting vorliegen, d. h. es dürfen keine relevanten Variablen ausgelassen worden sein, sonst wäre die Schätzung verzerrt (Backhaus et al., 2021). Zur Inspektion dieses möglichen Problems wird sich ebenfalls des Tukey-Anscombe-Plots bedient, bei dem

die Fehlerterme gegen die geschätzten y -Werte auf der x -Achse abgetragen werden (Backhaus et al., 2021, S. 107 f.). Für die einfache Regression können die Residuen gegen die x -Variable auf der x -Achse geplottet werden, da die geschätzten y -Werte Linearkombinationen von x sind (Backhaus et al., 2021, S. 108). Wünschenswert ist die zufällige und gleichmäßige Streuung (ohne systematische Muster) um die x -Achse.

Eine weitere Prämisse für die einfache oder multiple Regression ist, dass die Residuen normalverteilt sein müssten, da ansonsten Probleme bei der Gültigkeit von Signifikanztests resultieren könnten (Backhaus et al., 2021). Das Vorgehen für große Stichprobe, d. h. > 40 , so wie in vorliegendem Fall, ist die Inspektion des Histogramms, welches innerhalb der Regression in SPSS automatisch mit ausgegeben werden kann. Die Residuen sollten symmetrisch und annähernd normalverteilt um die Normalverteilungskurve abgebildet sein. Rechnerisch kann man sich zudem noch des Shapiro-Wilk-Testes bedienen, der in SPSS über die Funktion *explorative Datenanalyse* auswählbar ist. Der Test geht in der Nullhypothese davon aus, dass die Residuen normalverteilt sind. Ziel für die Annahme der Prämisse ist folglich, dass die Nullhypothese nicht verworfen wird, d. h., dass $p > .05$ ist.

Eine weitere Prämisse für das lineare Regressionsmodell ist, dass die Residuen eine konstante Varianz haben müssen, was als Homoskedastizität bezeichnet wird (Backhaus et al., 2021), ansonsten würden die Standardfehler der Regressionskoeffizienten verzerrt werden. Dies kann einerseits visuell durch Inspektion des Plots ermittelt werden. Hierfür werden die unstandardisierten Residuen gegen die unstandardisiert vorhergesagte Werte in einem Streudiagramm abgebildet. Ist Heteroskedastizität vorhanden, bilden die Daten ein Muster in der Art eines Trichters bzw. Dreiecks. Ist dem nicht so, liegen homoskedastische Residuen vor und die Prämisse ist nicht verletzt. Zudem kann dies auch rein rechnerisch in SPSS mithilfe des White-Tests oder des Breusch-Pagan-Tests ermittelt werden. Hierzu wird die Funktion *Univariate Analyse – Optionen* verwendet und die entsprechenden Tests ausgewählt. Der White-Test würde verwendet werden, wenn die Residuen nicht normalverteilt sind, der Breusch-Pagan-Test, wenn sie normalverteilt sind. Die Nullhypothese der beiden Tests geht von Homoskedastizität aus. Wenn der Test signifikant ist, kann die Nullhypothese nicht angenommen werden und es liegt Heteroskedastizität vor, andernfalls schon.

Weiterhin darf keine Multikollinearität vorliegen. Multikollinearität bedeutet, dass die unabhängigen Variablen zu stark miteinander korrelieren, sodass als Folge dessen die Effekte der beiden Variablen auf die abhängige Variable nicht mehr voneinander zu trennen wären (Backhaus et al., 2021). Für sowohl die lineare (Backhaus et al., 2021) als auch die logistische (Eid et al., 2017) Regression ist dies eine Prämisse, die nicht verletzt werden darf. Insofern werden alle unabhängigen Personen- und Webseitenmerkmale in einer bivariaten Korrelationsmatrix

dargestellt, um mögliche Multikollinearität zu entdecken. Betreffende Modelle sind folglich alle Modelle, in denen mehrere Prädiktoren gleichzeitig im Modell untersucht werden (in I-A, II-A, II-B, II-C, III-A, III-B, III-C sowie IV-A, V-A, VI-A, VI-B und VI-C). Von einem Multikollinearitätsproblem spricht man, wenn der Korrelationskoeffizient $r > .80$ (Dattalo, 2013, S. 14) bzw. auch $> .90$ (Tabachnick & Fidell, 2007, S. 89) ist. In vorliegendem Datensatz lagen keine signifikanten und derart hochkorrelierten bivariaten Zusammenhänge vor. Der Ausschluss der Multikollinearität mithilfe des Korrelationskoeffizienten kann jedoch nicht immer vollständig gewährleistet sein (Belsley et al., 1980, S. 89). Aus diesem Grund wird der Varianz-Inflations-Faktor (*VIF*) zur Bestimmung herangezogen (Backhaus et al., 2021, S. 123). Zur Berechnung des *VIF* wird jede unabhängige Variable auf die anderen unabhängigen Variablen regressiert. Daraus resultiert der Korrelationskoeffizient R_j^2 mit Hilfe dessen der *VIF* berechnet wird:

$$VIF_j = \frac{1}{1 - R_j^2}$$

Je größer dieser ist, desto wahrscheinlicher ist es, dass eine unabhängige Variable durch eine Linearkombination aus anderen unabhängigen Variablen darstellbar ist. Somit würde Multikollinearität vorliegen. Auch hier gibt es keine festen Regeln (Field, 2013, S. 325), jedoch gilt in der Praxis, dass ein $VIF > 10$ auf Multikollinearität hindeuten könnte (O'Brien, 2007).

In SPSS gibt es keine gesonderte Funktion für den *VIF* für logistische Regressionen. Aus diesem Grund wird sich des Programmes JASP bedient, welches für alle Regressionen diese Möglichkeit bereitstellt. JASP wird somit ergänzend zu den Analysen mit SPSS verwendet. Insofern werden für jede Modelle, seien es logistische oder lineare Regressionen, die *VIF*-Werte gesondert geprüft.

Zusätzlich zu den Ausreißern in den grundsätzlichen Datensätzen, die bereits thematisiert wurden und z. B. in Messfehlern zu Grunde liegen, müssen auch modellspezifische Ausreißer identifiziert werden. In SPSS ist die Entdeckung von Ausreißern möglich, indem in der Maske zur Durchführung der logistischen oder linearen Regression *Optionen – Fallweise Auflistung der Residuen sowie Ausreißer außerhalb 2 Standardabweichungen* ausgewählt wird. Der Output ist eine Liste mit Ausreißern $\pm 2 SD$. Ausreißer werden als solche charakterisiert, wenn ihre studentisierten Residuen (*SResid*) im Bereich $\pm 3 SD$ (Pardoe, 2021, S. 224) bzw. zwischen 2 und 3 *SD* (vom Hofe, 2010, S. 1266) liegen. Die Literatur ist sich hierin nicht final einig. Eine nähere Betrachtung der Liste mit den *SResid* ist in jedem Fall notwendig und erfordert ggf. weitere Analysen mithilfe von Cook's Distanz (Cook, 1977). Diese basiert auf den studentisierten Residuen und klassifiziert Ausreißer als einflussreiche Ausreißer, sofern ihre Cook's Distanz > 1 ist (Backhaus et al., 2021, S. 133). Werden in nachfolgenden Modellberechnungen

Ausreißer als einflussreich klassifiziert, sollte immer im Einzelfall entschieden werden, ob dieser gelöscht und die Analyse wiederholt wird oder ob er weiterhin inkludiert bleibt (Backhaus et al., 2021, S. 134). In beiden Fällen wird eine ausreichende Dokumentation dargeboten.

5.3.2. Einfluss der Personen- und Webseitenmerkmale auf die abhängigen Variablen der Suchstrategie

5.3.2.1. Modell I A-B mit Spezifität der Suchbegriffe

Modell I-A. Für Modell I-A wird mit den soziodemografischen Merkmalen als Kontrollvariablen und mit den Personenmerkmalen NfC und IO eine schrittweise multiple logistische Regression gerechnet, die die abhängige Variable, Spezifität der Suchbegriffeingabe, beeinflussen könnte. Die Prüfung der Signifikanz des Gesamtmodells folgt zunächst mithilfe der sogenannten Likelihood-Ratio-Statistik bzw. dem Omnibustest der Modellkoeffizienten, der einer Chi²-Verteilung folgt. Anschließend wird das Nagelkerkes R^2 als Pseudo-R-Quadrat-Statistik herangezogen. Anhand der ersten Modellberechnung werden die theoretischen Hintergründe zu der Likelihood-Ratio-Statistik und dem Nagelkerkes R^2 ausführlich erklärt. Für die danach folgenden Modelle werden lediglich die Ergebnisse präsentiert.

Die Likelihood-Ratio-Statistik (im Folgenden abgekürzt mit LLR) vergleicht die Wahrscheinlichkeit des untersuchten Modells (L_v : Soziodemografie und NfC, IO) mit dem des Nullmodells (L_0 ; nur Konstante) (Backhaus et al., 2021, S. 327) und bedient sich dabei folgender Formel:

$$LLR = -2 \times \ln \left(\frac{L_0}{L_v} \right) = -2 \times (LL_0 - LL_v)$$

SPSS berechnet bei Durchführung der logistischen Regression $-2LL$, was mithilfe einer Umformung der folgenden Formel die Berechnung des LL_0 und LL_v ermöglicht. Denn es gilt zudem:

$$-2LL = -2 \times LL$$

LL_0 und LL_v sind in vorliegendem Fall laut SPSS $LL_0 = 185.763$ und $LL_v = 178.40$. Dividiert durch 2, ergeben sich folgende Werte für LL_0 und LL_v

$$LL_0 = 185.763 \div 2 = 92.8815$$

$$LL_v = 178.4 \div 2 = 89.20$$

Unter Einsatz der Werte in die Formel zur Berechnung von LLR ergibt sich die Gleichung:

$$LLR = 2 \times (92.8815 - 89.20) = 7.363$$

Da LLR unter der Nullhypothese $H_0: \beta_1 = \beta_2 = \dots \beta_J = 0$ annähernd Chi²-verteilt ist und somit für vorliegenden Fall mit $df = 8$ Freiheitsgraden besitzt, kann damit die statistische Signifikanz des

angepassten Modells getestet werden (Backhaus et al., 2021, S. 328). Die Nullhypothese geht von Gleichheit aus. Der tabellierte χ^2 -Wert bei $\alpha = 0.05$ und $df = 8$ beträgt 15.51. Da der hier berechnete, empirische $LLR = 7.363$ und somit < 15.51 ist, kann die Nullhypothese nicht verworfen werden. Das finale Modell mit NfC und IO ist somit statistisch nicht signifikant. Der SPSS-Output zeigt direkt die Ergebnisse in der Tabelle *Omnibus-Test der Modellkoeffizienten*, welche somit identisch sind. Aus diesem Grund werden für die Berechnungen nachfolgender Modelle direkt die SPSS-Outputs zur Analyse der Modellgüte herangezogen. Die Berechnung wird hier lediglich einmal exemplarisch aufgezeigt, um die einzelnen Schritte zu verdeutlichen.

Ein weiterer Test für die Modellgüte, der ebenfalls automatisch in SPSS mit ausgegeben werden kann, ist der Hosmer-Lemeshow-Test. Dieser wird verwendet, sofern gilt (Rossi, 2022, S. 472):

$$p < \frac{n}{20}$$

D. h., die Anzahl der Parameter muss kleiner sein als die Stichprobengröße dividiert durch 20. Ist der Test nicht signifikant, könnte von einer ausreichenden Anpassungsgüte zwischen dem Modell und den beobachteten Daten ausgegangen werden (Rossi, 2022). Da allerdings der Hosmer-Lemeshow-Test keine neuen Erkenntnisse liefert, mit dem Likelihood-Ratio-Test bereits die Modellgüte getestet wird und zudem immer vorab die eben gezeigte Bedingung geprüft werden muss, werden lediglich der Omnibus- bzw. der Likelihood-Ratio-Test herangezogen und berichtet.

Ein weiteres Maß zur Anpassungsgüte ist das Nagelkerkes R^{255} , welches zwischen 0 und 1 liegt. In der logistischen Regression ist nicht, wie bei der linearen Regression, von einem Bestimmtheitsmaß R^2 die Rede, sondern von den Pseudo- R^2 -Statistiken. Diese wurden entwickelt, um ein ähnliches Maß zum R^2 zu schaffen. Sie ähneln diesem aufgrund der Tatsache, dass sie ebenfalls nur Werte zwischen 0 und 1 annehmen können und dass ein höherer Wert eine bessere Anpassung widerspiegelt. Unterschiedlich ist jedoch, dass sie keinen Anteil messen, sondern auf dem Verhältnis zweier Wahrscheinlichkeiten beruhen (Backhaus et al., 2021, S. 329). Im vorliegenden Fall liefert das Nagelkerkes $R^2 = .071$ (Tabelle 20) nur eine geringe Anpassungsgüte (Backhaus et al., 2006, S. 456)⁵⁶.

Tabelle 20: *Modellgüte des Modells I-A NfC und IO und Spezifität des Suchbegriffs.*

Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
$\chi^2(8) = 7.363, p = .498$	.071

⁵⁵ Es wurde an dieser Stelle dem Nagelkerkes R^2 vor dem Cox & Snell- R^2 und dem McFaddens R^2 den Vorzug gegeben, da das Nagelkerkes R^2 eine Modifikation der Formel vom Cox & Snell- R^2 ist und dadurch Werte bis 1 zulässt (Backhaus et al., 2021, S. 330), was eine bessere Vergleichbarkeit mit dem normalen R^2 aus der linearen Regression erlaubt.

⁵⁶ Einteilung von Backhaus et al., 2006, S. 456: $R^2 > .2$ = akzeptabel, klein; $> .4$ = gut, mittel; $> .5$ = sehr gut, groß.

Sofern das Modell signifikant ist, können eigentlich weitere Kriterien zur Beurteilung herangezogen werden, welche im Folgenden ebenfalls theoretisch erläutert werden.

Die durch das Modell geschätzten und im Datensatz beobachteten Gruppenzugehörigkeiten können herangezogen werden, um zukünftige Gruppenzugehörigkeiten zu prognostizieren (Backhaus et al., 2021). Die Prognose der Gruppenzugehörigkeit wird als Klassifikation bezeichnet. Der Trennwert, ob spezifische oder oberflächliche Suchbegriffe verwendet werden, wird standardmäßig bei $p = 0.5$ festgesetzt. Der Gesamtprozentsatz korrekter Klassifikation bzw. die Trefferquote ist 63.4 %. Die Trefferquote gibt den Anteil der korrekten Prognosen bezüglich der Fälle an (Backhaus et al., 2021). Übersetzt heißt das, dass in 63.4 % der Fälle eine korrekte Prognose, also die korrekte Zuteilung in die Gruppen (spezifische vs. oberflächliche Suchbegriffe) durch das Modell vorhergesagt wird. Zwei weitere, damit einhergehende Maße zur Beurteilung der Prognose durch das Modell sind die Sensitivität und die Spezifität (Backhaus et al., 2021). Sensitivität ist der Anteil der korrekt vorhergesagten Personen, die ein bestimmtes Merkmal auch aufweisen (Backhaus et al., 2021); hier im Fall von Personen, die spezifische Suchbegriffe verwenden. Diese liegt bei 68.7 %. Die Spezifität gibt den Anteil der korrekt vorhergesagten Personen an, die ein Merkmal nicht aufweisen (Backhaus et al., 2021). Hier in dem Fall wäre es das Gegenteil zur spezifischen Suchbegriffseingabe, also die oberflächliche Suchbegriffseingabe. Diese liegt bei 58.2 %. Das Modell sagt zu 68.7 % korrekt die Gruppenzugehörigkeit der Personen, die spezifische und zu 58.2 % korrekt die Gruppenzugehörigkeit der Personen, die oberflächliche Suchbegriffe verwenden, voraus.

Problematisch ist an der Beurteilung der Klassifikation jedoch, dass der selbst gewählte Trennwert zu Grunde gelegt wird (Backhaus et al., 2021, S. 309). Wird der Trennwert beispielsweise auf $p = 0.2$ festgesetzt, verändern sich die Gruppenzugehörigkeiten und Prognosen drastisch. Aus diesem Grund wird mit der *Receiver Operating Characteristics* (im Folgenden abgekürzt mit ROC)-Kurve zusätzlich ein verallgemeinertes Konzept zur Beurteilung angewandt, nämlich das Maß für die Güte der Prognosefähigkeit (Backhaus et al., 2021, S. 310). Jeder Punkt auf der ROC-Kurve repräsentiert einen gewählten Trennwert und die entsprechende Klassifizierungstabelle. Die 45°-Diagonale repräsentiert eine rein zufällig getätigte Prognose. Das Ziel ist es, einen durch die Daten abgebildeten Graph zu erhalten, der möglichst entfernt von der Diagonalen liegt. Rechnerisch kann dies mit dem Wert *Area under Curve* (im Folgenden abgekürzt mit AUC) bestimmt werden. Dieser berechnet die Fläche der ROC-Kurve und ist maximal 1 (Backhaus et al., 2021, S. 310).

Die Berechnung von AUC wird für vorliegendes Modell mithilfe von JASP getätigt und beträgt .633. Dieser Wert gilt lt. Literatur als ungenügend (Backhaus et al., 2021, S. 311)⁵⁷ und spricht für eine schlechte Güte der Prognosefähigkeit, was in Anbetracht der Insignifikanz des Modells nicht verwunderlich ist.

Die Prüfung des Gesamtmodells mithilfe des Omnibus-Tests zeigt also, dass das Modell nicht signifikant ist. In Summe sind überwiegend Prädiktoren im Modell enthalten, die nicht signifikant sind. Zugleich bedeutet es jedoch nicht, dass nicht alle Regressionskoeffizienten insignifikant sein müssen. Aus diesem Grund werden die Regressionskoeffizienten mithilfe des Likelihood-Quotienten-Tests (im Folgenden abgekürzt mit *LLQ*) überprüft, sodass Aussagen zu möglichen Einflüssen getätigt werden könnten.

Zur Überprüfung der Signifikanz der Variablen kann sich entweder auf den Wald-Test oder auf den *LLQ* referiert werden (Backhaus et al., 2021). Dem *LLQ* wird der Vorzug gegeben, da der Wald-Test immer systematisch größere *p*-Werte liefert und daher versagen könnte (Backhaus et al., 2021). In SPSS ist die Berechnung des *LLQs* nur über multinomial logistische Regression möglich. Ein Vergleich der beiden Outputs der binomial und der multinomial logistischen Regression hinsichtlich der Modellgüten zeigt, dass es keine Unterschiede in der Signifikanz oder der Schlussfolgerung gibt, weshalb sich für den *LLQ* über die Funktion der multinomialen logistische Regression entschieden wird. Da eine schrittweise Berechnung nicht automatisch möglich ist, werden manuell die einzelnen Variablen integriert. Außerdem werden zur weiteren Beurteilung der Regressionskoeffizient β sowie der $\text{Exp}(\beta)$ aus der binomialen Regressionsberechnung aufgeführt. $\text{Exp}(\beta)$ stellt die Odds Ratio (im Folgenden abgekürzt mit *OR*) dar, welche mithilfe von β ermittelt werden kann. Folgende Gleichung verdeutlicht den Zusammenhang zwischen dem Regressionskoeffizienten und der Odds Ratio:

$$OR = e^{\beta}$$

Die Odds Ratio ist ein Maß für das Verhältnis von zwei Chancen (Englisch: Odds) (Backhaus et al., 2021), hier im Beispiel für spezifische Suchbegriffe oder keine spezifischen, also oberflächliche, Suchbegriffe. Ein *OR* von 1 würde bedeuten, dass zwei Ereignisse gleich wahrscheinlich sind, ein $OR < 1$ bedeutet, dass es wahrscheinlicher ist, dass ein Ereignis nicht eintritt und ein $OR > 1$ bedeutet, dass es wahrscheinlicher ist, dass ein Ereignis eintritt im Vergleich zu nicht eintritt.

Tabelle 21 zeigt eine Zusammenstellung der entsprechenden zur Beurteilung der Koeffizienten heranzuziehenden Werte: der *LLQ* zur Beurteilung der Signifikanz aus der multinomial

⁵⁷ Folgende Abgrenzungen gelten lt. Backhaus et al. (2021, S. 311): $AUC < .7$ = ungenügend; $.7 \leq AUC < .8$ = akzeptabel; $.8 \leq AUC < 0.9$ = exzellent; $AUC \geq .9$ = außerordentlich.

logistischen Regression, der Regressionskoeffizient β und der $\text{Exp}(B)$ aus der automatischen, schrittweise binomial logistischen Regression. Die Tabelle wird trotz mangelnder Ergebnisse dargestellt, um zu verdeutlichen, welche Koeffizienten herangezogen werden. Für zukünftige Ergebnisdarstellung wird jedoch lediglich die Tabelle mit dem entsprechenden Modell dargestellt, welches auch zur Beschreibung der Ergebnisse genutzt wird. Die vollständigen Tabellen finden sich jedoch in Anhang A 3 wieder.

Tabelle 21: *Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell I-A NfC, IO und Spezifität der Suchbegriffseingabe.*

Logistische Regression mit AV Spezifität Suchbegriffe und UV NFCs und IO							
Mo- dellnr.	Variable	Regressionsko- effizientB	Standard- fehler	Likelihood-Quotienten- Test		Signif- ikanz	Exp(B)
				Chi ²	df		
1	Alter	.094	0.074	1.723	1	.189	1.099
	HZB-Note	-.218	0.359	0.373	1	.541	0.804
	Weiblich	-.330	0.424	0.607	1	.436	0.719
	Divers	.350	0.780	0.202	1	.653	1.419
	Lehramt	-.729	0.524	1.972	1	.160	0.483
	Jura	-.685	0.540	1.630	1	.202	0.504
	Konstante	-1.585	1.924	0.696	1	.404	0.205
2	Alter	.097	.074	1.825	1	.177	1.102
	HZB-Note	-.229	0.359	0.408	1	.523	0.796
	Weiblich	-.334	0.430	0.607	1	.436	0.716
	Divers	.358	0.795	0.204	1	.651	1.431
	Lehramt	-.742	0.526	2.033	1	.154	0.476
	Jura	-.684	0.546	1.590	1	.207	0.505
	NfC	-.037	0.182	0.041	1	.839	0.964
	IO	.082	0.183	0.200	1	.655	1.085
	Konstante	-1.647	1.932	0.746	1	.388	0.193

Wie zu erkennen ist, gibt es auch keine signifikanten Regressionskoeffizienten. Die Ergebnisse sprechen somit nicht für die Hypothesen H1a und H3a. Hohes NfC und niedriges IO führt nicht zur Verwendung von spezifischen Suchbegriffen.

Überprüfung der modellspezifischen Prämissen. Da das Modell unerwartet insignifikant ist und keinerlei Ergebnisse liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Für vorliegendes Modell wurde die entsprechende Variablentransformation sowie neue Berechnung der logistischen Regression durchgeführt. Anhand des Likelihood-Quotienten-Tests wurde keine Signifikanz der metrisch skalierten Variablen (Alter, HZB-Note, NfC, IO) und ihren Interaktionen identifiziert, sodass die Prämisse der Linearität erfüllt ist. Weiterhin darf kein Multikollinearitätsproblem vorliegen. Für die entsprechenden unabhängigen Variablen lag nach Betrachtung des Outputs kein $VIF > 10$ vor, weshalb für dieses Modell die Prämisse der Abwesenheit der Multikollinearität erfüllt ist. Die Analyse

möglicher Ausreißer fiel positiv in dem Sinne aus, als dass keinerlei Ausreißer mit $SResid \pm 2 - 3 SD$ vorhanden waren.

Modell I-B. Für Modell I-B wird in SPSS mit den soziodemografischen Merkmalen und dem Personenmerkmal NfCC eine schrittweise logistische Regression berechnet, die einen möglichen Einfluss auf die abhängige Variable, Spezifität der Suchbegriffseingabe, aufzeigen könnte. Die Anpassungsgüte des sowohl reduzierten als auch des finalen Modells ist nicht signifikant mit $\chi^2(5) = 4.718$, $p = .451$ für das reduzierte und $\chi^2(6) = 4.904$, $p = .556$ für das finale. Auch das Nagelkerkes R^2 mit $=.105$ würde gemäß der Literatur-Empfehlung (Backhaus et al., 2006) eine schlechte Varianzaufklärung darstellen. Trotz der insignifikanten Modellgüte wird der Likelihood-Quotienten-Test herangezogen, um signifikante Prädiktoren zu ermitteln.

Dieser wird analog zum Modell I-A mithilfe der Funktion *multinomial logistische Regression* durchgeführt, jedoch zeigt sich für keinen der Prädiktoren ein signifikantes Ergebnis, welches somit nicht für die entsprechende Hypothese H2a spricht. Ein hohes NfCC führt nicht zur Eingabe eines allgemeinen Suchbegriffs.

Überprüfung der modellspezifischen Prämissen. Da das Modell unerwartet insignifikant ist und keinerlei Ergebnisse liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Keine der metrisch skalierten Variablen (Alter, HZB-Note, NfCC) und ihren Interaktionen ist signifikant, sodass die Prämisse der Linearität erfüllt ist. Die Abwesenheit einer Multikollinearität wurde bestätigt, denn es lagen keine Variablen mit einem $VIF > 10$. Die Analyse möglicher Ausreißer zeigte keinerlei relevanten und einflussreichen Ausreißer auf.

5.3.2.2. Modell II A-C mit Nachrecherche

Modell II-A. Für Modell II-A wird mit NfC eine logistische Regression berechnet, mit dem Ziel, die Einflussfaktoren auf die abhängige Variable, Nachrecherche, zu ermitteln. Dies wird in einem hierarchischen Verfahren gemacht, d. h. schrittweise unter Einschluss weiterer Variablen. Modell 1 wird unter Einbezug der soziodemografischen Daten, Modell 2 unter Einbezug der unabhängigen Variablen (NfC, IO, Textglaubwürdigkeit, Medienglaubwürdigkeit etc.) und Modell 3 ist unter Einbezug der Interaktionsterme zwischen den beiden unabhängigen Variablen berechnet.

Es folgt zunächst die Prüfung des Gesamtmodells mithilfe der Likelihood-Ratio-Statistik und dem Nagelkerkes R^2 als Pseudo-R-Quadrat-Statistik. Die Modellgüte für das finale Modell ist mit $\chi^2(19) = 31.525$, $p = .035$ signifikant (Tabelle 22) und somit aussagekräftiger im Vergleich zu den reduzierten Modellen. Auch das Nagelkerkes R^2 mit $.290$ zeigt einen kleinen Effekt

(Backhaus et al., 2006). Da sich die beiden Maße als signifikant für die Modellgüte darstellen, können weitere Beurteilungen für die Qualität des Modells herangezogen werden.

Tabelle 22: *Modellgüte des Modells II-A NfC und Nachrecherche*

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(6) = 12.914, p = .044$	.127
2	Block: $\chi^2(7) = 5.571, p = .591$	.178
3	Modell: $\chi^2(13) = 18.485, p = .140$	.290
	Block: $\chi^2(6) = 13.039, p = .042$	
	Modell: $\chi^2(19) = 31.525, p = .035$	

Der Gesamtprozentsatz korrekter Klassifikation bzw. die Trefferquote ist 77.8 %. In 77.8 % der Fälle wird eine korrekte Prognose, d. h. die korrekte Zuteilung in die Gruppen (keine Nachrecherche vs. Nachrecherche) durch das Modell vorhergesagt. Die Sensitivität ist in dem Fall der Anteil der korrekt vorhergesagten Personen, die nachrecherchieren und liegt bei 47.7 %. Sie ist damit nicht wahrscheinlicher als der Ausgang eines Zufallsexperimentes mit zwei Möglichkeiten (z. B. Münzwurf). Die Spezifität gibt den Anteil der korrekt vorhergesagten Personen an, die nicht nachrecherchieren und ist mit 92.3 % sehr hoch. Das Modell sagt also zu 47.7 % korrekt die Gruppenzugehörigkeit der Personen, die nachrecherchieren, und zu 92.3 % korrekt die Gruppenzugehörigkeit der Personen, die nicht nachrecherchieren, voraus. Aufgrund der Problematik um den selbst gewählten Trennwert wird die AUC zur Beurteilung angewandt, welche .792 beträgt. Diese gilt lt. Literatur knapp als akzeptabel (Backhaus et al., 2021, S. 311).

Nachdem die Güte des Gesamtmodells überprüft wurde, können anschließend die einzelnen geschätzten Koeffizienten beurteilt und überprüft werden. Tabelle 23 zeigt das Ergebnis der schrittweisen logistischen Regression.

Tabelle 23: Ausschnitt der Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-A NfC und Nachrecherche

Logistische Regression mit AV Nachrecherche und UV NfC							
Mo- dellnr.	Variable	Regressi- onskoeffi- zientB	Stan- dard- fehler	Likelihood-Quotienten- Test			
				Chi ²	df	Signif- ikanz	Exp(B)
3	Alter	-.053	.069	.584	1	.445	.949
	HZB-Note	-.434	.446	.960	1	.327	.648
	Weiblich	-.158	.529	.089	1	.765	.854
	Divers	-1.062	1.058	1.103	1	.294	.346
	Lehramt	2.041	.714	9.198	1	.002	7.699
	Jura	.674	.666	1.016	1	.313	1.962
	NfC	.585	.384	2.488	1	.115	1.795
	Textglaubwürdigkeit	-.715	.388	3.635	1	.057	.489
	Textglaubwürdigkeit ²	.538	.329	2.774	1	.096	1.712
	Medienglaubwürdigkeit	-.214	.351	.375	1	.540	.807
	Medienglaubwürdigkeit	.073	.247	.087	1	.768	1.076
	Autorenglaubwürdigkeit	.798	.371	5.052	1	.025	2.220
	Autorenglaubwürdigkeit ²	-.046	.246	.036	1	.850	.955
	NfC*Textglaubwürdigkeit	.240	.423	.325	1	.569	1.272
	NfC* Textglaubwürdigkeit ²	-.586	.301	4.096	1	.043	.557
	NfC*Medieng- laubwürdigkeit	-.008	.249	4.229	1	.040	.992
	NfC* Medieng- laubwürdigkeit ²	.727	.363	0.001	1	.974	2.070
	NfC* Auto- renglaubwürdigkeit	-.625	.418	2.371	1	.124	.535
	NfC* Auto- renglaubwürdigkeit ²	.492	.362	1.935	1	.164	1.636
	Konstante	.175	1.969	0.008	1	.929	1.191

Der SPSS-Output zur Likelihood-Quotienten-Test gibt lediglich für einen soziodemografischen Prädiktor, die Domäne Lehramt ($\chi^2(1) = 10.158$, $p = .002$ und $OR = 7.669$), einen signifikanten Effekt aus (in Tabelle grau hinterlegt zur besseren Kennzeichnung). Das bedeutet, dass angehende Lehrer im Vergleich zu Medizinerinnen eine fast 8-fach größere Chance haben, dass sie nachrecherchieren.

Die Ergebnisse zeigen keinen direkten Effekt für NfC sowie für die einzelnen non-linearen Glaubwürdigkeitsindices. Lediglich ein direkter, jedoch linearer Effekt für den autorenbezogenen Glaubwürdigkeitsindex ist signifikant. Die Ergebnisse sprechen nicht für die Hypothesen H1b und H4a.

Es liegt jedoch ein Moderationseffekt vor, denn es zeigen sich zwei signifikante Interaktionseffekte: der non-lineare Effekt der textlichen Glaubwürdigkeit mit NfC sowie der lineare Effekt der mediumsbezogenen Glaubwürdigkeit mit NfC haben einen signifikant von Null verschiedenen Einfluss.

Für die Interpretation der Ergebnisse des NfC, den Glaubwürdigkeitsindices und ihren Interaktionen ist ein Zwischenschritt erforderlich, denn die komplexe Analyse- und

Modellierung erschwert eine direkte Aussage nur anhand der Vorzeichen aus den Daten. Die Ergebnisse werden daher mithilfe von R (Version 4.3.1) visualisiert und mit der wünschenswerten bzw. der hypothetisch hergeleiteten Visualisierung verglichen. Wie anhand der 2D-Grafiken in der Hypothesenherleitung zu erkennen, sollte der non-lineare Effekt eine nach unten geöffnete Parabel abbilden, d. h. konkav sein. Bei quadratischen Gleichungen ist der dazugehörige Koeffizient negativ. Bei konvexer und somit nach oben geöffneter Parabel wäre er positiv (Hayes, 2022a, S. 3). Der non-lineare Effekt sollte am Beispiel von NfC allerdings nur bei hohem NfC verwirklicht werden, wobei der lineare Effekt bei niedrigem NfC sichtbar würde. Da dies vom dazugehörigen Glaubwürdigkeitsindex abhängt und postuliert wird, dass sich dies unterschiedlich auf die Nachrecherche auswirkt, stellt die Visualisierung dieser Hypothesen ein 3D-Modell dar.

Das Plotten wird dabei wie folgt getätigt: Alle diejenigen Variablen, die laut *LR*-Test eine Signifikanz aufweisen, werden in die nachstehende Gleichung der logistischen Form (Kalisch & Meier, 2021, S. 19)

$$f(x) = \frac{e^y}{(1 + e^y)}$$

integriert, wobei $f(x)$ bzw. y die lineare Regressionsgleichung darstellt.

Aufgrund der Transformation in die logistische Form können alle signifikanten Variablen mit ihren Regressionskoeffizienten in die Gleichung eingefügt werden. Als Ergebnis werden je nach Ergebnis entweder ein 3D-Modell oder ein 2D-Modell dargestellt. Im 3D-Modell ist die x-Achse immer das Personenmerkmal (hier: NfC), die y-Achse die wahrgenommene Glaubwürdigkeit und die z-Achse die abhängige Variable (hier: Nachrecherche). Im 2D-Modell ist die y-Achse immer die abhängige Variable und die x-Achse die wahrgenommene Glaubwürdigkeit.

Zur besseren Interpretation werden die relevanten Erkenntnisse, d. h. aus der hypothetisierten Perspektive von NfC als Moderator, farblich hervorgehoben und mit E1 bis E5 (bzw. fortführend in den weiteren Modellen) als Abkürzung von verschiedenen Erkenntnissen durchnummeriert. Signifikante Interaktionseffekte werden nacheinander stets beginnend mit den textlichen Merkmalen berichtet. Danach folgen die mediumsbezogenen und abschließend autorenbezogenen Merkmale. Die Ergebnisse sind in Abbildung 19 visualisiert:

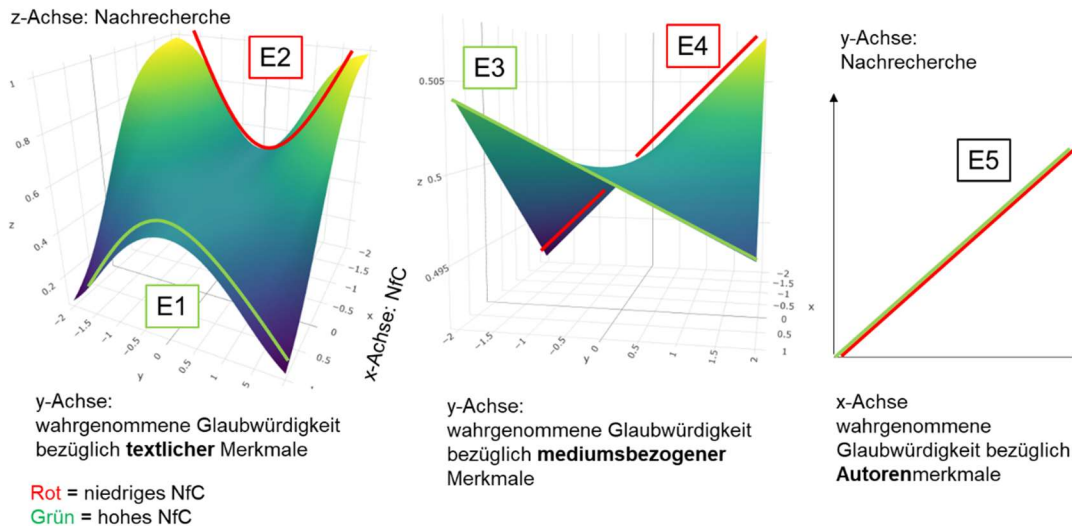


Abbildung 19: Effekte aus Modell II-A.

Erkenntnis 1 und 2. Personen mit hohem NfC (grüne Linie) recherchieren eher dann nach, wenn sie Webseiten mit einer unklaren oder mittelmäßigen Glaubwürdigkeit hinsichtlich textlicher Merkmale besuchen. Nachrecherchieren in diesem Kontext meint, dass entweder die zuvor gefundene Seite nochmals in einer Suchmaschine gesucht wurde oder dass auf der gefundenen Seite selbst nach Informationen über diese gesucht wird, durch Anschauen des Impressums oder die „Über Uns“-Beschreibung der Webseite selbst. Dies entspricht einem Teil der Moderationshypothese H4a. Somit ist ein non-linearer Effekt für die Probanden mit hohem NfC nachweisbar (E1). Personen mit hohem NfC recherchieren weniger nach (H4a), wenn sie Webseiten aufrufen, die aufgrund textlicher Merkmale als sehr glaubwürdig oder sehr unglaubwürdig wahrgenommen werden könnten (non-linearer Effekt). Es zeigt sich weiterhin, dass entgegen der postulierten Hypothese aus H1b ein non-linearer Effekt für Probanden mit niedrigem NfC nachweisbar ist (E2). Denn sie recherchieren umso mehr nach, wenn die Webseite entweder sehr viele oder sehr wenige textliche Merkmale für eine positive Glaubwürdigkeitsbeurteilung bereitstellt (E2). Ein linearer, konstanter Effekt wird somit nicht, wie in H1b postuliert, gefunden, weshalb diese abzulehnen ist.

Erkenntnisse 3 und 4. Bei Personen mit hohem als auch mit niedrigem NfC liegt ein linearer Effekt in Bezug auf mediumsbezogene Merkmale vor. Personen mit hohem NfC recherchieren umso mehr nach, je unglaubwürdiger die besuchte Webseite bezüglich mediumsbezogener Merkmale ist (E3). Das Ergebnis spricht somit nicht für die entsprechende Moderationshypothese H4a zum non-linearen Effekt. Hinsichtlich niedrigem NfC wurde vermutet, dass diese Individuen nicht nachrecherchieren – ungeachtet der besuchten Webseiten und ihren Merkmalen. Jedoch zeigt E4, dass Personen mit niedrigem NfC umso mehr nachrecherchieren, wenn Webseiten besucht werden, die bezüglich Medienmerkmale als glaubwürdig

wahrgenommen werden könnten. Dieses Ergebnis spricht somit erneut gegen die postulierte Moderationshypothese.

Erkenntnis 5. Da der Effekt von der Autorglaubwürdigkeit, d. h. der wahrgenommenen positiven Glaubwürdigkeit aufgrund vieler Autorenmerkmale, ein direkter Effekt ist und keinerlei Interaktion mit NfC vorliegt, wird dieser nicht in R geplottet, sondern zweidimensional dargestellt. Unabhängig von NfC wird umso mehr nachrecherchiert, wenn Webseiten besucht wurden, die viele Autorenmerkmale für eine positive Glaubwürdigkeitsbeurteilung bereitstellen. Das ist insofern gegen die postulierte Hypothese, da bei Personen mit hohem NfC ein non-linearer Effekt postuliert wurde. Bei Personen mit niedrigem NfC wurde zwar ein linearer Effekt postuliert, jedoch in dem Sinne, als dass sie nicht nachrecherchieren und die Linie daher konstant parallel zur x-Achse verläuft. Das Ergebnis spricht gegen die Hypothese, denn mit steigender Anzahl der besuchten Webseite, die Autorenmerkmale für eine positive Glaubwürdigkeitsbeurteilung beinhalten, recherchieren Personen unabhängig vom NfC nach.

Zusammenfassend ist zu sagen, dass die Ergebnisse weder für den direkten Effekt von NfC (H1b) noch für den direkten non-linearen Effekt aller unterschiedlicher Webseitenmerkmalsblöcke (H4a) sprechen. Es gibt zwar einen direkten Effekt mit den Autorenmerkmalen, dieser ist jedoch linear, was nicht postuliert wurde. Bezüglich des Moderationseffektes von NfC auf den Effekt der verschiedenen Merkmalsblöcke auf die Nachrecherche kann statuiert werden, dass die Ergebnisse nur teilweise für H4a für hohes NfC und die textliche Glaubwürdigkeit sprechen. Alle restlichen Hypothesen können nicht angenommen werden, was in nachstehender Tabelle 24 nochmals übersichtlich dargestellt ist:

Tabelle 24: *Hypothesenentscheidung Modell II-A.*

Direkte Effekte auf Nachrecherche	Status	Interaktionseffekte mit NfC auf Nachrecherche	Status
NfC (H1b)	✗	-	
Textglaubwürdigkeit	-	Niedriges NfC: H1b (E2)	✗
Textglaubwürdigkeit ² (H4a)	✗	Hohes NfC: H4a (E1)	✓
Medienglaubwürdigkeit	-	Niedriges NfC: H1b (E4)	✗
Medienglaubwürdigkeit ² (H4a)	✗	Hohes NfC: H4a (E3)	✗
Autorenglaubwürdigkeit	-	Niedriges NfC: H1b (E5)	✗
Autorenglaubwürdigkeit ² (H4a)	✗	Hohes NfC: H4a	✗

Anmerkung.

Legende: ✗ = Ergebnis spricht gegen Hypothese; ✓ = Ergebnis spricht für Hypothese; - = wurde nicht postuliert

Überprüfung der modellspezifischen Prämissen. Die Überprüfung der Linearität ist eigentlich eine erste Prämisse. Da allerdings sowohl das Modell in Summe als auch der postulierte Prädiktor signifikant waren, kann von einer ausreichenden Linearität ausgegangen werden, was die Prüfung hierfür hinfällig macht. Die Analyse von *VIF* zur Aufdeckung eines möglichen Multikollinearitätsproblems erzielte kein $VIF > 10$. Die Analyse der Ausreißer zeigte vier Fälle, deren *SResid* eine Standardabweichung $> \pm 2 SD$ haben und zu einem möglichen Problem

werden könnten. Aus diesem Grund wurde Cook's Distanz herangezogen, welche im Output als zusätzliche Werte abgespeichert werden können. Werte > 1 werden als einflussreiche Ausreißer betrachtet. Von den vier identifizierten Ausreißern könnte lediglich einer einflussreich sein. Es wurde daher unter Ausschluss des Ausreißers testweise eine neue Berechnung des Modells gemacht. Das neue Modell war ebenfalls signifikant, wobei sogar weitere signifikante Ergebnisse erzielt wurden. Auf folgendem Grund wurde jedoch das ursprüngliche Modell weiter untersucht: Es wurde bereits eine Ausreißer-Diagnostik auf Basis der Variablen durchgeführt, bei der keine relevanten Ausreißer gefunden wurden. Die hier gefundenen Ausreißer sind daher nur als modellbedingt zu betrachten. Ein Ausschluss dieser würde zu einer Anpassung des Modells an die Daten und an die Verformung der ursprünglichen Stichprobe führen, was nicht Ziel dieser Forschungsarbeit ist. Der Arbeit liegt keine Modelloptimierung zu Grunde, sondern die Beantwortung einer Forschungsfrage und dem damit verbundenen Hypothesen auf Basis einer Stichprobe.

Modell II-B. Für Modell II-B wird mit IO eine schrittweise logistische Regression berechnet, die einen möglichen Einfluss auf die abhängige Variable, Nachrecherche, nachweisen könnte. Modell 1 ist unter Einschluss der soziodemografischen Daten, Modell 2 mit Einbezug der unabhängigen Variablen (IO, Textglaubwürdigkeit, Medienglaubwürdigkeit etc.) und Modell 3 ist mit Einbezug der Interaktionsterme zwischen den beiden unabhängigen Variablen.

Die Anpassungsgüte des finalen Modells ist mit $\chi^2(19) = 30.107$, $p = .050$ signifikant (Tabelle 25) und somit aussagekräftiger im Vergleich zu den reduzierten Modellen. Auch das Nagelkerkes R^2 mit .279 zeigt einen akzeptablen, kleinen Effekt (Backhaus et al., 2006).

Tabelle 25: *Modellgüte des Modells II-B IO und Nachrecherche*

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(6) = 12.914$, $p = .044$	.127
2	<u>Block:</u> $\chi^2(7) = 5.514$, $p = .598$ <u>Modell:</u> $\chi^2(13) = 18.428$, $p = .142$	.178
3	<u>Block:</u> $\chi^2(6) = 11.679$, $p = .070$ <u>Modell:</u> $\chi^2(19) = 30.107$, $p = .050$	.279

Der Gesamtprozentsatz korrekter Klassifikation bzw. die Trefferquote ist 76.3 %. In 76.3 % der Fälle wird eine korrekte Prognose, d. h. die korrekte Zuteilung in die Gruppen (keine Nachrecherche vs. Nachrecherche) durch das Modell vorhergesagt. Die Sensitivität, sprich der Anteil der korrekt vorhergesagten Personen, die nachrecherchieren, liegt bei 45.5 % und ist damit sogar noch unwahrscheinlicher als der Ausgang eines Zufallsexperimentes mit zwei

Möglichkeiten (z. B. Münzwurf). Die Spezifität, der Anteil der korrekt vorhergesagten Personen, die nicht nachrecherchieren, ist mit 91.2 %. sehr hoch. Das Modell sagt demnach lediglich zu 45.5 % korrekt die Gruppenzugehörigkeit der Personen, die nachrecherchieren, und zu 91.2 % korrekt die Gruppenzugehörigkeit der Personen, die nicht nachrecherchieren, voraus. Aufgrund der Problematik um den selbst gewählten Trennwert wird die AUC zur Beurteilung der Prognosefähigkeit angewandt, welche für vorliegenden Fall .763 beträgt. Diese gilt lt. Literatur knapp als akzeptabel (Backhaus et al., 2021, S. 311).

Weiterhin werden die einzelnen Prädiktoren auf Signifikanz überprüft, was in nachstehender Tabelle 26 berichtet wird:

Tabelle 26: Ausschnitt der Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-B IO und Nachrecherche.

Logistische Regression mit AV Nachrecherche und UV IO							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
3	Alter	-.082	.076	1.223	1	.269	.922
	HZB-Note	-.260	.447	.340	1	.560	.771
	Weiblich	-.723	.531	1.863	1	.172	.485
	Divers	-2.017	1.058	4.138	1	.042	.133
	Lehramt	2.208	.754	9.710	1	.002	9.101
	Jura	.792	.666	1.405	1	.236	2.208
	IO	-.550	.405	1.925	1	.165	.577
	Textglaubwürdigkeit	-.594	.417	2.099	1	.147	.552
	Textglaubwürdigkeit ²	.289	.313	.868	1	.352	1.335
	Medienglaubwürdigkeit	.563	.388	2.302	1	.129	1.756
	Medienglaubwürdigkeit ²	-.321	.297	1.264	1	.261	.725
	Autorenglaubwürdigkeit	.239	.344	.485	1	.486	1.270
	Autorenglaubwürdigkeit ²	-.030	.236	.016	1	.898	.970
	IO*Textglaubwürdigkeit	.845	.479	3.386	1	.066	2.329
	IO*Textglaubwürdigkeit ²	.977	.409	8.160	1	.004	2.656
	IO*Medienglaubwürdigkeit	-.199	.430	.215	1	.643	.819
	IO*Medienglaubwürdigkeit ²	-.646	.325	4.938	1	.026	.524
	IO*Autorenglaubwürdigkeit	-.541	.389	2.007	1	.157	.582
	IO*Autorenglaubwürdigkeit ²	.033	.305	.012	1	.914	1.033
	Konstante	1.508	2.125	0.523	1	.469	4.517

Es zeigt sich, dass im finalen Modell die Domäne Lehramt und das diverse Geschlecht einen signifikanten Einfluss haben. Für erstere Variable bedeutet es mit den Werten $\chi^2(1) = 9.710$, $p = .002$; $OR = 9.101$, dass angehende Lehrkräfte eine um etwa 9-fache Chance haben im Vergleich zur Referenzgruppe der Mediziner, dass sie nachrecherchieren. Personen, die sich als

divers identifizieren, haben im Vergleich zur Referenzgruppe der Männer eine verringerte Chance, dass sie nachrecherchieren.

Es ergeben sich keinerlei direkten Effekte für IO sowie die einzelnen non-linearen Glaubwürdigkeitsindices. Die Ergebnisse sprechen nicht für die entsprechenden Hypothesen H3b und H4a der direkten Effekte.

Es könnte allerdings ein Moderationseffekt vorliegen, denn sowohl der Interaktionsterm mit IO und dem non-linearen, textlichen Glaubwürdigkeitsindex ($\chi^2(1) = 8.160, p = .004; OR = 2.656$) als auch der Interaktionsterm mit IO und dem non-linearen, mediumsbezogenen Glaubwürdigkeitsindex ($\chi^2(1) = 4.938, p = .026; OR = 0.524$) sind signifikant. Für die Interpretation und mögliche Bestätigung eines Moderationseffektes wird sich aufgrund der Komplexität auf die Visualisierung der Effekte in R gestützt (Abbildung 20). Bezüglich der Autorenmerkmale werden keine Effekte gefunden.

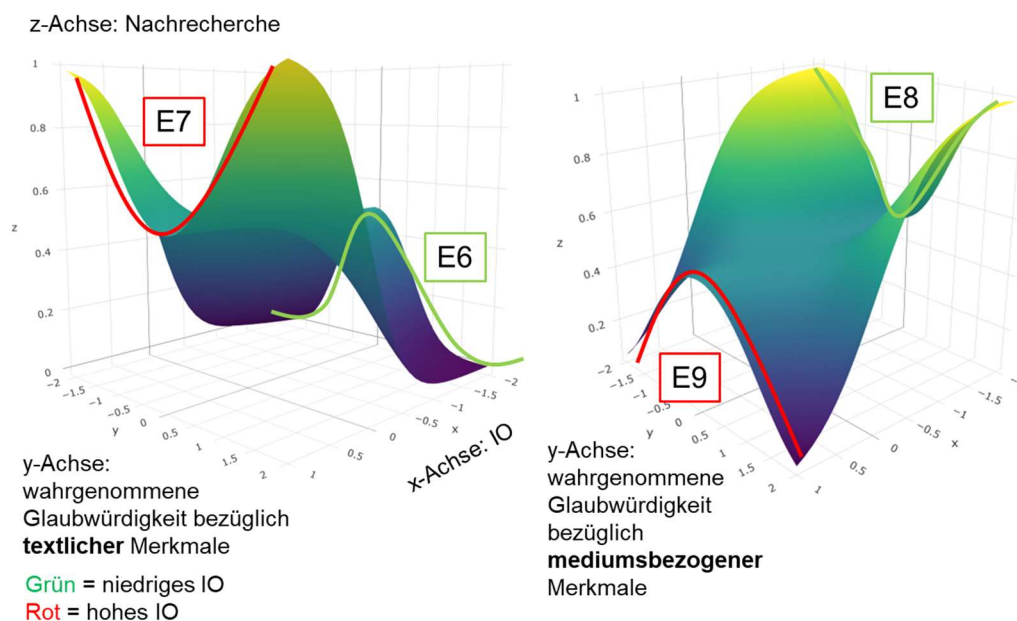


Abbildung 20: Effekte aus Modell II-B.

Erkenntnisse 6 und 7. Ähnlich wie im vorherigen Modell liegt bei beiden Gruppen (hoch/niedrig) ein non-linearer Effekt bezüglich textlicher Glaubwürdigkeitsmerkmale vor. Es wurde postuliert, dass Personen mit niedrigem IO nur dann nachrecherchieren, wenn sie Webseiten mit unklarer Glaubwürdigkeitsbeurteilung aufrufen. Dies könnte gemäß der Visualisierung (grüne Linie) zu erkennen sein (E6), weshalb die Ergebnisse somit für die entsprechende Hypothese H4a teilweise sprechen.

Es zeigt sich weiterhin, dass entgegen der postulierten Moderationshypothese H3b auch ein non-linearer Effekt für Probanden mit hohem IO (rote Linie) vorliegt (E7). Sie recherchieren

beim Aufrufen von sehr glaubwürdigen und sehr unglaubwürdigen Webseiten bzgl. textlicher Merkmale nach, weshalb dieses Ergebnis gegen die ursprüngliche Hypothese spricht.

Erkenntnisse 8 und 9. Für hohes als auch niedriges IO liegt ein non-linearer Interaktionseffekt mit den mediumsbezogenen Merkmalen vor. Probanden mit niedrigem IO recherchieren nur dann nach, wenn sie sehr glaubwürdige oder sehr unglaubwürdige Webseiten bzgl. der Medienmerkmale besuchen (E8). Dies zeigt zwar einen non-linearen Effekt, jedoch genau umgekehrt als angenommen, weshalb das Ergebnis nicht für H4a spricht. Probanden mit hohem IO recherchieren nur dann nach, wenn sie Webseiten besuchen, die bezüglich mediumsbezogener Merkmale eine unklare oder mittelmäßige Glaubwürdigkeit aufweisen (E9). Dies ist ebenfalls gegen die Moderationshypothese, da diese von einem linearen Effekt ausgeht. Das Ergebnis spricht nicht für die Hypothese H3b.

Zusammenfassend ist zu sagen, dass kein direkter Effekt von IO (H3b) und kein non-linearer direkter Effekt der Glaubwürdigkeitsmerkmale (H4a) nachgewiesen wird. Bezüglich des Moderationseffektes von IO auf den Effekt der verschiedenen Merkmalsblöcke auf die Nachrecherche kann statuiert werden, dass H4a nur teilweise für niedriges IO und die textliche Glaubwürdigkeit angenommen werden kann. Alle übrigen Hypothesen können nicht angenommen werden, was in nachstehender Tabelle 27 nochmals übersichtlich dargestellt ist:

Tabelle 27: *Hypothesenentscheidung Modell II-B.*

Direkte Effekte auf Nachrecherche	Status	Interaktionseffekte mit IO auf Nachrecherche	Status
IO (H3b)	✗	-	
Textglaubwürdigkeit	-	Hohes IO: H3b (E7)	✓
Textglaubwürdigkeit ² (H4a)	✗	Niedriges IO: H4a (E6)	✗
Medienglaubwürdigkeit	-	Hohes IO: H3b (E9)	✗
Medienglaubwürdigkeit ² (H4a)	✗	Niedriges IO: H4a (E8)	✗
Autorenglaubwürdigkeit	-	Hohes IO: H3b	✗
Autorenglaubwürdigkeit ² (H4a)	✗	Niedriges IO: H4a	✗

Anmerkung.

Legende: ✗ = Ergebnis spricht gegen Hypothese; ✓ = Ergebnis spricht für Hypothese; - = wurde nicht postuliert

Überprüfung der modellspezifischen Prämissen. Da das Modell signifikant ist, wird an der vorausgesetzten Prämisse der Linearität weiterhin festgehalten. Es existierten im entsprechenden Output zur Analyse eines möglichen Multikollinearitätsproblems keine *VIF*-Werte > 10, was somit gegen Multikollinearität spricht. Die Analyse der Ausreißer ergab für dieses Modell sieben Fälle, deren *SResid* > ± 2 *SD* waren. Diese wurden unter Berücksichtigung der Cook's Distanz betrachtet. Das Ergebnis der Einzelfallbetrachtung lieferte zwei mögliche einflussreiche Ausreißer. Diese wurden testweise ausgeschlossen und das Modell neu berechnet, was eine signifikante Modellgüte und einen weiteren signifikante Prädiktor mit gleichzeitigem Verlust eines vorherigen signifikanten Prädiktors aufwies. Dennoch wurde sich dafür entschieden, die zwei Probanden nicht auszuschließen. Denn diese sind einzig für dieses Modell als

Ausreißer klassifiziert, nicht jedoch in weiteren Modellen. Das Löschen der Ausreißer aus diesem Modell hätte zur Folge, dass diese generell aus der Stichprobe entfernt werden müssten und dies könnte wiederum Auswirkungen auf die anderen gefundenen Effekte haben. Da die zuvor getätigte Ausreißer-Analyse ergab, dass alle Probanden eine Daseinsberechtigung in der Stichprobe haben, müssen die hier gefundenen Ausreißer somit weiter inkludiert bleiben. Ebenso erzielte sich zwar ein neuer Prädiktor, gleichzeitig wurde ein zuvor signifikanter Prädiktor jedoch insignifikant.

Modell II-C. Für Modell II-C wird mit den soziodemografischen Daten, NfCC, den Glaubwürdigkeitsindices sowie ihren Interaktionstermen eine schrittweise logistische Regression berechnet, die einen möglichen Einfluss auf die abhängige Variable, Nachrecherche, ermitteln könnte. Hierzu wird die Stichprobe zunächst auf diejenigen Probanden gefiltert, welche auch Werte für NfCC aufweisen ($n = 64$).

Die Modellgüte für das finale Modell ist mit $\chi^2(18) = 20.143$, $p = .325$ nicht signifikant (Tabelle 28), weshalb keine weiteren Prüfungen zum Modell getätigt werden. Dennoch werden die Regressionskoeffizienten aus dem finalen Modell Nr. 3 auf deren möglichen signifikanten Einfluss mithilfe des Likelihood-Quotienten-Tests überprüft.

Tabelle 28: *Modellgüte von Modell II-C NfCC und Nachrecherche.*

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(5) = 3.898$, $p = .564$	.089
2	<u>Block:</u> $\chi^2(7) = 4.449$, $p = .727$ <u>Modell:</u> $\chi^2(12) = 8.347$, $p = .757$	.184
3	<u>Block:</u> $\chi^2(6) = 11.797$, $p = .067$ <u>Modell:</u> $\chi^2(18) = 20.143$, $p = .325$	.405

Tabelle 29 berichtet das Ergebnis der Prädiktorenüberprüfung:

Tabelle 29: *Ausschnitt der Ergebnisse der Berechnung der logistischen Regression von Modell II-C NfCC und Nachrecherche.*

Logistische Regression mit AV Nachrecherche und UV NfCC							
Mo- delln r.	Variable	Regres- sionsko- effi- zientB	Stan- dardfeh- ler	Likelihood-Quotienten- Test		Signif- ikanz	Exp(B)
				Chi²	df		
3	Alter	.104	.137	.575	1	.448	1.109
	HZB-Note	-.443	.962	.214	1	.643	.642
	Weiblich	.942	1.090	.789	1	.374	2.564
	Lehramt	1.245	1.516	.673	1	.412	3.473
	Jura	.186	1.346	.019	1	.890	1.204
	NfCC	-.262	.853	.095	1	.757	.769
	Textglaubwürdigkeit	-.159	1.102	.021	1	.885	.853
	Textglaubwürdigkeit ²	1.642	.839	4.923	1	.026	5.164
	Medienglaubwürdigkeit	-.141	.866	.026	1	.872	.868
	Medienglaubwürdigkeit ²	-.669	.542	1.654	1	.198	.512
	Autorenglaubwürdigkeit	-.117	.809	.021	1	.885	.890
	Autorenglaubwürdigkeit ²	-.665	.850	.709	1	.400	.514
	NfCC*Textglaubwürdigkeit	-.386	1.249	.097	1	.755	.680
	NfCC*Textglaubwürdigkeit ²	1.915	1.341	5.063	1	.024	6.786
	NfCC*Medienglaubwürdigkeit	.584	1.033	.329	1	.566	1.793
	NfCC*Medienglaubwürdigkeit ²	-.513	.933	.301	1	.583	.598
	NfCC*Autorenglaubwürdigkeit	-1.093	1.291	.741	1	.389	.335
	NfCC*Autorenglaubwürdigkeit ²	-2.381	1.473	6.369	1	.012	.092
	Konstante	-4.446	3.818	1.396	1	.237	.012

Es finden sich keine signifikanten Effekte zu der Soziodemografie oder NfCC in der Analyse. Die Ergebnisse sprechen nicht für die Hypothese, dass Personen mit hohem NfCC weniger nachrecherchieren (H2b). Es befindet sich allerdings ein signifikanter, direkter Effekt für den non-linearen, textlichen Glaubwürdigkeitsindex wieder ($\chi^2(1) = 4.923$, $p = .026$; $OR = 5.164$). Dieser ist jedoch gegen die postulierte Hypothese des direkten Effektes, da der Regressionskoeffizient positiv und die Parabel nach oben geöffnet ist. In H4a wurde angenommen, dass der Aufruf von wahrscheinlich sehr glaubwürdigen und sehr unglaubwürdigen wahrgenommenen Webseiten bezüglich textlicher Merkmale zu weniger Nachrecherche führt und umgekehrt. Wie sich das Ergebnis darstellt, ist es allerdings genau umgekehrt der Fall, weshalb das Ergebnis nicht für den direkten Effekt aus H4a spricht.

Hinsichtlich eines möglichen Moderationseffektes sind zwei Interaktionsterme, der non-lineare textliche Glaubwürdigkeitsindex ($\chi^2(1) = 5.063$, $p = .024$; $OR = 6.786$) und der non-lineare

autorenbezogene Glaubwürdigkeitsindex ($\chi^2(1) = 6.369, p = .012; OR = 0.092$) signifikant. Für die Visualisierung mit R wird bezüglich ersterem auch der direkte Effekt mit integriert, was folgend begründet wird: Wie anhand des reduzierten Modells 2 zu erkennen ist, ist dort noch kein Effekt signifikant. Erst mit Hinzufügen der Interaktionsterme wird auch der direkte Effekt signifikant. Er scheint also im Zusammenhang mit der Interaktionsterm-Variable zu stehen, weshalb sie gemeinsam betrachtet werden sollten. Es muss jedoch angemerkt werden, dass die Interaktionsterme beim Plotten die direkten Effekte unterdrücken könnten. Allerdings, wie eben beschrieben, scheinen sie ohnehin im Zusammenspiel mit den Interaktionen relevant zu werden, weshalb der Plot als Ganzes interpretiert wird. Abbildung 21 zeigt die gefundenen Effekte. Zur Betrachtung und Interpretation anzumerken ist, dass R die Richtungen der x-Achsen der beiden Grafiken vertauscht hat.

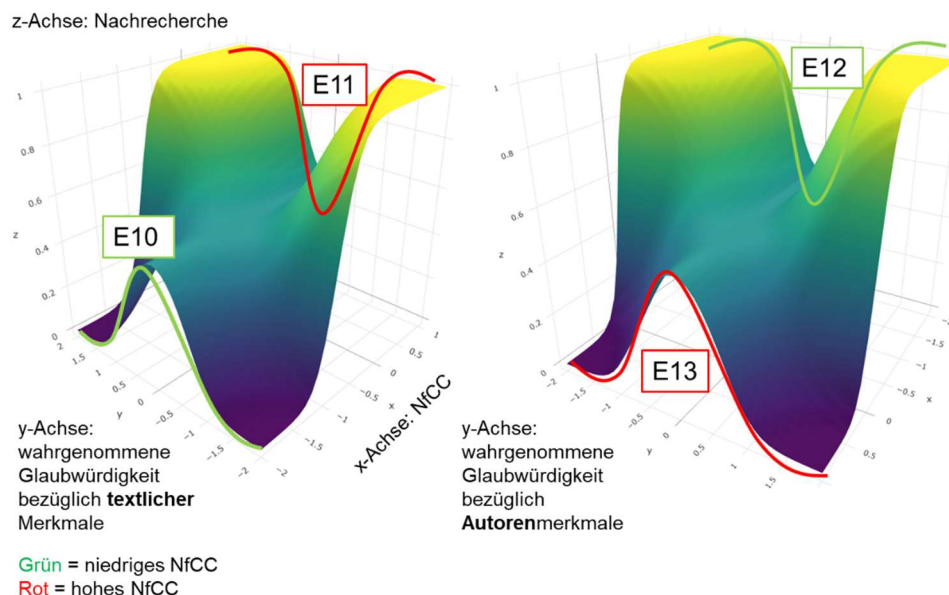


Abbildung 21: Effekte aus Modell II-C

Erkenntnisse 10 und 11. Probanden mit niedrigem NfCC (grüne Linie) recherchieren nach, wenn sie Webseiten aufrufen, die eine mittelmäßige oder unklare Glaubwürdigkeit bezüglich textlicher Merkmale aufweisen (E10). Das Ergebnis spricht für die postulierte Hypothese H4a. Probanden mit hohem NfCC (rote Linie) recherchieren dann nach, wenn sie sehr glaubwürdige oder sehr unglaubwürdige Webseiten bezüglich textlicher Merkmale aufrufen (E11). Dies spricht gegen die postulierte Moderationshypothese H2b, denn der gefundene Effekt ist nicht konstant bzw. linear, sondern non-linear.

Erkenntnisse 12 und 13. Probanden mit niedrigem NfCC (grüne Linie) recherchieren genau dann nach, wenn sie Webseiten aufrufen, die als sehr glaubwürdige oder sehr unglaubwürdige zu bewertende Autorenmerkmale bereitstellen (E12). Das bedeutet es wird nachrecherchiert, wenn sie durchschnittlich betrachtete Webseiten aufrufen, die entweder sehr viele oder sehr

wenige Autorennamen angeben. Dies ist zwar non-linear, aber entgegen der postulierten Hypothese H4a, weshalb das Ergebnis nicht für die Hypothese spricht. Probanden mit hohem NfCC (rote Linie) recherchieren dann nach, wenn sie Webseiten aufrufen, die bezüglich textlicher Merkmale als mittelmäßig oder unklar glaubwürdig beurteilt werden könnten (E13). Dies ist ebenfalls nicht entsprechend der Hypothese, da dort ein linearer Effekt bzw. parallel konstanter Effekt (sie machen überhaupt keine Nachrecherche) postuliert wurde. Das Ergebnis spricht daher nicht für H2b.

Zusammenfassend ist zu sagen, dass der direkte Effekt von NfCC (H2b) sowie der direkte non-lineare Effekt aller unterschiedlicher Webseitenmerkmalsblöcke (H4a) nicht angenommen werden kann. Es gibt zwar einen direkten, non-linearen Effekt mit den textlichen Webseitenmerkmalen, dieser ist jedoch in umgekehrter Richtung, weshalb dies nicht für die postulierte Hypothese spricht. Bezüglich des Moderationseffektes von NfCC auf den Effekt der verschiedenen Merkmalsblöcke auf die Nachrecherche kann statuiert werden, dass die Ergebnisse nur teilweise für H4a für niedriges NfCC und die textliche Glaubwürdigkeit sprechen. Alle restlichen Hypothesen können nicht angenommen werden, was in nachstehender Tabelle 30 nochmals übersichtlich dargestellt ist:

Tabelle 30: *Hypothesenentscheidung Modell II-C.*

Direkte Effekte auf Nachrecherche	Status	Interaktionseffekte mit NfCC auf Nachrecherche	Status
NfCC (H2b)	✗	-	
Textglaubwürdigkeit	-	Hohes NfCC: H2b (E11)	✗
Textglaubwürdigkeit ² (H4a)	✗	Niedriges NfCC: H4a (E10)	✓
Medienglaubwürdigkeit	-	Hohes NfCC: H2b	✗
Medienglaubwürdigkeit ² (H4a)	✗	Niedriges NfCC: H4a	✗
Autorenglaubwürdigkeit	-	Hohes NfCC: H2b (E13)	✗
Autorenglaubwürdigkeit ² (H4a)	✗	Niedriges NfCC H4a (E12)	✗

Anmerkung.

Legende: ✗ = Ergebnis spricht gegen Hypothese; ✓ = Ergebnis spricht für Hypothese; - = wurde nicht postuliert

Überprüfung der modellspezifischen Prämissen. Da das Modell unerwartet insignifikant ist und keinerlei Ergebnisse liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Ob Linearität gegeben ist, wurde durch Bildung der Interaktionsterme mit den logarithmierten metrischen unabhängigen Variablen ermittelt. Eine Interaktion war signifikant, jedoch war es lediglich Alter als soziodemografisches Merkmal, was somit wenig Aussagekraft für den Effekt der eigentlich untersuchten unabhängigen Variablen NfCC hat. Die Analyse des *VIFs* zeigte, dass keine Werte mit $VIF > 10$ in der Stichprobe enthalten waren. Die Analyse der Ausreißer ergab für dieses Modell sechs Fälle, deren $SResid > \pm 2 SD$ waren. Cook's Distanz identifizierte aus den sechs Ausreißern alle als einflussreich, jedoch wurde sich bereits ohne eine weitere testweise Analyse gegen den Ausschluss der Probanden entschieden. Das hat den Grund, dass mit sechs Probanden fast 10 % ($n = 64$) aus der

Modellberechnung entfernt worden wäre, was weitere Probleme durch abnehmende Stichprobengröße verursachen würde. Außerdem hätte der Ausschluss der Ausreißer einen Einfluss auf die anderen Modelle, weil sie auch von denen ausgeschlossen werden müssten. Diese decken sich aber nicht mit den bereits identifizierten Ausreißern aus den vorherigen Modellen, weshalb die Probanden in den anderen Modellen eine Daseinsberechtigung haben. Daher werden keine Ausreißer ausgeschlossen.

5.3.2.3. Modell III A-C mit Weiterverlinkung anklicken

Modell III-A. Für Modell III-A wird mit NfC eine schrittweise logistische Regression berechnet, die einen möglichen Einfluss auf die abhängige Variable, Anklicken der Weiterverlinkung, ermitteln könnte. Die Analyse zeigt, dass alle Modelle signifikant sind (Tabelle 31). Die anschließende Überprüfung der Koeffizienten zeigt, dass im reduzierten Modell 2 ein Parameter mehr, HZB-Note, signifikant ist. Da somit das große Modell 3 keinerlei neuen Erkenntnisse liefert und sich an dem Prinzip der Sparsamkeit orientiert wird, wird lediglich Modell 2 nachstehend weiterhin überprüft und berichtet.

Tabelle 31: Modellgüte von Modell III-A NfC und Weiterverlinkung anklicken

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(6) = 18,335, p = .005$	.186
2	<u>Block:</u> $\chi^2(7) = 21.333, p = .003$ <u>Modell:</u> $\chi^2(13) = 39.668, p < .001$	.374
3	<u>Block:</u> $\chi^2(6) = 3.821, p = .701$ <u>Modell:</u> $\chi^2(19) = 43.489, p < .001$	.404

Die Modellgüte für das finale, zweite Modell ist mit $\chi^2(13) = 39.668, p < .001$ signifikant. Auch das Nagelkerkes R^2 mit .374 zeigt einen kleinen Effekt (Backhaus et al., 2006).

Der Gesamtprozentsatz korrekter Klassifikation bzw. die Trefferquote ist 79.3 %. In 79.3 % der Fälle wird eine korrekte Prognose, d. h. die korrekte Zuteilung in die Gruppen (keine Weiterverlinkung anklicken vs. Weiterverlinkung anklicken) durch das Modell vorhergesagt. Die Sensitivität liegt bei 40.0 % und gibt den Anteil der korrekt vorhergesagten Personen, die Weiterverlinkung anklicken, an. Sie ist damit nicht sehr hoch, denn sie ist sogar weniger wahrscheinlicher als der Ausgang eines Zufallsexperimentes mit zwei Möglichkeiten (z. B. Münzwurf). Die Spezifität gibt den Anteil der korrekt vorhergesagten Personen an, die nicht die Weiterverlinkung anklicken, an. Diese ist mit 93.0 % sehr hoch. Das Modell sagt zu 40.0 % korrekt die Gruppenzugehörigkeit der Personen, die Links weiterverfolgen, und zu 93.0 % korrekt die Gruppenzugehörigkeit der Personen, die Links nicht anklicken, voraus. Aufgrund der

Problematik mit dem selbst gewählten Trennwert wird zudem die AUC betrachtet, welche für vorliegenden Fall .793 beträgt. Das Modell hat lt. Literatur also eine akzeptable Prognosefähigkeit (Backhaus et al., 2021, S. 311).

Eine Signifikanz der Variablen kann mithilfe des Likelihood-Ratio-Tests für Quotienten für die HZB-Note nachgewiesen werden ($\chi^2(1) = 5.070$, $OR = 0.278$; $p = .024$), sowie für die beiden direkten, linearen Effekte der textlichen ($\chi^2(1) = 4.961$, $OR = 3.175$; $p = .026$) und mediumsbezogenen ($\chi^2(1) = 6.839$, $OR = 2.643$ $p = .009$) Glaubwürdigkeit. Somit liegen lediglich lineare Zusammenhänge der Glaubwürdigkeitsindices und solche ohne Einfluss der Personenmerkmale vor (Tabelle 32).

Tabelle 32: Ausschnitt der Ergebnisse der schrittweisen Berechnung von Modell III-A NfC und Weiterverlinkung.

Logistische Regression mit AV Weiterverlinkung und UV NFC							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
2	Alter	.055	.073	.568	1	.451	1.057
	HZB-Note	-1.278	.609	5.070	1	.024	.278
	Weiblich	.309	.602	.267	1	.605	1.362
	Divers	-19.428	12508.4	2.475	1	.116	.000
			30				
	Lehramt	.252	.813	.096	1	.756	1.287
	Jura	.224	.702	.101	1	.750	1.251
	NfC	.020	.246	.006	1	.936	1.020
	Textglaubwürdigkeit	1.155	.575	4.961	1	.026	3.175
	Textglaubwürdigkeit ²	-.727	.486	2.557	1	.110	.484
	Medienglaubwürdigkeit	.972	.414	6.839	1	.009	2.643
	Medienglaubwürdigkeit ²	-.055	.371	.022	1	.881	.947
	Autorenglaubwürdigkeit	-.636	.383	2.864	1	.091	.530
	Autorenglaubwürdigkeit ²	-.079	.314	.064	1	.800	.924
	Konstante	-.538	2.014	0.071	1	.790	.584
	NfC*Autorenglaubwürdigkeit ²	.168	.431	.151	1	.697	1.182
	Konstante	-1.316	2.203	.357	1	0.550	.268

Das Ergebnis bezüglich der HZB-Note sagt aus, dass die Chance, die Weiterverlinkung anzuklicken, mit besserem Abiturschnitt steigt. Das ist insofern zu erklären, als dass bessere Abiturienten womöglich auch bessere Lerner sind, die die Glaubwürdigkeit erkennen und die Weiterverlinkung verfolgen und eine tiefere Suche anstreben. Dennoch wird das Ergebnis der Vollständigkeit wegen nur kurz erwähnt, jedoch nicht nochmal in der Diskussion aufgegriffen. Grund hierfür ist, dass das Ergebnis wenig Relevanz und Implikation für die Hochschulbildung hat.

Hinsichtlich der Merkmale bedeutet das Ergebnis, dass sich die Chance, die Weiterverlinkung anzuklicken, um circa 3.1 bzw. 2.6 erhöht, wenn sich die wahrgenommene Glaubwürdigkeit

bezüglich sowohl der textlichen Merkmale als auch der mediumsbezogenen Merkmale um eine Einheit erhöhen. Anders ausgedrückt bedeutet dies, dass ein Proband eher die Weiterverlinkung anklickt, wenn er Webseiten besucht hat, die mehr textliche bzw. mediumsbezogene Merkmale zur Glaubwürdigkeitswahrnehmung bereitstellten als welche, die weniger dieser Merkmale bereitstellten. Das Ergebnis spricht somit nicht für den non-linearen, direkten Effekt der Hypothese H4b, dass hohe und niedrige Glaubwürdigkeit zu weniger Anklicken der Weiterverlinkung führe (Tabelle 33).

Zudem gibt es keinen direkten Zusammenhang zwischen NfC und der Weiterverlinkung. Das Ergebnis spricht nicht für die Hypothese H1c. Hohes NfC führt nicht eher dazu, dass Probanden die Weiterverlinkung anklicken.

Tabelle 33: *Hypothesenentscheidung aus Modell III-A.*

Direkte Effekte auf Weiterverlinkung	Status	Interaktionseffekte mit NfC auf Weiterverlinkung	Status
NfC (H1c)	✗	-	
Textglaubwürdigkeit	-	Niedriges NfC: H1c	✗
Textglaubwürdigkeit ² (H4b)	✗	Hohes NfC: H4b	✗
Medienglaubwürdigkeit	-	Niedriges NfC: H1c	✗
Medienglaubwürdigkeit ² (H4b)	✗	Hohes NfC: H4b	✗
Autorenglaubwürdigkeit	-	Niedriges NfC: H1c	✗
Autorenglaubwürdigkeit ² (H4b)	✗	Hohes NfC: H4b	✗

Anmerkung.

Legende: ✗ = Ergebnis spricht gegen Hypothese; ✓ = Ergebnis spricht für Hypothese; - = wurde nicht postuliert

Überprüfung der modellspezifischen Prämissen. Da das Modell signifikante Ergebnisse liefert, wird weiterhin an der vorausgesetzten Prämisse der Linearität festgehalten und keine Überprüfung getätigt. Die Analyse von *VIF* zur Aufdeckung eines möglichen Multikollinearitätsproblems ergab kein $VIF > 10$, sodass die Prämisse erfüllt ist. Die Analyse der Ausreißer ergab, dass vier studentisierte Residuen mit $\pm 2 - 3 SD$ existierten. Einer davon wurden mithilfe von Cook's Distanz als möglicher einflussreicher Ausreißer identifiziert. Dieser wurde testweise aus dem Modell ausgeschlossen. Eine neue Berechnung erzielte zwar keine bessere Modellgüte, allerdings wurde der lineare autorenbezogene Glaubwürdigkeitsindex signifikant. Es wurde sich dennoch entschieden, das nicht modifizierte Modelle zu interpretieren, was folgende Gründe hat: Die identifizierten Ausreißer waren für dieses Modell als solche erkannt, haben allerdings in den anderen bereits gerechneten Modellen eine ausreichende Daseinsberechtigung, da sie dort nicht als Ausreißer identifiziert wurden. Ein Ausschluss würde Modelle und die bereits gefundenen Effekte zunehmend verändern. Weiterhin würde ein Ausschluss der Ausreißer die Daten lediglich an das Modell anpassen, was eine Modelloptimierung darstellt. Dies ist allerdings nicht das Ziel des Forschungsvorhabens, denn dieses soll konkrete Hypothesen mit der vorhandenen Stichprobe überprüfen und nicht Modelloptimierungen vornehmen. Daher werden die Ausreißer zwar kritisch gewürdigt, aber nicht weiter thematisiert.

Modell III-B. Für Modell III-B wird mit IO eine logistische Regression berechnet, die einen möglichen Einfluss auf die abhängige Variable, Weiterverlinkung anklicken, ermitteln könnte. Die Modellgüte für das dritte Modell ist mit $\chi^2(19) = 72.992$, $p < .001$ signifikant (Tabelle 34). Auch das Nagelkerkes R^2 mit 0.613 zeigt einen sehr guten, großen Effekt (Backhaus et al., 2006). Dieses wird daher nachstehend beschrieben.

Tabelle 34: Modellgüte des Modells III-B IO und Weiterverlinkung anklicken.

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(6) = 18.335$, $p = .005$	.186
2	<u>Block:</u> $\chi^2(7) = 23.268$, $p = .002$ <u>Modell:</u> $\chi^2(13) = 41.604$, $p < .001$	.389
3	<u>Block:</u> $\chi^2(6) = 31.389$, $p < .001$ <u>Modell:</u> $\chi^2(19) = 72.992$, $p < .001$	.613

Der Gesamtprozentsatz korrekter Klassifikation bzw. die Trefferquote ist 85.9 %. In 85.9 % der Fälle wird eine korrekte Prognose, d. h. die korrekte Zuteilung in die Gruppen (keine Weiterverlinkung anklicken vs. Weiterverlinkung anklicken) durch das Modell vorhergesagt. Die Sensitivität liegt bei 62.9 % und die Spezifität bei 94.0 %. das Modell sagt zu 62.9 % korrekt die Gruppenzugehörigkeit der Personen, die Links weiterverfolgen, und zu 94.0 % korrekt die Gruppenzugehörigkeit der Personen, die Links nicht, voraus. Die AUC beträgt .917 und gilt als exzellente Prognosefähigkeit (Backhaus et al., 2021, S. 311).

Tabelle 35 zeigt die detaillierte Koeffizientenstatistik.

Tabelle 35: Ausschnitt der Ergebnisse der schrittweisen Berechnung von Modell III-B IO und Weiterverlinkung.

Logistische Regression mit AV Weiterverlinkung und UV IO							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Sig.	Exp(B)
				Chi ²	df		
3	Alter	.071	.099	.497	1	.481	1.073
	HZB-Note	-2.048	.830	7.827	1	.005	.129
	Weiblich	-.180	.735	.060	1	.806	.835
	Divers	-20.543	11679.472	4.237	1	.040	.000
	Lehramt	1.748	1.112	2.649	1	.104	5.742
	Jura	-.157	.970	.026	1	.871	.854
	IO	-1.564	.727	6.003	1	.014	.209
	Textglaubwürdigkeit	2.110	.770	9.497	1	.002	8.250
	Textglaubwürdigkeit ²	-.838	.575	2.291	1	.130	.433
	Medienglaubwürdigkeit	1.689	.610	11.437	1	.001	5.413
	Medienglaubwürdigkeit ²	-.266	.557	.248	1	.619	.766
	Autorenglaubwürdigkeit	-1.765	.610	10.469	1	.001	.171
	Autorenglaubwürdigkeit ²	-.314	.411	.605	1	.437	.731
	IO*Textglaubwürdigkeit	2.645	1.050	8.572	1	.003	14.089
	IO*Textglaubwürdigkeit ²	2.267	.823	6.866	1	.009	9.647
	IO*Medienglaubwürdigkeit	-1.514	.696	5.261	1	.022	.220
	IO*Medienglaubwürdigkeit ²	.362	.600	.370	1	.543	1.436
	IO*Autorenglaubwürdigkeit	-.868	.814	1.200	1	.273	.420
	IO*Autorenglaubwürdigkeit ²	-1.771	.687	8.902	1	.003	.170
	Konstante	.165	2.706	0.004	1	.951	1.180

Aus soziodemografischer Betrachtungsweise ist lediglich die HZB-Note ($OR = 0.129$; $p = .003$) signifikant⁵⁸. Das Ergebnis sagt aus, dass mit jeder zusätzlichen Einheit der HZB-Note die Chance für die Links zu verfolgen um 87 % ($1 - 0.129$) reduziert wird. Anders ausgedrückt bedeutet dies, dass Personen mit einer besseren HZB-Note eher die Links weiterverfolgen als Personen mit einer schlechteren Note. Ein niedriger (und somit besserer) Durchschnittswert hinsichtlich der HZB-Note um eine Einheit erhöht also die Wahrscheinlichkeit, dass Links weiterverfolgt werden.

Es findet sich weiterhin ein direkter Effekt von IO ($OR = 0.209$, $p = .014$) sowie den linearen textlichen ($OR = 8.250$; $p = .002$), mediumsbezogenen ($OR = 5.413$; $p = .001$) und autorenbezogenen ($OR = 0.171$; $p = .001$) Glaubwürdigkeitsindices in den Daten. Werden lediglich die direkten Effekte betrachtet, wären folgende Rückschlüsse zu ziehen: Mit steigendem IO sinkt

⁵⁸ Obwohl das diverse Geschlecht eigentlich ebenfalls ein signifikantes Ergebnis ($OR = 0.000$; $p = .040$) vorweist, kann es dennoch nicht valide als solches bezeichnet werden. Denn in Anbetracht des sehr hohen Standardfehlers und der geringen Anzahl an Personen ($n = 9$), die sich als divers identifizieren, scheint die Berechnung des OR nicht valide zu sein. Daher wird dieses Ergebnis nicht näher thematisiert.

die Chance, dass nachrecherchiert wird. Dies entspräche der Hypothese H3c. Außerdem wird ungeachtet des IO umso mehr nachrecherchiert, wenn Webseiten aufgerufen werden, die eine hohe textliche und mediumsbezogene und eine niedrige autorenbezogene Glaubwürdigkeit aufweisen. Da es sich hierbei um lineare Effekte handelt, spräche das Ergebnis somit nicht für die korrespondierende Hypothese H4b mit dem postulierten non-linearen Effekt.

Bei den Aussagen ist jedoch Vorsicht geboten. Denn zum einen sind die direkten Effekte zwar auch schon im reduzierten Modell 2 (s. Anhang A 3) signifikant, verbessern sich jedoch mit Hinzufügen der Interaktionseffekte. Zum anderen besteht die Möglichkeit, dass die Interaktion disordinal vorliegt. Somit wären die direkten Haupteffekte nicht uneingeschränkt interpretierbar (Huber et al., 2014). Dies wäre jedoch erst durch die Visualisierung zu erkennen. Aus diesen Gründen werden die direkten Effekte zur Visualisierung mit R in die Gleichungen mit integriert, um eine holistische Aussage zu den einzelnen Merkmalsblöcken treffen zu können. Wichtig anzumerken ist jedoch, dass die Interaktionsterme beim Plotten die direkten Effekte unterdrücken könnten. Allerdings, wie eben beschrieben, scheinen sie ohnehin erst im Zusammenspiel mit den Interaktionen eine größere Rolle zu erhalten.

Folgende Interaktionsterme sind weiterhin signifikant und werden nacheinander ausführlich und in Abbildung 22 grafisch dargestellt: Interaktionsterm zwischen IO und dem linearen ($OR = 14.089$; $p = .003$) und non-linearen ($OR = 9.647$; $p = .009$) textlichen Glaubwürdigkeitsindex, zwischen IO und dem linearen mediumsbezogenen Glaubwürdigkeitsindex ($OR = 0.220$; $p = .022$) sowie zwischen IO und dem non-linearen autorenbezogenen Glaubwürdigkeitsindex ($OR = 0.170$; $p = .003$).

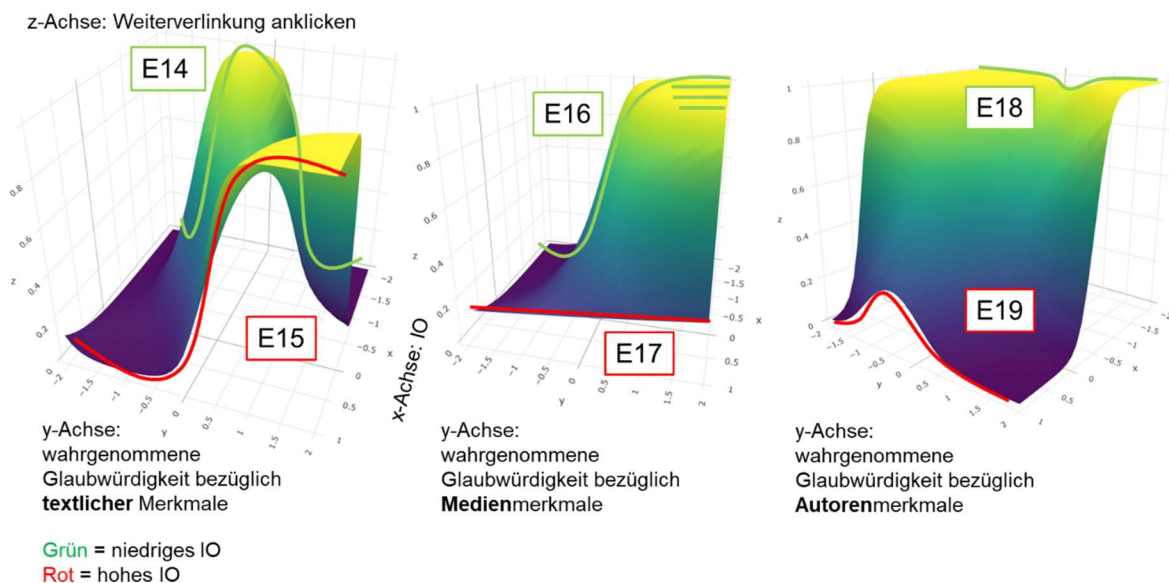


Abbildung 22: *Effekte aus Modell III-B.*

Erkenntnisse 14 und 15. Die Hypothese besagt, dass Personen mit niedrigem IO seltener weiterführende Links anklicken (H4b) bei sehr glaubwürdigen oder sehr unglaubwürdigen Webseiten (non-linearer Effekt) und, dass Personen mit hohem IO ungeachtet der Glaubwürdigkeit weniger Links anklicken (H3c) (linearer Effekt). Die Analyse zeigt: Personen mit niedrigem IO (grüne Linie) klicken dann Links an, wenn sie Webseiten besuchen, die unklare oder mittelmäßig textliche Merkmale für eine positive Glaubwürdigkeitsbeurteilung bereitstellen (E14). Der Kurvenverlauf könnte als Ergebnisbestätigung herangezogen werden. Das Ergebnis spricht somit für die Hypothese H4b und die textliche Glaubwürdigkeit. Probanden mit hohem IO (rote Linie) klicken Links an, wenn sie Webseiten besuchen, die viele textliche Glaubwürdigkeitsmerkmale bereitstellen und deshalb auch als glaubwürdiger wahrgenommen werden könnten (E15). Dies spricht nicht für die postulierte Hypothese. Denn in dieser wurde postuliert, dass Probanden mit hohem IO ungeachtet der wahrgenommenen Glaubwürdigkeit weniger Links verfolgen.

Erkenntnisse 16 und 17. Personen mit niedrigem IO klicken eher nur dann Links an, wenn sie Webseiten besuchen, die viele Medienmerkmale bereitstellen und somit als glaubwürdiger wahrgenommen werden könnten. Das Ergebnis spricht somit gegen die Hypothese H4b, da dies keinen non-linearen Effekt repräsentiert (E16). Bei Personen mit hohem IO spricht das Ergebnis tendenziell für die Hypothese H3c. Sie klicken ungeachtet der wahrgenommenen Glaubwürdigkeit weniger die Links an (E17).

Erkenntnisse 18 und 19. Bei Personen mit niedrigem IO spielt der Aufruf der glaubwürdigen oder unglaubwürdigen Webseiten bzgl. der Autorenmerkmale eher weniger eine Rolle (E18). Sie klicken die Weiterverlinkung auf einem sehr hohen Niveau und ungeachtet der Autorenmerkmale an. Das Ergebnis spricht zunächst gegen die Hypothese H4b. Jedoch besteht die leichte Tendenz, dass beim Aufruf von Webseiten mit unklaren Glaubwürdigkeitsmerkmalen weniger die Weiterverlinkung angeklickt wird. Dies ist jedoch nur leicht erkennbar, weshalb in Summe die Hypothese H4b eher abgelehnt werden sollte. Personen mit hohem IO verfolgen lediglich dann die Links, wenn sie Webseiten mit einer unklaren oder mittelmäßigen Glaubwürdigkeit bzgl. Autorenmerkmale besuchen (E19). Dies geschieht jedoch auf einem sehr geringen Niveau. Da sich daher in Summe kein linearer Effekt abzeichnet und die Probanden zumindest teilweise die Weiterverlinkung anklicken, spricht das Ergebnis auch nicht für die entsprechende Hypothese H3c.

Zusammenfassend ist somit festzuhalten, dass der direkte Effekt von IO entsprechend der Hypothese H3c zwar gefunden wird, sich aber verstärkt erst mit Hinzufügen der Interaktionsterme zeigt, weshalb diesem direkten Effekt nicht viel Wert beigemessen werden sollte. Es werden darüber hinaus direkte, aber lineare Effekte der Webseitenmerkmale gefunden. Diese

entsprechen nicht den postulierten, non-linearen Effekten aus den postulierten Hypothese H4b.

Bezugnehmend zu den Moderationseffekten ist jedoch zu sagen, dass die Ergebnisse zumindest teilweise für die Hypothesen der textlichen und mediumsbezogenen Merkmale sprechen: Probanden mit niedrigem IO klicken dann die Links an, wenn sie Webseiten mit einer mittelmäßigen oder unklaren Glaubwürdigkeit bzgl. textlicher Merkmale besuchen. Außerdem klicken Probanden mit hohem IO scheinbar ungeachtet der mediumsbezogenen Merkmale nicht die Links an. Alle restlichen Hypothesen können nicht angenommen werden, was in nachstehender Tabelle 36 nochmals übersichtlich dargestellt ist:

Tabelle 36: *Hypothesenentscheidung aus Modell III-B.*

Direkte Effekte auf Weiterverlinkung anklicken	Status	Interaktionseffekte mit IO auf Weiterverlinkung anklicken	Status
IO (H3c)	✓	-	
Textglaubwürdigkeit	-	Hohes IO: H3c (E15)	✗
Textglaubwürdigkeit ² (H4b)	✗	Niedriges IO: H4b (E14)	✓
Medienglaubwürdigkeit	-	Hohes IO: H3c (E17)	✓
Medienglaubwürdigkeit ² (H4b)	✗	Niedriges IO: H4b (E16)	✗
Autorenglaubwürdigkeit	-	Hohes IO: H3c (E19)	✗
Autorenglaubwürdigkeit ² (H4b)	✗	Niedriges IO: H4b (E18)	✗

Anmerkung.

Legende: ✗ = Ergebnis spricht gegen Hypothese; ✓ = Ergebnis spricht für Hypothese; - = wurde nicht postuliert

Überprüfung der modellspezifischen Prämissen. Die Überprüfung der Prämissen zeigte kein Multikollinearitätsproblem, da keine Werte mit $VIF > 10$. Die Analyse der Ausreißer ergab fünf Probanden mit $SResid > \pm 2-3 SD$, welche mithilfe von Cook's Distanz nochmal näher betrachtet wurden. Davon kristallisierten sich vier als mögliche einflussreiche Ausreißer heraus, welche jedoch nicht testweise ausgeschlossen wurden. Denn bei den Ausreißern handelte es sich erneut um bislang nicht entdeckte Ausreißer, welche somit nur individuell für das Modell als solche gelten würden. Wie bereits mehrfach erwähnt, ist die Anpassung der Daten an das Modell nicht das Ziel der Arbeit, weshalb die Ausreißer zwar identifiziert und kritisch gewürdigt werden, jedoch aber nicht ausgeschlossen werden.

Modell III-C. Für das Modell wird eine schrittweise logistische Regression berechnet, die einen möglichen Einfluss auf die abhängige Variable, Anklicken der Weiterverlinkung, ermitteln könnte. Keine der hierarchisch berechneten Modelle ist signifikant (Tabelle 37). Dennoch ist das reduzierte Modell mit einem p -Wert von .055 fast signifikant und somit vielversprechender. Im Hinblick des Sparsamkeitsprinzips wird daher nachstehend das Modell weiterhin auf signifikante Regressionskoeffizienten untersucht.

Tabelle 37: Modellgüte des Modells III-C NfCC und Weiterverlinkung anklicken.

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(5) = 5.820, p = .324$	.138
2	<u>Block:</u> $\chi^2(7) = 14.890, p = .037$ <u>Modell</u>	.437
3	$\chi^2(12) = 20.710, p = .055$ <u>Block:</u> $\chi^2(6) = 5.982, p = .425$ <u>Modell:</u> $\chi^2(18) = 26.692, p = .085$	.537

Der Likelihood-Ratio-Test ermittelt lediglich für den linearen autorenbezogenen ($OR = 0.173$; $p = .008$) und textlichen ($OR = 8.089$; $p = .047$) Glaubwürdigkeitsindex signifikante Ergebnisse (Tabelle 38). Das bedeutet, dass sich Chance um circa den Faktor 8 erhöht, die Weiterverlinkung anzuklicken, wenn sich die wahrgenommene Glaubwürdigkeit bezüglich textlicher Merkmale um eine Einheit erhöht. Anders ausgedrückt bedeutet dies, dass eher dann die Verlinkung angeklickt wird, wenn Webseiten besucht werden, die mehr textliche Merkmale zu einer positiven Glaubwürdigkeitswahrnehmung bereitstellen als welche, die weniger dieser Merkmale bereitstellen. Hinsichtlich der Autorenmerkmale bedeutet dies, dass umso mehr die Weiterverlinkungen angeklickt werden, je weniger Autorenmerkmale auf den Webseiten bereitgestellt werden.

Es wird kein direkter, signifikanter Zusammenhang mit NfCC oder den non-linearen Glaubwürdigkeitsmerkmalen und der abhängigen Variablen gefunden. Die Ergebnisse sprechen nicht für die Hypothesen H2c und H4b: Niedriges NfCC führt nicht dazu, dass die Probanden eher die Weiterverlinkungen anklicken. Der Aufruf von als sehr glaubwürdig oder sehr unglaubwürdig wahrgenommenen Webseiten bezüglich diverser Merkmale führt nicht dazu, dass weniger die Weiterverlinkungen angeklickt werden.

Tabelle 38: Ausschnitt der Ergebnisse der schrittweisen Modellberechnung III-C NfCC und Weiterverlinkung anklicken.

Logistische Regression mit AV Weiterverlinkung und UV NfCC							
Mo- dellnr.	Variable	Regressi- onskoeffi- zientB	Stan- dard- fehler	Likelihood-Quotienten- Test			
				Chi ²	df	Sig.	Exp(B)
2	Alter	.182	.184	1.100	1	.294	1.200
	HZB-Note	-.826	1.023	.689	1	.406	.438
	Weiblich	-2.010	1.107	3.691	1	.055	.134
	Lehramt	-.106	1.383	.006	1	.939	.899
	Jura	-.177	1.435	.015	1	.901	.837
	NfCC	.601	.477	1.728	1	.189	1.824
	Textglaubwürdigkeit	2.091	1.180	3.933	1	.047	8.089
	Textglaubwürdigkeit ²	-.788	.831	.953	1	.329	.455
	Medienglaubwürdigkeit	.522	.718	.556	1	.456	1.686
	Medienglaubwürdigkeit ²	-.959	.832	1.462	1	.227	.383
	Autorenglaubwürdigkeit	-1.756	.756	6.991	1	.008	.173
	Autorenglaubwürdigkeit ²	-.309	.480	.438	1	.508	.734
	Konstante	-2.522	4.369	0.341	1	.559	.080

Überprüfung der modellspezifischen Prämissen. Da das Modell unerwartet keinerlei Ergebnisse hinsichtlich NfCC liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Für vorliegendes Modell wurde die entsprechende Variablentransformation sowie neue Berechnung der logistischen Regression durchgeführt. Für die genannten unabhängigen Variablen und ihre Interaktionen lag kein signifikantes Ergebnis vor, weshalb die Prämisse der Linearität erfüllt ist. Auch die Prämisse der Abwesenheit der Multikollinearität ist erfüllt, da keine Werte mit $VIF > 10$ existieren. Die Analyse der Ausreißer ergab zwei $SResid > \pm 2 - 3 SD$. Die Cook's Distanz-Analyse zeigte, dass diese mögliche einflussreiche Ausreißer darstellen könnten, weshalb sie testweise aus der Analyse ausgeschlossen wurden. Der Ausschluss erzielte eine signifikante Modellgüte ($\chi^2(1) = 30.236$, $p = .003$), jedoch wurde sich auch hier gegen den Ausschluss der beiden Probanden entschieden, da diese erneut „neu“ sind und die Liste aller auszuschließenden Probanden weiter verlängern würde. Ein Ausschluss hätte eine Auswirkung auf die anderen Modelle.

5.3.3. Einfluss der Personen- und Webseitenmerkmale auf die abhängigen Variablen der Evaluationsstrategie

5.3.3.1. Modell IV A-B mit Korrektheit der Quelleneinschätzung

Modell IV-A. Für das Modell wird schrittweise zunächst der Einfluss von Alter, Geschlecht, HZB-Note, Domäne und weiterhin den unabhängigen Personenmerkmalen NfC und IO auf die Korrektheit der Quelleneinschätzung ermittelt. Hierzu wird sich einer multiplen Regression, welche auf der Methode der kleinsten Quadrate (KQ-Methode) basiert, bedient (Backhaus et al., 2021). Die Modellgüte wird dabei über einen F -Test sowie das korrigierte

Bestimmtheitsmaß R^2_{kor} bestimmt, was im Nachstehenden für dieses Modell einmal ausführlich theoretisch erläutert, für darauffolgende Modelle aber nur noch kurz berichtet wird.

Das Ergebnis des F -Tests wird in SPSS bei Durchführung der Regression automatisch in der Tabelle ANOVA ausgegeben. Der F -Test prüft die Nullhypothese, dass keine der angenommenen unabhängigen Variablen einen Einfluss auf die abhängige Variable in der Grundgesamtheit hat (Backhaus et al., 2021).

Zur Durchführung des F -Tests wird ein empirischer Wert der F -Statistik F_{emp} benötigt, welche über folgende Formel berechnet werden kann (Backhaus et al., 2021, S. 91):

$$F_{\text{emp}} = \frac{\text{erklärte Varianz}}{\text{nicht erklärte Varianz}} = \frac{MSE}{MSR}$$

Für das finale Modelle erhält man nach Einsetzen der entsprechenden Varianzen folgenden Wert für F_{emp} :

$$F_{\text{emp}} = \frac{2.016}{0.553} = 3.644$$

Unter Betrachtung der F -Verteilungstabelle mit $\alpha = 0.05$ erhält man für $df(1) = 8$ und $df(2) = 134$ den kritischen F -Wert 2.008. Da $F_{\text{emp}} = 3.644 > F_{\text{krit}} = 2.008$ ist, kann die Nullhypothese nicht angenommen werden. Es liegt keine Gleichheit vor.

Dies ist auch so aus dem ANOVA-Output erkennbar und wird für nachfolgende Modelle direkt abgelesen werden (Tabelle 39). Es zeigt sich für den Fall also, dass das finale Modell mit $F(8,134) = 3.644$, $p < .001$ die Korrektheit der Quelleneinschätzung statistisch signifikant voraussagt, sodass eine ausreichende Güte vorliegt.

Das korrigierte Bestimmtheitsmaß R^2_{kor} wird zur Beurteilung der Modellgüte herangezogen, da es, anders als das R^2 , die Anzahl der Beobachtungen und unabhängigen Variablen berücksichtigt (Backhaus et al., 2021, S. 92 ff.). Als Folge daraus können validere Aussagen zur Modellgüte getroffen werden. Die Varianzaufklärung ist sowohl beim reduzierten Modell mit $R^2_{\text{kor}} = .13$ als auch beim finalen Modell mit $R^2_{\text{kor}} = .13$ nach Cohen (1988, S. 413 f.)⁵⁹ zufolge eine mittlere Varianzaufklärung. 13.0 % der Varianz werden somit durch die Modelle erklärt.

⁵⁹ Einteilung: R^2 hat eine kleine Effektstärke mit .02, eine mittlere Effektstärke mit .13 und eine große Effektstärke mit .50.

Tabelle 39: Modellgüte von Modell IV-A.

Modell Nr.	F-Test	R^2_{kor}
1	$F(6,136) = 4.540, p < .001$	.130
2	$F(8,138) = 3.644, p < .001$	.130

Im nächsten Schritt werden die geschätzten Regressionskoeffizienten auf ihre Signifikanz und Präzision überprüft (Backhaus et al., 2021). Ob ein Regressionskoeffizient signifikant ist, kann mithilfe des t -Tests im Output von SPSS identifiziert werden (Backhaus et al., 2021). Es wird einmal ausführlich eine manuelle Berechnung und der damit verbundenen theoretischen Hintergründe dargestellt. Für die Überprüfung der Signifikanz der Regressionskoeffizienten wird sich der t -Statistik bedient, welche der t -Verteilung folgt. Hierzu wird im ersten Schritt der empirische t -Wert t_{emp} durch Division des Regressionskoeffizienten b_j durch den Standardfehler $SE(b_j)$ ermittelt (Backhaus et al., 2021, S. 97):

$$t_{\text{emp}} = \frac{b_j}{SE(b_j)}$$

Eine exemplarische Berechnung für den Prädiktor „Alter“ ergibt folgenden empirischen t -Wert:

$$t_{\text{emp}} = \frac{-0.081}{0.021} = \text{ca. } -3.786$$

Der t -Test ist in diesem Fall ein zweiseitiger Test. Die t -Statistik mit $df(2) = 134$ und $\alpha = 0.05$ liefert den kritischen t -Wert: $\pm 1.978^{60}$. Der empirische t -Wert für Alter liegt mit -3.79 außerhalb des Bereiches $+1.978$ und -1.978 , weshalb dieser bei $\alpha = 0.05$ als einziger Prädiktor statistisch signifikant ist ($\beta = -.081$; $p \leq .001$) (Tabelle 40). Ob der gefundene Effekt auch eine gute Präzision aufweist, kann mithilfe des Standardfehlers SE interpretiert werden. Der wahre Effekt des Parameters liegt innerhalb des Intervalls $-.081 \pm 0.042$ mit der Wahrscheinlichkeit von 95 %. Der gefundene Effekt bedeutet somit, dass jüngere Probanden signifikant besser bei der Einschätzung der Quellenglaubwürdigkeit abschneiden. Für jedes Jahr mehr „Alter“, sprich für ein Jahr älter, sinkt die Korrektheit der Quelleneinschätzung durchschnittlich um 0.081. Das Ergebnis zum Alter wird der Vollständigkeit wegen nur kurz erwähnt und nicht nochmal in der Diskussion aufgegriffen. Grund hierfür ist, dass das Ergebnis in diesem Forschungsvorhaben wenig Relevanz und Implikation für die Bildungsforschung hat.

Da darüber hinaus keine weiteren signifikanten Ergebnisse gefunden werden, ist zu schlussfolgern, dass die Ergebnisse nicht für die Hypothesen H5a und H7a sprechen. Hohes NfC und

⁶⁰ Mithilfe von Excel und der Funktion $T.INV.2S(0,05;134)$ berechnet.

niedriges IO führt nicht zu einem besseren Abschneiden in der korrekten Quelleneinschätzung, d. h. zur korrekten Identifikation glaubwürdiger bzw. unglaubwürdiger Quellen.

Tabelle 40: *Ergebnisse der linearen Regression vom finalen Modell IV-A NfC und IO auf Korrektheit der Quelleneinschätzung.*

Lineare Regression mit AV Korrektheit der Quelleneinschätzung und UV NfC und IO					
	Regressionsko- effizient	SE	Beta	T	Sig.
Konstante	5.251	.572		9.179	<.001
Alter	-.081	.021	-.356	-3.786	<.001
HZB-Note	.023	.124	.019	.185	.853
Weiblich	-.133	.148	-.079	-.894	.373
Divers	.067	.287	.020	.233	.816
Lehramt	-.195	.180	-.114	-1.086	.280
Jura	-.025	.191	-.013	-.133	.894
NfC	-.070	.065	-.088	-1.072	.286
IO	-.066	.064	-.081	-1.022	.309

a. Abhängige Variable: Korrektheit Quelleneinschätzung

Überprüfung der modellspezifischen Prämissen. Die Parameter im Modell müssen linear sein (Backhaus et al., 2021). Bei der hier vorliegenden multiplen Regression kann zur Erkennung von Nicht-Linearität die abhängige Variable gegen jede unabhängige Variable aufgetragen werden (Backhaus et al., 2021). Die Streudiagramme der jeweiligen unabhängigen Variablen NfC und IO zeigten, dass Linearität vorliegt. Weiterhin darf kein Underfitting vorliegen. Es dürfen also keine relevanten Variablen ausgelassen worden sein, sonst wäre die Schätzung verzerrt (Backhaus et al., 2021). Zur Inspektion dieses möglichen Problems wird der Tukey-Anscombe-Plot verwendet, bei dem die Fehlerterme gegen die geschätzten y-Werte auf der x-Achse abgetragen werden (Backhaus et al., 2021). Wünschenswert ist die zufällige und gleichmäßige Streuung (ohne systematische Muster) um die x-Achse, was in diesem Modell nicht der Fall ist. Die Residuen bilden ein Muster, was somit für die Verletzung der Prämisse spricht. Weiterhin müssen die Fehlerterme normalverteilt sein (Backhaus et al., 2021). Zur Überprüfung dieser Prämisse werden die Residuen grafisch in einem Histogramm abgetragen. Idealerweise sind diese symmetrisch verteilt, da die Normalverteilung ebenfalls eine symmetrische Verteilung ist. Es bestünde weiterhin die Möglichkeit, spezialisierte Wahrscheinlichkeitsplots, wie den Q-Q-Plot oder P-P-Plot, zu verwenden. Dies wird allerdings bei kleinen Stichproben, d. h. ≤ 40 empfohlen, was in vorliegendem Modell nicht in Frage kommt, da $n = 143 > 40$. Die Inspektion des Histogramms der standardisierten Residuen zeigte eine leichte Tendenz, weshalb sich mit einer rechnerischen Analyse mittels des Shapiro-Wilks-Tests abgesichert wurde. Dieser ergab ein $p = .002$, sodass die Nullhypothese verworfen werden muss und somit keine Normalverteilung vorliegt. Die soeben thematisierten Fehlerterme müssen gemäß einer weiteren Prämisse eine konstante Varianz aufweisen, d.h. homoskedastisch sein (Backhaus et al., 2021). Hierfür wurde erneut eine visuelle Inspektion getätigt, indem die unstandardisierten, vorhergesagten Werte von Y gegen die unstandardisierten Residuen geplottet wurden.

Heteroskedastizität ist an einem dreieckigen oder trichterförmigen Muster zu erkennen. Da dies in manchen Fällen schwierig zu erkennen ist, kann zudem der White-Test (da nicht normalverteilt) zur Beurteilung der Homoskedastizität herangezogen werden. Wenn der Test signifikant ist, kann die Nullhypothese nicht angenommen werden und es liegt Heteroskedastizität vor. Dies ist bei vorliegendem Datensatz nicht der Fall ($p = .439$). Die Prämisse der Homoskedastizität ist nicht verletzt. Die Liste *Fallweise Diagnose* zeigte, dass keine $SResid > \pm 2 - 3 SD$ in den Daten vorlagen. Die Inspektion des Residuen-Plots zeigte dennoch zwei vom Gros der Probanden weit entfernte Residuen auf, welche deshalb testweise ausgeschlossen wurden. Der Ausschluss erzielte jedoch keine Verbesserung der Modellgüte, der Signifikanzen und damit zusammenhängend der Schlussfolgerungen, weshalb die Probanden weiterhin im Modell geführt wurden. Weiterhin wäre als Grund für den Nicht-Ausschluss zu nennen, dass sich die beiden Probanden nicht mit den bereits zuvor identifizierten Ausreißern überschneiden, sodass hieraus erneut Probleme für die anderen Modelle entstehen würden. Im Hinblick auf ein mögliches Multikollinearitätsproblem wurden die *VIF*-Werte begutachtet. Keine der Variablen im Modell wies ein $VIF > 10$ auf.

Modell IV-B. Für die schrittweise Berechnung von Modell IV-B wird der Einfluss von Alter, Geschlecht, HZB-Note, Domäne und im zweiten Schritt NfCC auf die Korrektheit der Quelleneinschätzung ermittelt. Das reduzierte Modell zeigt dabei nicht nur eine bessere Varianzaufklärung ($R^2_{\text{kor}} = .099$), sondern auch eine signifikante Modellgüte mit $F(5,58) = 2.378$, $p = .005$ (Tabelle 41). Das Hinzufügen des Parameters NfCC verschlechtert demnach das Modell. Dennoch befinden sich sowohl im reduzierten als auch im finalen Modell keine signifikanten Regressionskoeffizienten. Das Ergebnis spricht nicht für die entsprechende Hypothese H6a. Niedriges NfCC führt nicht zu einem besseren Abschneiden in der korrekten Quelleneinschätzung.

Tabelle 41: *Modellgüte des Modells IV-B.*

Modell Nr.	F-Test	R^2_{kor}
1	$F(5,58) = 2.378$, $p = .050$	.099
2	$F(6,57) = 1.951$, $p = .088$	.083

Überprüfung der modellspezifischen Prämissen. Im Sinne des Sparsamkeitsprinzips würde eigentlich das reduzierte Modell weiter berichtet und interpretiert werden. Da das reduzierte Modell lediglich die Kontrollvariablen enthält und laut Hypothese zumindest ein Effekt von NfCC denkbar ist, werden sich jedoch mithilfe der Prämissenüberprüfungen mögliche Erklärungen für die nicht signifikante Modellgüte im Gesamtmodell hergeleitet. Die nachstehende Prämissenüberprüfung bezieht sich daher auf das gesamte Modell mit NfCC. Das Streudiagramm von NfCC auf der x-Achse und der Korrektheit der Quelleneinschätzung auf der y-

Achse zeigte, dass Linearität vorliegt und somit die Prämisse nicht verletzt ist. Eine visuelle Inspektion der Residuen gegen die UV NfCC zeigte eine zufällige, gleichmäßige Streuung um die x-Achse, weshalb die Prämisse des Underfittings somit nicht verletzt ist. Eine vorhandene Normalverteilung der Residuen wurde durch Inspektion des Histogramms der Residuen und rechnerisch mithilfe des Shapiro-Wilk-Tests verifiziert, da $p = .318$ und somit $> .05$. Zur Überprüfung der Homoskedastizität wurde der Breusch-Pagan-Test (da normalverteilt) verwendet. Da der Test mit $p = .003$ signifikant ist, kann die Nullhypothese nicht angenommen werden und es liegt Heteroskedastizität vor. Die Prämisse der Homoskedastizität ist verletzt. Der SPSS-Output zeigt, dass zwei Ausreißer zur näheren Betrachtung mit $\pm 2 - 3 SD$ enthalten waren. Diese wurden testweise ausgeschlossen, jedoch erzielte sich hierdurch keine Verbesserung i. S. v. Signifikanz der Modellgüte, weshalb u. a. die Ausreißer weiterhin beibehalten wurden. Weitere Gründen für den Nicht-Ausschluss von Ausreißern wurde bereits in den vorherigen Modellen angeführt. Multikollinearität konnte ebenfalls ausgeschlossen werden, da die unabhängigen Variablen keine *VIF*-Werte > 10 aufwiesen.

5.3.3.2. Modell V A-B mit Systematische Begründung

Modell V-A. Für Modell V-A wird der Einfluss von Alter, Geschlecht, HZB-Note, Domäne und im zweiten Schritt NfC sowie IO auf die systematische Begründung der Quelleneinschätzung ermittelt. Modell 1 ist unter Einschluss der soziodemografischen Daten und Modell 2 unter Einschluss der unabhängigen Variablen. Sowohl das reduzierte als auch das finale Modell sind signifikant (Tabelle 42). Da das Nagelkerkes R^2 im finalen Modell etwas höher ausfällt ($R^2 = .174$) und somit einen größeren, fast akzeptablen Effekt hat im Vergleich zum reduzierten Modell, wird nachstehend das finale Modell weiter interpretiert.

Tabelle 42: *Modellgüte des Modells V-A NfC, IO und systematische Begründung.*

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(6) = 16.884, p = .010$	.165
2	<u>Block:</u> $\chi^2(2) = 0.946, p = .623$	.174
	<u>Modell:</u> $\chi^2(8) = 17.830, p = .023$	

Der Gesamtprozentsatz korrekter Klassifikation bzw. die Trefferquote ist 71.9 %. In 71.9 % der Fälle sagt das Modell eine korrekte Prognose, d. h. eine korrekte Zuteilung in die Gruppen (systematische vs. heuristische Begründung) voraus. Die Sensitivität liegt bei 89.2 % und die Spezifität bei 33.3 %. Das Modell sagt folglich zu 89.2 % korrekt die Gruppenzugehörigkeit der Personen, die eine systematische Begründung angeben, und zu 33.3 % korrekt die Gruppenzugehörigkeit der Personen, die eine heuristische Begründung angeben, voraus. Letztere

Prognosewahrscheinlichkeit ist allerdings kritisch zu betrachten, da diese nur sehr gering ist. Die Berechnung der AUC zeigt dennoch einen Wert von .720, welche als akzeptabel gilt.

Nachdem die Güte des Gesamtmodells überprüft und als zumindest akzeptabel testiert wurde, können anschließend die einzelnen geschätzten Koeffizienten beurteilt und überprüft werden.

Der Likelihood-Quotienten-Test zeigt, dass im finalen Modell lediglich die Domäne Lehramt einen signifikanten Einfluss hat (Tabelle 43). Für diese Variable bedeutet es mit den Werten $\chi^2(1) = 5.157$, $p = .024$; $OR = 0.284$, dass es bei angehenden Lehrkräften im Vergleich zur Referenzgruppe der Mediziner wahrscheinlicher ist, dass sie eher heuristische Begründungen zur Quellenglaubwürdigkeit angeben. Da somit weder NfC noch IO signifikant sind, spricht das Ergebnis nicht für die korrespondierenden Hypothesen H5b und H7b: Personen mit hohem NfC und niedrigem IO führen nicht eher Begründungen zur Quellenglaubwürdigkeit aus der systematischen Kategorie an.

Tabelle 43: Ausschnitt der Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-A NfC, IO und systematische Begründung.

Logistische Regression mit AV Systematische Begründung und UVs IO und NfC							
Mo- delln r.	Variable	Regressi- onskoeffi- zientB	Standard- fehler	Likelihood-Quotienten- Test		Signif- ikanz	Exp(B)
				Chi ²	df		
2	Alter	.105	.097	1.217	1	.270	1.111
	HZB-Note	-.295	.424	.487	1	.485	.744
	Weiblich	.465	.465	.988	1	.320	1.592
	Divers	-.505	.864	.343	1	.558	.604
	Lehramt	-1.274	.576	5.157	1	.023	.280
	Jura	.671	.682	1.017	1	.313	1.956
	NfC	.029	.205	.020	1	.887	1.030
	IO	.205	.212	0.946	1	.331	1.227
	Konstante	-1.446	2.449	.357	1	.550	.235

Überprüfung der modellspezifischen Prämissen. Da das Modell unerwartet keinerlei signifikante Ergebnisse hinsichtlich der UVs liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Für die genannten unabhängigen Variablen und ihre Interaktionen lag jedoch kein signifikantes Ergebnis vor, weshalb die Prämisse der Linearität erfüllt ist. Die Analyse von *VIF* zur Aufdeckung eines möglichen Multikollinearitätsproblems ergab keine $VIF > 10$, was somit gegen ein Multikollinearitätsproblem spricht. Die Analyse der Ausreißer SPSS ergab für dieses Modell einen Fall, dessen $SResid > \pm 2 SD$ war, jedoch nach Betrachtung von Cook's Distanz nicht als ein einflussreicher Ausreißer identifiziert werden konnte.

Modell V-B. Für Modell V-B wird der Einfluss von Alter, Geschlecht, HZB-Note, Domäne und NfCC auf die Kategorie (systematische oder heuristische) der Quelleneinschätzung ermittelt.

Modell 1 ist mit Einbezug der soziodemografischen Daten und Modell 2 mit Einbezug von NfCC. Sowohl das reduzierte als auch das finale Modell sind nicht signifikant (Tabelle 44).

Tabelle 44: *Modellgüte des Modells V-B NfCC und systematische Begründung*

Modell Nr.	Omnibus-Test der Modellkoeffizienten	Nagelkerkes R^2
1	$\chi^2(5) = 10.437, p = .064$	.216
2	Block: $\chi^2(1) = 0.042, p = .838$ Modell: $\chi^2(6) = 10.478, p = .106$	.217

Trotz der insignifikanten Modellgüte könnten vereinzelt Prädiktoren signifikant sein. Aus diesem Grund wird der Likelihood-Quotienten-Test herangezogen. Dieser zeigt, dass im finalen Modell lediglich die Domäne Lehramt einen signifikanten Einfluss hat (Tabelle 45). Angehende Lehrkräfte haben eine verringerte Chance ($OR = 0.190$, da < 1) im Vergleich zur Referenzgruppe der Mediziner, dass sie systematische Begründungen zur Quellenglaubwürdigkeit angeben. Da somit NfCC nicht signifikant ist, spricht das Ergebnis nicht für die korrespondierende Hypothese H6b. Personen mit niedrigem NfCC führen nicht eher Begründungen zur Quellenglaubwürdigkeit aus der systematischen Kategorie an.

Tabelle 45: *Ausschnitt der Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-B NfCC und systematische Begründung.*

Logistische Regression mit AV Systematische Begründung und UV NfCC							
Modellnr.	Variable	Regressionskoeffizient B	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
2	Alter	-.093	.153	.368	1	.544	.911
	HZB-Note	.775	.724	1.216	1	.270	2.170
	Weiblich	.723	.645	1.251	1	.263	2.062
	Lehramt	-1.659	.881	3.828	1	.050	.190
	Jura	.146	1.076	.019	1	.892	1.157
	NfCC	-.065	.316	.042	1	.838	.937
	Konstante	1.961	3.772	0.269	1	.604	7.105

Überprüfung der modellspezifischen Prämissen. Da das Modell insignifikant ist und keinerlei Ergebnisse hinsichtlich der UV liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Für die genannten unabhängigen Variablen und ihre Interaktionen lag lediglich ein signifikantes Ergebnis für Alter vor. Da Alter jedoch lediglich eine Kontrollvariable und nicht von Interesse ist, ist dies als nicht problematisch zu betrachten. Die Analyse von *VIF* zur Aufdeckung eines möglichen Multikollinearitätsproblems zeigte kein $VIF > 10$, was somit gegen Multikollinearität spricht. Die Analyse der Ausreißer ergab für dieses Modell zwei Fälle deren $SResid > \pm 2 SD$ war, jedoch nach Betrachtung von Cook's Distanz nicht als einflussreiche Ausreißer identifiziert werden konnten.

5.3.3.3. Modell VI A-C mit Argumentationsstruktur

Da es sich bei Modell VI um ein moderiertes Mediationsmodell mit einer dichotomen Variable als AV handelt, wird sich der Software PROCESS von Andrew Hayes (Hayes, 2022b) in SPSS bedient. Obwohl die gängigen Modelle von Hayes auf eine lineare Regression ausgelegt sind, erkennt das entsprechende Makro die dichotome Variable automatisch und passt die Berechnung dementsprechend an (Hayes & Matthes, 2009, S. 932). PROCESS erlaubt allerdings nur einen Prädiktor gleichzeitig, deshalb werden die Berechnungen separat für jedes Personenmerkmal (NfC, IO, NfCC) gerechnet. Das generelle Vorgehen hierfür ist identisch, weshalb es im Folgenden kurz erläutert und in Abbildung 23 grafisch dargestellt wird.

Gemäß den vier Schritten (Baron & Kenny, 1986) zur Erkennung einer Mediation werden die Zusammenhänge sukzessive in PROCESS überprüft: Zunächst muss festgestellt werden, ob ein direkter Effekt c vorliegt, der somit überhaupt mediiert werden kann (Schritt 1). Anschließend muss ein Zusammenhang zwischen X und dem Mediator M ermittelt werden, welcher im Pfad a ausgedrückt wird (Schritt 2). Weiterhin muss der Mediator mit der abhängigen Variablen Y über den Pfad b signifikant sein (Schritt 3). Zuletzt muss im gesamten Modell der Pfad c seine Signifikanz verlieren, sodass eine vollständige Mediation vorliegen kann. Ist der Pfad c weiterhin signifikant, ist von partieller Mediation die Rede. Aktuellere Forschende kritisieren die vier Schritte des Vorgehens jedoch. Die Signifikanz des Zusammenhangs c (Schritt 1) zwischen UV und AV wird zwar von Baron und Kenny (1986) noch als ein erster notwendiger Schritt zur Überprüfung einer Mediation genannt. Jüngere Forschungen (z. B. Rucker et al., 2011; Zhao et al., 2010) weisen aber darauf hin, dass sich allein auf c zu fokussieren die Forschung behindern kann. Sie empfehlen daher, trotz Abwesenheit eines direkten Effekts c weiterhin die indirekten Effekte zu untersuchen und nicht nach Versagen im Schritt 1 abzubrechen.

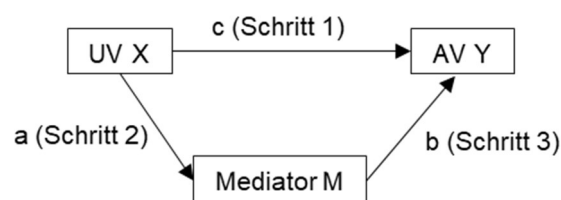


Abbildung 23: Allgemeines Mediationsmodell.

Zur Umsetzung wird in PROCESS die Modellspezifikation 4 ausgewählt, weil es sich um ein Mediationsmodell inklusive eines Moderators handelt, wobei der Moderator kein externer ist, sondern sich bereits als Personenmerkmal und somit unabhängige Variable im Modell befindet. Mit der Zusatzfunktion in PROCESS „Test for X by M interaction(s)“ erlaubt das System, dass X , in diesem Fall die unabhängige Variable und somit Personenmerkmal z. B. NfC,

zusätzlich auf eine Interaktion mit dem Mediator, der wahrgenommenen Textglaubwürdigkeit der selektierten Webseiten, getestet wird. Das repräsentiert die moderierte Mediation. Analog für NfC (Modell VI-A) werden nachstehend alle Modelle mit den jeweiligen Personenmerkmalen (IO, NfCC) sowie Soziodemografie (Alter, Geschlecht, HZB-Note, Domäne) berechnet und dargestellt.

Modell VI-A. In Schritt 1 wird überprüft, ob der direkte Effekt von NfC und Argumentationsstruktur (Pfad c) signifikant ist. Die korrespondierende Hypothese H5c postuliert, dass hohes NfC eher zu einer besseren Argumentationsstruktur führe. Dieser Pfad ist jedoch nicht signifikant (Abbildung 24). Das Ergebnis spricht nicht für die Hypothese H5c. Hohes NfC führt nicht eher zu einer guten Argumentationsstruktur. Der Prozess müsste nach dem ursprünglichen Vorgehen (Baron & Kenny, 1986, S. 1177) eigentlich abgebrochen werden. Wie bereits erläutert, wird dieser Empfehlung jedoch nicht gefolgt. Es zeigt sich, dass der Effekt a von NfC auf Textglaubwürdigkeit ($\beta = -.086$, $p = .291$) nicht signifikant ist. Das Ergebnis spricht somit nicht für die postulierte Hypothese H8: Personen mit hohem NfC selektieren nicht solche Webseiten, die bezüglich textlicher Merkmale eine höhere Glaubwürdigkeit aufweisen. Es ist jedoch ein Ergebnis zur Soziodemografie in Pfad a zu verzeichnen. Angehende Lehrkräfte selektieren im Vergleich zu der Referenzgruppe (angehende Mediziner) Webseiten, die hinsichtlich der textlichen Merkmale als eher unglaubwürdig wahrgenommen werden könnten ($\beta = -.539$, $p = .018$). Es zeigt sich außerdem, dass der Effekt b von Textglaubwürdigkeit auf die Argumentationsstruktur ebenfalls nicht signifikant ist ($OR = -0.191$, $p = .334$). Das bedeutet, dass die Selektion von Webseiten, die hinsichtlich textlicher Merkmale eine hohe Glaubwürdigkeit aufweisen bzw. suggerieren, nicht dazu führt, dass eine bessere Argumentationsstruktur vorliegen kann. Das Ergebnis spricht nicht für H11. Die Textglaubwürdigkeit als Mediator ist ebenfalls nicht signifikant ($OR = 0.016$), erkennbar an Konfidenzintervallen, welche 0 mit einbeziehen, d. h. hier im Fall 95 % KI [-0.046, 0.086]. Letztlich prüft die Zusatzfunktion „Likelihood ratio test of X by M interaction“, ob die moderierte Mediation signifikant ist. Diese ist ebenfalls nicht signifikant ($\chi^2(1) = 0.005$, $p = .945$). Das Ergebnis spricht gegen H12, dass in Abhängigkeit des NfC der Effekt der wahrgenommenen Glaubwürdigkeit der selektierten Webseiten auf die Argumentationsstruktur variieren würde.

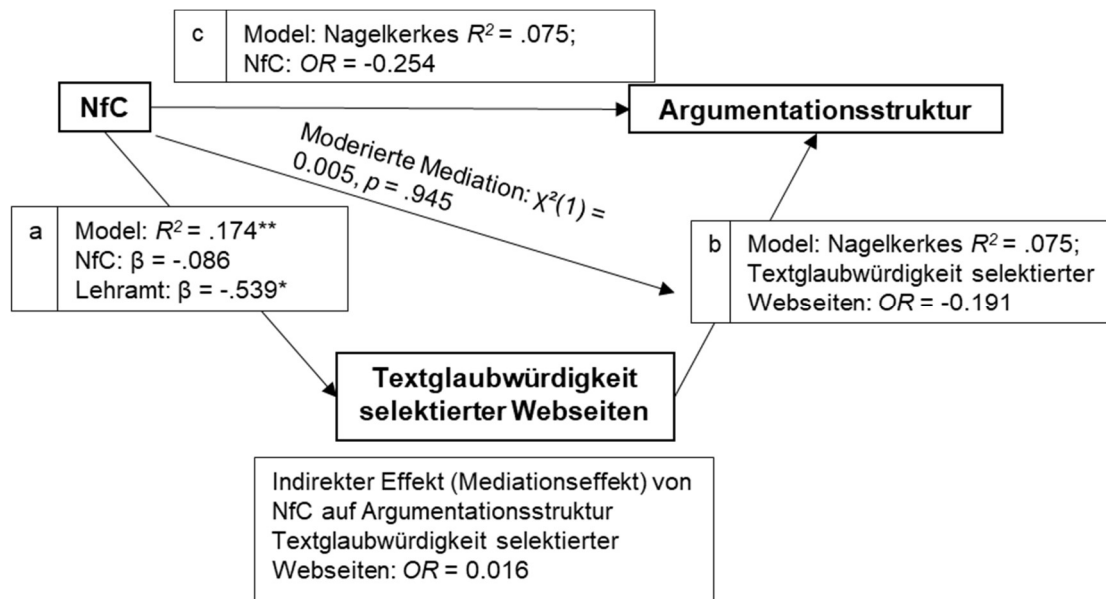


Abbildung 24: Mediationsmodell VI-A NfC und Argumentationsstruktur.

$**p < .01$; $*p < .05$

Überprüfung der modellspezifischen Prämissen. Da es sich bei diesem Modell trotz Mediation grundsätzlich um eine logistische Regression handelt, müssen auch die entsprechenden für eine Regression gültigen Prämissen überprüft werden: Linearität, keine Multikollinearität sowie keine Ausreißer. Da das Modell unerwartet insignifikant ist und keinerlei Ergebnisse liefert, wird die Linearität mithilfe der LN-Werte der metrischen Variablen und ihren Interaktionen überprüft. Für vorliegendes Modell wurde die entsprechende Variablentransformation ($\ln_Alter$; $\ln_NfC$; $\ln_HZB$; $\ln_PIXsText$) in SPSS durchgeführt. Die neue Berechnung wurde ebenfalls in SPSS über die Funktion der binomial logistischen Regression vorgenommen, da PROCESS lediglich für die Berechnung der Signifikanzen und Effekte herangezogen wurde. Es lag kein signifikantes Ergebnis für die Interaktionen mit den $\ln_Dummys$ vor, sodass die Prämisse erfüllt ist. Die Analyse von VIF zur Aufdeckung eines möglichen Multikollinearitätsproblems resultierte in keinen Werten mit einem $VIF > 10$, sodass die Prämisse als erfüllt gilt. Die Analyse der Ausreißer in SPSS ergab keinerlei $SResid > \pm 2 SD$.

Modell VI-B. In Schritt 1 wird überprüft, ob der direkte Effekt von IO und Argumentationsstruktur (Pfad c) signifikant ist (Abbildung 25). Dieser ist jedoch nicht signifikant, was gegen H7c spricht. Personen mit hohem IO haben eher keine schlechte Argumentationsstruktur. Auch die Effekte a von IO auf Textglaubwürdigkeit ($\beta = -.073, p = .360$) sowie b von Textglaubwürdigkeit auf Argumentationsstruktur sind ebenfalls nicht signifikant ($OR = -0.176, p = .373$). Das Ergebnis spricht gegen H10 und H11: Personen mit hohem IO selektieren nicht Webseiten, die bezüglich textlicher Merkmale eine geringere Glaubwürdigkeit aufweisen (H10). Und die Selektion von Webseiten mit hoher Textglaubwürdigkeit führt nicht dazu, dass eine bessere

Argumentationsstruktur vorliegen kann (H11). Folglich ist die Textglaubwürdigkeit als Mediator nicht signifikant ($OR = 0.013$) erkennbar am 95 % KI [-0.040, 0.083]. Analog zum Modell mit NfC ist jedoch die gleiche Erkenntnis bezüglich der Domäne festzustellen: Angehende Lehrkräfte selektieren im Vergleich zu der Referenzgruppe, angehenden Medizinerinnen, Webseiten, die hinsichtlich der textlichen Merkmale als eher unglaubwürdig wahrgenommen werden könnten ($\beta = -.543$, $p = .018$). Auch die moderierte Mediation ist nicht signifikant ($\chi^2(1) = 0.193$; $p = .661$). Das Ergebnis spricht nicht für H14, dass in Abhängigkeit von IO der Effekt der wahrgenommenen Glaubwürdigkeit der selektierten Webseite auf die Argumentationsstruktur variieren würde.

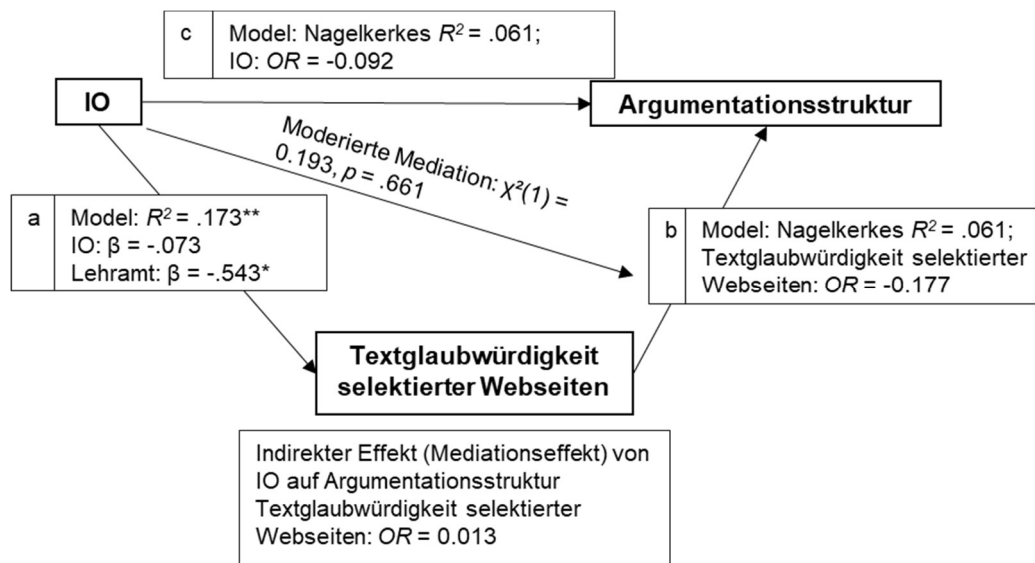


Abbildung 25: Mediationsmodell VI-B IO und Argumentationsstruktur.

** $p < .01$; * $p < .05$

Überprüfung der modellspezifischen Prämissen. Die Linearität zwischen den metrischen UVs und der Logit-Transformation der AV ist nicht vorhanden, da ein signifikantes Ergebnis für die Interaktionen zwischen den UVs und ihren LN-Werten vorlag. Die Prämisse ist somit nicht erfüllt. Die Analyse von *VIF* zur Aufdeckung eines möglichen Multikollinearitätsproblems ergab keine Variablen mit einem $VIF > 10$ entdeckte. Die Analyse der Ausreißer in SPSS ergab keine $SResid > \pm 2 SD$.

Modell VI-C. In Schritt 1 wird überprüft, ob der direkte Effekt von NfCC und Argumentationsstruktur (Pfad c) signifikant ist (Abbildung 26). Dieser ist jedoch nicht signifikant, was somit gegen H6c spricht. Personen mit hohem NfCC weisen nicht eine schlechtere Argumentationsstruktur auf. Weiterhin zeigt sich, dass die Effekte a von NfCC auf Textglaubwürdigkeit ($\beta = .169$, $p = .152$) sowie b von Textglaubwürdigkeit auf Argumentationsstruktur ebenfalls nicht signifikant sind ($OR = -0.196$, $p = .537$). Dies bedeutet, dass Personen mit hohem NfCC nicht Webseiten mit hoher Textglaubwürdigkeit selektieren. Das Ergebnis spricht nicht für H9. Das

gilt ebenso erneut für H11 in diesem Modell: Die Selektion von Webseiten mit hoher Textglaubwürdigkeit führt nicht dazu, dass eine bessere Argumentationsstruktur vorliegt. Folglich ist die Textglaubwürdigkeit als Mediator nicht signifikant ($OR = 0.033$) erkennbar an 95 % KI $[-0.338; 0.105]$. Auch in diesem Modell gibt es jedoch ein Ergebnis zur Soziodemografie: Angehende Lehrkräfte selektieren im Vergleich zu der Referenzgruppe, angehenden Medizinern, Webseiten, die hinsichtlich der textlichen Merkmale als eher unglaublich wahrgenommen werden könnten ($\beta = -.907$, $p = .011$). Die moderierte Mediation ist nicht signifikant ($\chi^2(1) = 0.042$; $p = .837$). Das Ergebnis spricht nicht für die Annahme von H13. In Abhängigkeit des NfCC variiert der Effekt der wahrgenommenen Glaubwürdigkeit der selektierten Webseiten auf die Argumentationsstruktur nicht.

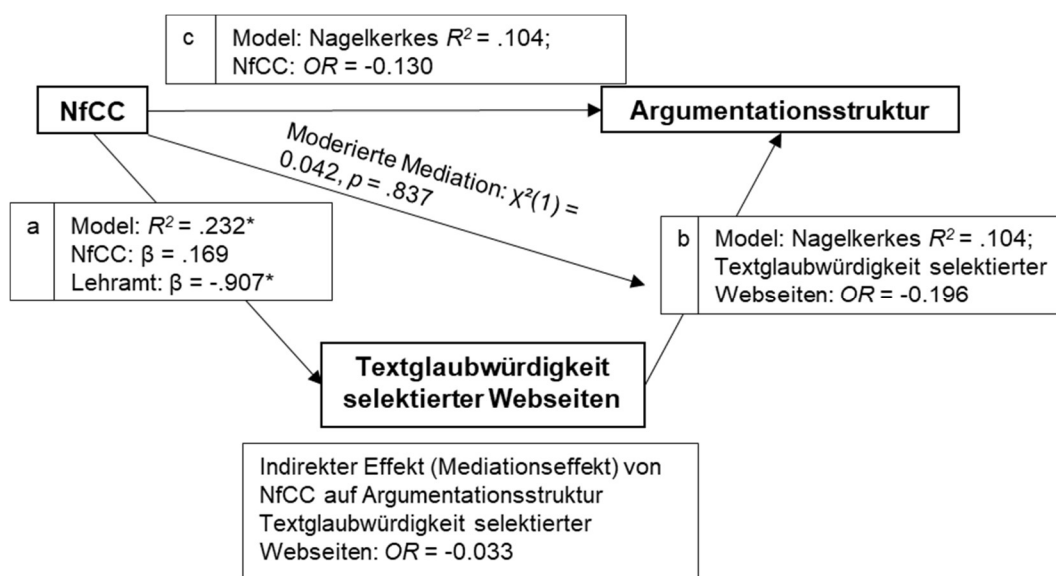


Abbildung 26: Mediationsmodell VI-C NfCC und Argumentationsstruktur.

** $p < .01$; * $p < .05$

Überprüfung der modellspezifischen Prämissen. Die Überprüfung zeigt, dass es lediglich einen Ausreißer mit $SResid > \pm 2 SD$ und Cook's Distanz > 1 gab. Aus diesem Grund wurde dieser testweise aus einer zweiten Analyse ausgeschlossen. Die neue Analyse erzielte keinerlei Verbesserungen i. S. v. signifikante Modellgüte oder Effekte. Aufgrund u. a. dieser Tatsache und weil der Ausschluss der Probanden weitere, bereits erläuterte Probleme mit sich ziehen würde, wurde sich gegen den Ausschluss entschieden.

6. Diskussion und Interpretation der Befunde

Wie sich anhand des vorherigen Ergebniskapitels zeigt, sind einige Ergebnisse, die für die Annahme von den Hypothesen sprechen, vorhanden. Daraus können modellübergreifende Erkenntnisse gewonnen werden, welche nachstehend thematisiert werden: Das erste Unterkapitel diskutiert die Ergebnisse der Regressionsberechnung. Das zweite Unterkapitel

diskutiert die deskriptiven sowie weitere domänenspezifische Ergebnisse und ordnet sie in den Forschungskontext ein.

6.1. Non-linearer Zusammenhang und Glaubwürdigkeitsbeurteilung

Zwei (II, III) der sechs Modelle zeigen Ergebnisse, welche teilweise für die postulierten Hypothesen sprechen: So gibt es beispielsweise einen Moderationseffekt mit hohem NfC und der Textglaubwürdigkeit bei der Nachrecherche im Modell II-A. Ähnlich verhält es sich mit den anderen Personen- und Webseitenmerkmalen im Modell III. Da die erneute Darstellung der Ergebnisse wiederholend und deren Diskussion zu detailliert wäre, werden die Ergebnisse konsolidiert betrachtet und lediglich anhand Modell III exemplarisch diskutiert. Denn sie haben folgende Schlussfolgerung gemein: Die unterschiedlichen Personenmerkmale haben entgegen den Hypothesen jeweils unterschiedliche Einflüsse (d. h. z. B. direkt, in Interaktion oder keinen) auf die Nachrecherche bzw. das Anklicken der Weiterverlinkung. Auch die jeweiligen Webseitenmerkmale beeinflussen auf verschiedene Weise (z. B. eher vermehrt Textglaubwürdigkeit, sehr wenig Einfluss von Autorenglaubwürdigkeit) das Verhalten, obwohl dies in den Hypothesen anders postuliert wurde. Zudem gibt es Moderationseffekte, die jedoch ebenfalls unterschiedlich je nach Personenmerkmal (z. B. nur bei IO und nicht bei NfCC) ausfallen und wiederum unterschiedlich im Effekt (non-linear vs. linear) der Webseitenmerkmale auf das Suchverhalten sind. Das heißt, dass sich mit zunehmenden Aufrufen von als glaubwürdig zu bewertenden Webseiten nicht immer ein linearer Effekt, d. h. mehr Nachrecherchieren oder mehr Klicken der Weiterverlinkung, einstellt. Vielmehr wird ersichtlich, dass die Probanden unter bestimmten Voraussetzungen (z. B. hohes NfC und Aufrufen textlicher Merkmale) trotz zunehmender Glaubwürdigkeit nicht das entsprechende (und eigentlich erwartete, lineare) Suchverhalten zeigen. Wie die Ergebnisse zeigen, ist die Betrachtung von Webseiten- und Personenmerkmalen innerhalb eines Modells komplexer als erwartet und erfordert ein Umdenken der gängigen linearen Zusammenhänge, wenn z. B. ein Prozess zur Lösung eines Informationsproblems erforscht werden soll. Folgende anschauliche Ergebnisdarstellung von Modell III verdeutlicht die Komplexität und Notwendigkeit zum Umdenken:

Modell III testet, ob höheres NfC, niedriges IO und niedriges NfCC zu mehr Linkverfolgung führt, ob ein non-linearer direkter Effekt der Webseitenmerkmale vorliegt und ob NfC, IO und NfCC jeweils den Effekt der wahrgenommenen Glaubwürdigkeit der besuchten Webseiten mit unterschiedlichen Merkmalen auf die Weiterverlinkung moderiert. Das heißt, Personen mit hohem NfC, niedrigem IO oder niedrigem NfCC klicken eher Links an, wenn sie sehr glaubwürdige oder sehr unglaubwürdige Webseiten besucht haben, wohingegen Personen mit niedrigem NfC, hohem IO und hohem NfCC generell weniger Links anklicken unabhängig von der Glaubwürdigkeit. Wie die Ergebnisdarstellung zeigt, sprechen die Ergebnisse nur teilweise für

die Hypothesen und hängen von der jeweiligen Betrachtung der Merkmalsblöcke ab. Die Annahme, dass die einzelnen Merkmale äquivalent zur Glaubwürdigkeitsbeurteilung beitragen, hat sich nicht bestätigt. Vielmehr zeigt sich, dass je nach Merkmalsblock unterschiedliche Verhaltensweise resultieren. Für einige der Personenmerkmale sind diese ähnlich, für andere wiederum nicht. Abbildung 27 zeigt die gefundenen Effekte übersichtlich in einer anderen (zweidimensionalen) Darstellungsweise je nach Webseitenmerkmalsblock (Horizontale) und Personenmerkmale (Vertikale) auf, um eine bessere Vergleichbarkeit zu ermöglichen. Die erneute kurze Betrachtung der Ergebnisse in veränderter Darstellung ist an dieser Stelle notwendig, um die anschließende Diskussion und Interpretation besser nachvollziehen zu können.

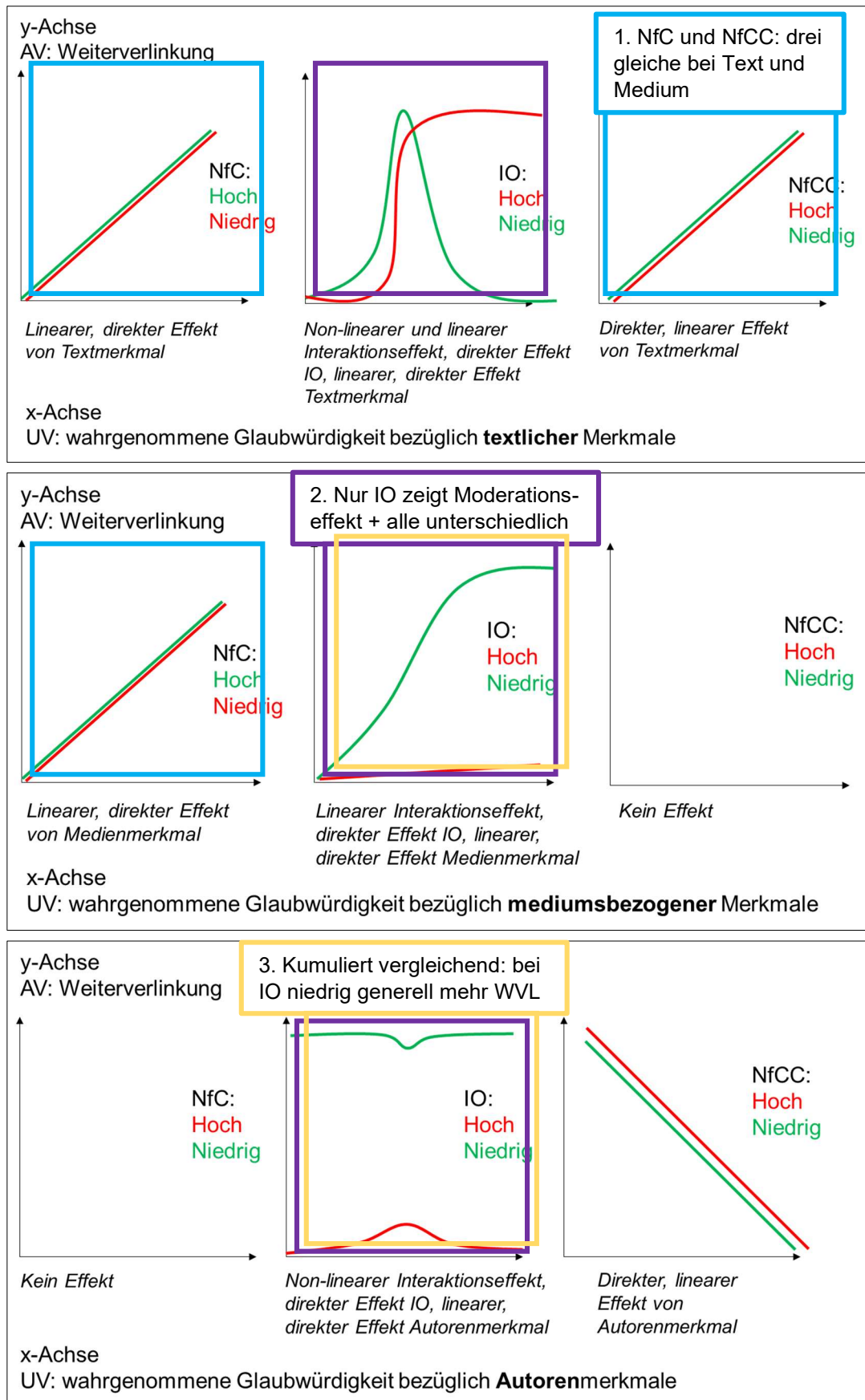


Abbildung 27: Gesamtbetrachtung der Ergebnisse aus Modell III.

Um eine ausreichende Basis zur Diskussion und Interpretation der Ergebnisse sicherzustellen, werden nachstehend zunächst Vorüberlegungen über die Intention zum Anklicken der Weiterverlinkungen angestellt:

Intention 1: allgemeine Strategie. Die Probanden verfolgen Links weiter, weil sie dies ohnehin, unabhängig ihrer Dispositionen und der (wahrgenommenen) Merkmale der besuchten Webseiten, getan hätten. Es ist ihre übliche Strategie. Wenn diese Begründung jedoch zutreffen würde, würden sich alle Probanden unabhängig ihrer Disposition und über alle Webseitenmerkmalsblöcke hinweg gleich verhalten. Dies ist nicht der Fall, weshalb die Begründung nicht anwendbar ist.

Intention 2: Überprüfung: Die gefundenen Informationen und die Quelle werden eingeschätzt und möglicherweise auch als glaubwürdig oder unglaubwürdig bewertet. Die Glaubwürdigkeit der Quelle bzw. der Informationen kann allerdings nicht mit abschließender Sicherheit beurteilt werden. Sie werden daher zurückverfolgt, um sie auf Glaubwürdigkeit oder weitere Qualitätsmerkmale zu überprüfen.

Intention 3: um weitere, glaubwürdige Informationen zu finden. Die gefundenen Informationen werden im ersten Schritt als glaubwürdig eingeschätzt. Unter der Annahme der glaubwürdigen Informationen und den damit verbundenen Links, werden diese in der Hoffnung angeklickt, mehr glaubwürdige Informationen zu diesem Thema finden zu können.

Intention 4: Effizienz/um Ressourcen zu sparen. Die gefundenen Informationen werden nicht ausgiebig oder gar nicht beurteilt. Die Linkverfolgung dient dazu Ressourcen zu sparen, dadurch effizient zu handeln und direkt unter Zuhilfenahme weiterer Informationen die Glaubwürdigkeit einschätzen zu können.

Eine Kombination der verschiedenen Intentionen ist ebenso denkbar. Mithilfe der Intentionen ist es nachstehend möglich, einige der Gemeinsamkeiten und Unterschiede zu interpretieren und zu erklären. Es ist jedoch vorab Folgendes anzumerken: Obwohl diese Arbeit keine qualitative Arbeit darstellt, werden entsprechende Probandenantworten und Beschreibungen ihres Verhaltens anhand der Logdaten herangezogen. Grund hierfür ist, dass sie die Interpretationsüberlegungen der Forschenden konkreter und einschlägiger darstellen und diese untermauern können. Es ist jedoch bewusst, dass sie lediglich Einzelfälle darstellen und die Verallgemeinerung bzw. Übertragung dieser Interpretation auf andere Probanden eingeschränkt ist.

Die erste **Auffälligkeit ist die Gemeinsamkeit**, dass auf der Ebene der textlichen Merkmale bei den Modellen mit NfC und NfCC sowie auf der Ebene der Medienmerkmale bei dem Modell mit NfC ein ähnliches Verhalten vorliegt und, dass dieses unabhängig von den Personenmerkmalen ist. Für IO ist dies jedoch im Vergleich nicht unabhängig von den Personenmerkmalen

und auf allen Ebenen unterschiedlich. Die gefundenen signifikanten Effekte in den Modellen mit NfC bzw. NfCC belaufen sich lediglich auf direkte Effekte der Webseitenmerkmale. Das Ergebnis spricht somit nicht für die Hypothesen bezüglich des direkten Effektes der Personenmerkmale sowie des Moderationseffektes. Es liegt hinsichtlich der Webseitenmerkmale zwar ein direkter Effekt auf das Anklicken der Weiterverlinkung vor, dieser wurde ursprünglich jedoch als non-linearer Effekt postuliert. Deshalb ist auch die entsprechende Hypothese zu verwerfen.

Mögliche Erklärung. Warum die Probanden ungeachtet der Personenmerkmale von NfC und NfCC, aber unter Beachtung von IO derart handeln, kann mit dem in der Situation zum Tragen kommenden Personenmerkmal erklärt werden. Obwohl NfC und NfCC auch situativ hervorgerufen werden können, sind die beiden Merkmale hauptsächlich als stabile Persönlichkeitsdispositionen zu verstehen (Schlink & Walther, 2007; Webster & Kruglanski, 1994; Petty et al., 2009). Sich schnell überladen zu fühlen (hohes IO), ist zwar ebenfalls eine Disposition, könnte aber in dieser Situation (20 Minuten, Zeitdruck und Stress) stärker forciert worden sein im Vergleich zu NfC und NfCC. Das heißt, dass der Zeitdruck bzw. Stress das IO im Vergleich zu NfC und NfCC stärker ausgelöst hatte, sodass sich unter verschiedenen IO-Werten unterschiedliche Verhaltensweisen zeigen konnten.

Warum die Probanden umso mehr Links verfolgen, wenn sie Webseiten besuchen, die bezüglich textlicher und mediumsbezogener Merkmale als glaubwürdiger eingeschätzt werden könnten, könnte wie folgt *interpretiert* werden: Sie bewerten die Webseite anhand der textlichen und mediumsbezogenen Merkmale und schätzen sie als glaubwürdig ein. Sie nutzen die vorhandenen Links, um weitere gute (glaubwürdige) Quellen bzw. Informationen darauf zu finden (Intention 3).

Proband *ID170* beispielsweise hat ein **hohes NfC** und ruft im Durchschnitt Webseiten auf, die eine hohe textliche und mediumsbezogene Glaubwürdigkeit aufweisen. Er klickt u. a. die Weiterverlinkung von *Quarks* zu *CJ Sportsmed*. Beides sind Webseiten mit hoher Textglaubwürdigkeit. In seiner Antwort bewertet er *Quarks* als glaubwürdig und nutzt die angegebenen Links, um weitere glaubwürdige Informationen, wie im Artikel von *CJ Sportsmed*, zu finden. Diese verwendet er ebenso in seiner Antwort und beurteilt sie als glaubwürdig:

„<https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/> Ich halte diese URL für zuverlässig: Quarks ist bekannt für seine genaue Literaturrecherche, bezieht sich auf aktuelle Studien der MHH und Norwegen. Quellenangaben und Links zu den Studien befanden sich am Ende der Seite. https://journals.lww.com/cjsportsmed/Abstract/2018/05000/Effect_of_E_Bike_Versus_Bike_Commuting_on.2.aspx
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7370133/> Beide halte ich für zuverlässig. Studien unter kontrollierten Bedingungen.“ [sic.] (*ID170*)

Das Gegenargument für diese Interpretation wäre, dass die Probanden nur deshalb die Links weiterverfolgen, weil sie nur genau solche Webseiten aufgerufen haben, die Links überhaupt erst zur Verfügung stellten. D. h., dass lediglich als glaubwürdig zu betrachtende Webseiten Links zu Verfügung stellen und sie deshalb die Links weiterverfolgt haben. Tabelle 15 zeigt jedoch, dass auch unglaubliche Webseiten anklickbare Weiterverlinkungen zur Verfügung stellen, welche wiederum angeklickt werden.

Eine zweite Auffälligkeit ist, dass ausschließlich IO den erwünschten Moderationseffekt zeigt und dieser je nach Webseitenmerkmalsblock auch unterschiedlich ist. Zwei Teilhypothesen könnten basierend auf dem Ergebnis angenommen werden: Probanden mit **niedrigem IO** klicken dann die Links an, wenn sie Webseiten mit einer mittelmäßigen oder unklaren Glaubwürdigkeit bzgl. textlicher Merkmale besuchen. Probanden mit **hohem IO** klicken ungeachtet der bewerteten Glaubwürdigkeit bzgl. mediumsbezogener Merkmale und Autorenmerkmale nicht die Links an.

IO und Text. Probanden mit **niedrigem IO** klicken genau dann die Weiterverlinkung an, wenn sie Webseiten mit einer mittelmäßigen Glaubwürdigkeit bzgl. textlicher Merkmale aufrufen.

Erklärung und Interpretation. Alle von ihnen besuchten Webseiten sind beispielsweise mittelmäßig verständlich und die Weiterverlinkung wird innerhalb des Rechercheprozesses angeklickt. Personen mit **geringem IO**-Empfinden erreichen den Schwellenpunkt, an dem Informationen überladend wirken, später (Chewning & Harrell, 1990; Roetzel, 2019). Sie haben möglicherweise größere kognitive Kapazitäten und könnten daher Webseiten mit hoher oder niedriger Glaubwürdigkeit bzgl. textlicher Merkmale direkt erkennen und weiterverwenden bzw. ausschließen. Dieses Verhalten entspräche somit genau der postulierten Hypothese, was sich beispielsweise anhand von Proband *ID156* belegen lässt. Er ruft im Durchschnitt als sehr glaubwürdig zu bewertende Webseiten hinsichtlich textlicher Merkmale auf, verfolgt aber die Links nicht weiter. Er befindet sich im Graphen rechts unten. Er stützt sein Urteil über die Glaubwürdigkeit auf textliche Merkmale (Transparenz, Quellenherkunft). Für ihn ist sein Urteil dahingehend ausreichend und er sieht keine Notwendigkeit, weitere Schritte zu tätigen, was die o. g. Vermutung stützt:

„1. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6249962/> - Peer reviewed review bzgl. der Gesundheitseffekte von E-Bikes 2. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5644161/> - exemplarische Primärstudie zu den Gesundheitseffekten der E-Bikes.“ [*sic.*] (*ID156*)

Für die Webseiten, die eine mittelmäßige oder unklare Textglaubwürdigkeit aufweisen, wollen sich Personen mit **niedrigem IO** entweder rückversichern und erneut überprüfen (Intention 2) oder haben vielleicht sogar aufgrund anderer Merkmale die Webseite als ausreichend

glaubwürdig eingeschätzt, um durch die Linkverfolgung weitere, gute, glaubwürdige Webseiten zu finden (Intention 3).

Proband *ID149* geht von *Ärzteblatt Artikel 1* zu *Ärzteblatt Artikel 2* weiter. Er besucht im Durchschnitt Webseiten, die eine mittelmäßige Textglaubwürdigkeit aufweisen und befindet sich auf dem Maximum der Kurve. Der *Ärzteblatt Artikel 1* selbst hat ebenfalls eine mittelmäßige Textglaubwürdigkeit. Er schätzt aber den Artikel aufgrund eines anderen Merkmals (Bekanntheit, Medium) als ausreichend glaubwürdig ein, weshalb hier Intention 3 zutreffen könnte. Er will durch die ihm bekannte und als glaubwürdig eingeschätzte Quelle weitere glaubwürdige Informationen finden:

„[...] <https://www.aerzteblatt.de/nachrichten/96943/E-Bikes-gesund-und-gefaehrlich-zugleich-Zuverlaessig>. Das Ärzteblatt ist eine mir bekannte Zeitschrift, wo man sich erste Eindrücke verschaffen kann und auch Links zu eventuellen Studien bekommt, die dem Artikel zugrundeliegen. <https://emotion-technologies.de/e-bike-infos/vorteile/gesundheit/> Nicht zuverlässig. Mir scheint, diese Seite möchte einem selbst etwas verkaufen und stellt es in ein positives Licht. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5644161/> Zuverlässig. Aber nicht genau das, was ich eigentlich gesucht habe. [https://www.aerzteblatt.de/nachrichten/96745/Studie-E-Biker-als-Pendler-koerperlich-\(fast\)-fitter-als-Radfahrer](https://www.aerzteblatt.de/nachrichten/96745/Studie-E-Biker-als-Pendler-koerperlich-(fast)-fitter-als-Radfahrer) Zuverlässig. Siehe Ärzteblatt oben. [...]“ [sic.] (*ID149*)

Proband *ID145* klickt ebenfalls die Weiterverlinkung an und ruft im Durchschnitt mittelmäßig glaubwürdige Webseiten bzgl. textlicher Merkmale auf. Er klickt den Link von *Focus* auf *Nature* und nutzt die Erkenntnis zur Beurteilung der Zuverlässigkeit von *Focus*. Die Angabe, dass im Artikel von *Focus* Quellen zu *Nature* enthalten sind, die er zurückverfolgen kann („Quellenangaben“), scheint ihm ausreichend, um seine anfängliche Skepsis aufgrund der Unübersichtlichkeit zu überwinden.

„[...] https://www.focus.de/gesundheitspraxistipps/gelber-bluetenstaub-dicke-luft-problem-fuer-allergiker-folgen-des-klimawandels_id_8851576.html -> unübersichtliche Webseite, viel Werbung die beim Lesen verwirrt, Quellenangaben -> wahrscheinlich zuverlässig, aber unübersichtlicher. [...]“ [sic.] (*ID145*)

Die Probandenantworten belegen demnach die vermutete Hypothese H4b, dass die Probanden bei Aufrufen von mittelmäßig glaubwürdigen Webseiten Links weiterverfolgen. Wie die unterschiedlichen Probandenantworten zeigen, kann der Grund für das Verhalten entweder die Rückversicherung der ursprünglich zuerst getätigten Evaluation (Intention 2) oder aber die Hoffnung sein, weitere glaubwürdige Quellen zu finden (Intention 3).

Probanden, die **schnell IO** empfinden, zeigen ein anderes Verhalten. Sie klicken die Weiterverlinkung bei Webseiten mit hoher textlicher Glaubwürdigkeit an. Bei Webseiten mit einer

niedrigen bis hin zu einer unklaren Glaubwürdigkeit werden keine Weiterverlinkungen angeklickt.

Obwohl Probanden mit **hohem IO** schnell eine Überlastung empfinden und deshalb oberflächlicher (Simons, 1988) suchen, zeigt sich doch solch ein Verhalten. Sie erkennen trotz ihrer eher heuristischen Informationsverarbeitung textliche Merkmale einer Quelle und ziehen sie zur Beurteilung der Glaubwürdigkeit heran. Eine mögliche Erklärung hierfür könnte sein, dass die Probanden mit hohem IO teilweise trotzdem systematisch verarbeiten, wenn sie z. B. hierfür noch ausreichend Kapazitäten haben. Dies deckt sich auch mit den Erkenntnissen von Chen et al. (2012, S. 343), die feststellten, dass auch innerhalb ihres Probandenkreises mit hohem IO einige oberflächlich (und somit heuristisch) und andere tiefergehend (und somit systematisch) Informationen verarbeiten. Das bedeutet, dass selbst bei hohem IO Unterschiede innerhalb dieser Teilstichprobe vorliegen können, welche hier nicht näher differenziert sind, aber das Ergebnis erklären könnten.

Das Verhalten kann folgendermaßen *interpretiert* werden: Die Weiterverlinkung anzuklicken hat möglicherweise das Ziel, Ressourcen zu sparen und direkt Quellen zu finden, die schnell Aufschluss über die Glaubwürdigkeit geben oder die Probanden zu weiteren guten Quellen führen könnten (Intention 4).

Proband *ID89* hat ein **hohes IO**, ruft im Durchschnitt Webseiten mit hoher Textglaubwürdigkeit auf und verfolgt die Links von *PubMed Artikel 5* zu *NCBI Artikel 2*. Trotz ebenfalls hoher textlicher Glaubwürdigkeit von *PubMed Artikel 5* verwendet er diesen Artikel selbst nicht, sondern lediglich den von NCBI. Das zeigt, dass er den Link deshalb verfolgt haben könnte, um Ressourcen zu sparen und entweder direkt weitere gute oder aber einfachere Informationen zu finden.

Wenn sie erkennen, dass die Quelle nicht ausreichend glaubwürdig ist, weil sie z. B. nicht aktuell oder intransparent ist, schließen sie diese aus dem Prozess aus und sehen keine Notwendigkeit, ihre Links zu untersuchen oder zu verwenden. Proband *ID162* beispielsweise besucht im Durchschnitt Webseiten mit geringer Textglaubwürdigkeit und klickt keine Weiterverlinkung an, weshalb er sich links unten auf der Kurve befindet. Er kommentiert die Glaubwürdigkeit von der textlich unglaubwürdigen Webseite *Elektrofahrrad24.de* und schlussfolgert aufgrund textlicher Merkmale (Unausgewogenheit der Argumente⁶¹), dass sie nicht als glaubwürdig beurteilt werden kann. Er schließt sie dann aus dem Prozess aus.

⁶¹ Argumentqualität bzw. Ausgewogenheit der Argumente ist zwar in der Forschung kein untersuchtes Merkmal, dennoch lässt sich anhand dieser Aussage schließen, dass der Proband textliche Merkmale erkannt hat. Es kann geschlussfolgert werden, dass er auch andere textliche Merkmale wahrgenommen hat.

Entgegen der vermuteten Hypothese H3c klicken die Probanden also nicht weniger die Weiterverlinkungen an. Sie betreiben trotz IO einen Zusatzaufwand, weil sie möglicherweise Ressourcen sparen (Intention 4) oder weitere glaubwürdige Quellen bzw. einfachere Quellen finden wollen (Intention 3).

IO und Medium. Probanden mit **niedrigem IO** klicken eher dann die Weiterverlinkung an, wenn sie Webseiten aufrufen, die eine hohe mediumsbezogene Glaubwürdigkeit aufweisen, d. h., wenn sie ein gewisses Genre aufweisen, wenig kommerzielles Interesse haben oder bekannt sind.

Mögliche Erklärung und Interpretation. Dies könnte damit erklärt werden, dass die Probanden aufgrund ihrer systematischen Informationsverarbeitung die Webseite relativ einfach und aufgrund von Heuristiken (z. B. Bekanntheitsheuristik) als glaubwürdig beurteilen konnten (H. Hong & Kim, 2020). Da sie sich generell weniger schnell überladen fühlen und die Beurteilung anhand von klar glaubwürdigen Webseiten kaum kognitiven Aufwand verbraucht, können sie möglicherweise noch Zusatzschritte machen. Sie klicken die Weiterverlinkung an in der Hoffnung, ebenfalls weitere gute Quellen zu finden (Intention 3).

Proband *ID129* besucht durchschnittlich sehr glaubwürdige Webseiten bzgl. der Medienmerkmale, weshalb er sich rechts oben auf der Kurve befindet. Er klickt von *Quarks* auf den *CJ Sportsmed Artikel*. Anhand seiner Stellungnahme wird klar, dass er die Webseite Quarks aufgrund des Genres und der Reputation (Medienmerkmale) beurteilt und als zuverlässig einstuft. Aufgrund dieser Einschätzung klickt er möglicherweise den Link an, um weitere gute Quellen zu finden:

„[...] <https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/> - macht auf mich einen seriösen eindruck, die seite auch als medium öffentlicher Nachrichten bekannt, über dei ARD finanziert soweit ich weiß, pro und contra gut abgewogen, Verweis auf Paper die das wissenschaftlich darstellen <https://www.egms.de/static/de/meetings/dkou2015/15dkou230.shtml> - Internetauftritt einer Fachgesellschaft, halte ich für relativ zuverlässig https://journals.lww.com/cjsportsmed/Abstract/2018/05000/Effect_of_E_Bike_Versus_Bike_Commuting_on.2.aspx - liest sich gut, gute Darstellung , Impactfactor von 3,.. (20219) nicht von der Hand zu weisen [...]“ [*sic.*] (*ID129*)

Bei den Webseiten, die sie direkt als unglaublich einstufen, investieren sie keinerlei weitere Ressourcen und untersuchen oder verwenden sie nicht weiter, wie anhand Proband *ID204* zu erkennen ist: Er ruft durchschnittlich unglaubliche Webseiten bezüglich Medienmerkmale auf und verfolgt keine Links weiter. Er beurteilt zwei seiner besuchten Webseiten aufgrund des Genres und des kommerziellen Interesses als unglaublich und schließt sie nicht in seine Stellungnahme ein.

Es zeigt sich daher, dass die Probanden mit **niedrigem IO** somit direkt eine erste Einschätzung anhand von schnell verfügbaren Entscheidungsregeln tätigen und zusätzlich ausreichend kognitive Kapazität vorhanden ist, Links zu verfolgen.

Anders sieht es bei den Probanden mit **hohem IO** aus. Gemäß der postulierten Hypothese klicken Probanden mit hohem IO ungeachtet der wahrgenommenen Glaubwürdigkeit bzgl. der Medienmerkmale selten die Weiterverlinkung an.

Mögliche Begründung und Interpretation. Erklärt werden könnte es damit, dass die Probanden mit hohem IO zwar die mediumsbezogenen Merkmale wahrnehmen und richtig einschätzen, indem sie z. B. ihnen bekannte Heuristiken anwenden. Sie sind aber bereits sehr überladen und fokussieren sich auf die Aufgabe. Da diese nicht vorsieht, dass zusätzliche Schritte i. S. v. Anklicken der Weiterverlinkung getätigt werden sollten, sehen sie hier auch keine Notwendigkeit das zu tun. Gemäß ihrer Disposition, sich schnell überladen zu fühlen (Chewning & Harrell, 1990; Roetzel, 2019), tun sie lediglich das Notwendige zur Aufgabenbearbeitung, um so ihre Kapazitäten nicht zu verschwenden.

Zwei Probanden, die **hohes IO** haben und keine Weiterverlinkung anklicken, liefern Belege hierfür. Proband *ID25* ruft im Durchschnitt sehr glaubwürdige Webseiten bzgl. der Medienmerkmale auf und befindet sich daher rechts auf der Geraden. Er beurteilt die Glaubwürdigkeit anhand des Genres und eines möglichen Interessenskonfliktes (kommerzielles Interesse). Er klickt keine Weiterverlinkung an, denn er ist möglicherweise an dieser Stelle bereits überlastet und möchte sich auf die eigentliche Aufgabenbearbeitung konzentrieren.

„<https://www.zeitschrift-sportmedizin.de/e-bikes-und-der-einfluss-auf-das-gesundheitsverhalten-von-risikogruppen/> Denke eher zuverlässig, da es ein Journal von Sportmedizinern zu sein scheint <https://pubmed.ncbi.nlm.nih.gov/29037235/> zuverlässig, da Artikel auf pubmed und kein Conflict of interest <https://pubmed.ncbi.nlm.nih.gov/30497998/> zuverlässig, da Artikel auf pubmed und kein Conflict of interest.“ [*sic.*] (*ID25*)

Proband *ID152* ruft im Durchschnitt sehr unglaubwürdige Webseiten bzgl. der Medienmerkmale auf und befindet sich daher links auf der Geraden. Er argumentiert für die Glaubwürdigkeit seiner Webseiten, u. a. damit, weil er sie kennt (Bekanntheit, Reputation).

„<https://www.theverge.com/2019/11/13/20963426/ebike-exercise-fitness-heart-rate-study-science-electric-bike> <https://formative.jmir.org/2019/3/e13643/> <https://www.forbes.com/sites/carltonreid/2018/11/23/riding-electric-bicycles-boon-to-health-and-not-cheating-confirms-literature-review/?sh=4ce3330735cc> für mich ist die Website (the Verge) eine bekannte Website für zuverlässige Nachrichten aus der technologische Branche, die ich regelmäßig genutzt habe, um Entscheidungen beim kauf eines Handys/Tablets genutzt habe. Sie und eine andere bekannte Website aus der USA (Forbes) haben beide die gleiche Studie erwähnt.“ [*sic.*] (*ID152*)

Die Belege zeigen also, dass die Probanden mit **hohem IO** ungeachtet der Webseitenmerkmale keine Weiterverlinkung anklicken, weil sie möglicherweise schon überladen sind und sich auf die eigentliche Aufgabenstellung konzentrieren wollen.

IO und Autor. Probanden mit **niedrigem IO** klicken ungeachtet des Autorenmerkmals die Weiterverlinkungen an. Es ist vielmehr ein linearer als ein non-linearer Effekt, wie postuliert. Dies spricht nicht für die entsprechende Hypothese H4b. Probanden mit **hohem IO** klicken ungeachtet des Autorenmerkmals weniger die Weiterverlinkungen an, wobei eine Tendenz besteht, dass bei mittelmäßiger bzw. unklarer Glaubwürdigkeit vermehrt die Weiterverlinkung angeklickt wird. Da dies allerdings auf einem sehr geringen Niveau geschieht, kann die Hypothese H3c tendenziell angenommen werden.

Für die Diskussion des Ergebnisses wird direkt die globale Betrachtung, also der Vergleich von hohem und niedrigem IO, herangezogen, denn sie sind beide tendenziell konstant linear. Obwohl bei Probanden mit **hohem IO** eine Tendenz besteht, dass sie bei mittelmäßiger oder unklarer Glaubwürdigkeit bzgl. des Autorenmerkmals eher die Weiterverlinkung anklicken, wird diesem Ergebnis nicht viel Wert beigemessen, da er nur auf einem sehr geringen Niveau stattfindet. Der Effekt ist überwiegend linear:

Mögliche Erklärung und Interpretation. Trotz dessen, dass ein Interaktionseffekt und ein direkter Effekt des Merkmals Autor vorliegen, scheint er dennoch keinen wesentlichen Einfluss auf das Verhalten der Probanden zu haben. Das Autorenmerkmal wird entweder nicht wahrgenommen und deshalb gibt es keine Unterschiede. Oder aber es wird wahrgenommen, die Probanden klicken aber aus anderen Gründen die Weiterverlinkung an bzw. nicht an. Da keinerlei Kommentare zu Autorenmerkmalen oder Autoren in den Stellungnahmen enthalten sind, lässt sich an dieser Stelle nichts Genaueres darüber urteilen.

Die **kumulierte, vergleichende Perspektive zeigt eine dritte Gemeinsamkeit** auf. Probanden mit **niedrigem IO** verfolgen zumindest für die Medien- und Autorenmerkmale generell häufiger die Links weiter im Vergleich zu Probanden mit **hohem IO**. Dies ist gemäß ihrer Disposition.

Mögliche Erklärung und Interpretation. Wenn Personen mit **niedrigem IO**, wie bereits mehrfach thematisiert, generell weniger überlastet sind, haben sie eventuell noch genügend Kapazitäten, um ihre Suche und entsprechend die Links weiter zu verfolgen. Personen mit **hohem IO** sind da möglicherweise bereits überladen und tätigen nur noch notwendige Schritte, um die Aufgabe zu bearbeiten. Oder aber sie tätigen eine komplett neue Suchanfrage. Eine neue Suchanfrage zu starten, statt eine bereits gefundene zu verfolgen, könnte als eine weniger systematische Suchstrategie interpretiert werden. Diese weniger systematische Vorgehensweise zeigt sich anhand der Logdaten. Von 75 Probanden, die **schnell IO** empfinden,

besuchen 72 Probanden als erstes oder als zweites eine Webseite, welche anklickbare Links zur Verfügung hat. Statt die Links auf den Webseiten jedoch weiter zu verfolgen, gehen sie zur Suchmaschine zurück und ändern ihren Suchbegriff bzw. starten eine neue Suchanfrage.

Beispielsweise hat Proband *ID126* ein **hohes IO** und ruft sechs Webseiten auf, von denen fünf anklickbare Links zu Verfügung stellen. Er ruft immer im Wechsel eine Webseite auf, kehrt zurück zur Suchmaschine und stellt eine neue Suchanfrage. Die Forschung liefert die Bestätigung für genau solch ein Verhalten: Probanden mit **hohem IO** neigen dazu, eine weniger systematische und gründliche Suchstrategie zu verfolgen (Swain & Haka, 2000). Darauf angewandt könnte das bedeuten, dass sie eine neue Suchanfrage starten oder die vorherige gänzlich beenden (da sie bereits überladen sind), statt tiefer zu gehen. Insofern lässt sich damit auch weitestgehend dieses Ergebnis erklären. Die Logdaten geben Hinweise darauf, dass dieses Verhalten eben genauso stattfindet.

Fazit. Modell II und III weisen je nach Betrachtung sowohl auf Webseitenmerkmalsebene als auch unter Betrachtung der Moderatoren (Personenmerkmale) unterschiedliche Ergebnisse in Form von Suchverhalten (Nachrecherchieren, Weiterverlinkung anklicken) auf. Unterschiedlich ist jedoch, dass bei Modell III nur mit IO Interaktionseffekte mit den Webseitenmerkmalen vorliegen. NfC und NfCC moderieren nicht den Effekt der Webseitenmerkmale auf das Anklicken der Weiterverlinkung, IO jedoch schon. IO scheint in Bezug auf dieses Suchverhalten somit ein großer, situationsbedingter Moderator zu sein. Obwohl NfC und NfCC auch situationsbedingt entstehen können, haben sie auf das Verhalten, anders als bei Modell II, keinen Einfluss. Ob die Weiterverlinkung angeklickt wird oder nicht, kann anhand dieser Ergebnisse überwiegend auf situationsbedingtes IO kombiniert mit der Wahrnehmung verschiedener Webseitenmerkmalsblöcke zurückzuführen sein. In Modell II gibt es weitere, diverse unterschiedliche Interaktionseffekte, die nicht weiter vertieft werden. Die Quintessenz aus dieser Erkenntnis ist allerdings, dass je nach Betrachtung der Webseiten- und Personenmerkmale unterschiedliche Verhaltensweisen und Effekte (non-linear, linear) resultieren.

Weiterhin ist festzustellen, dass die textlichen Merkmale ungeachtet der abhängigen Variablen (d. h. unabhängig von Modell II oder III) und der Personenmerkmale die meisten Effekte zeigen. Es liegen Interaktionseffekte und direkte Effekte auf den textlichen Merkmalsebenen vor. Die mediumsbezogenen und Autorenmerkmale zeigten teilweise keine, oder nur lineare direkte oder vereinzelt Interaktionseffekte auf. Im direkten Vergleich könnte daher geschlussfolgert werden, dass die textlichen Merkmale, stets unter der Prämisse der Wahrnehmung und korrekten Interpretation (Fogg, 2003), am ehesten einen Einfluss auf zumindest zwei Verhaltensweisen aus dem Suchprozess haben. Das verdeutlicht die Notwendigkeit, den Prozess zur Wahrnehmung und Interpretation verschiedener Webseitenmerkmale nochmals genauer

zu untersuchen und weitere Unterschiede herauszustellen. Es zeigt aber auch, dass die ursprüngliche Annahme, dass die unterschiedlichen einzelnen Merkmale als Äquivalente für die Glaubwürdigkeit stünden, nicht korrekt ist. Denn unterschiedliche Merkmale bzw. Merkmalsblöcke haben unterschiedliche Effekte auf die Verhaltensweisen. Sie sind folglich nicht alle einheitlich mit Glaubwürdigkeit gleichzusetzen. Möglicherweise wird einem Merkmal mehr oder weniger Wert bezüglich der Glaubwürdigkeit zugerechnet, was erneut vom Individuum abhängt. Mangels der Differenzierung in dieser Untersuchung bietet es eine Möglichkeit für zukünftige Forschungen, die unterschiedlichen Merkmale näher zu durchleuchten, was in den Implikationen folgend näher ausgeführt wird. Es sollte nicht nur untersucht werden, welche Merkmale zur Glaubwürdigkeitsbeurteilung und folglich unterschiedlichen Verhaltensweisen mit Online-Informationen beitragen, sondern auch wie (z. B. ob non-linear oder linear).

Letztlich verdeutlicht das Ergebnis, dass die angenommene Linearität zwischen Verhalten und Glaubwürdigkeitsbeurteilung überdacht werden muss. Unter bestimmten Merkmalen (z. B. Textglaubwürdigkeit oder IO) ist der Zusammenhang zwischen Glaubwürdigkeit und Suchverhalten nicht linear. Anders ausgedrückt bedeutet dies, dass sich mit zunehmender Bereitstellung von als glaubwürdig zu bewertenden Webseitenmerkmalen das wünschenswerte Suchverhalten (z. B. Weiterverlinkung verfolgen) nicht einstellt. Mehr ist nicht gleich besser. Aus diesem Grund trägt diese Arbeit dazu bei aufzuzeigen, dass Webseitenmerkmale unter verschiedenen Bedingungen zu verschiedenen Verhaltensweisen führen, welche nicht immer einer Linearität folgt.

6.2. Defizite bezüglich COR

Die Verteilung der abhängigen Variablen zeigen, dass die Probanden unzureichend die wünschenswerten Schritte zum effizienten Umgang mit Online-Informationen tätigen, wie es beispielsweise das IPS-I (Brand-Gruwel et al., 2009) oder COR (Molero et al., 2020) beschreiben.

Hinsichtlich der *Spezifität der Suchbegriffseingabe* ist zu erkennen, dass genau die Hälfte der Probanden jeweils oberflächliche bzw. spezifische Suchbegriffe eingeben. Oberflächliche Suchbegriffe repräsentieren dabei die Kopie der Fragestellung oder die Verwendung der Schlagwörter (E-Bike, Gesundheitsförderung) ohne weitere Zusätze. Spezifische Suchbegriffe sind alles darüber Hinausgehende: umformulierte Suchanfragen oder solche mit zusätzlichen Wörtern (Effekt, Outcome, Studie). Im Zuge eines effizienten und idealen Prozesses zur kritischen Recherche und zum Umgang mit Online-Informationen ist wünschenswert, dass Suchbegriffe spezifiziert werden (Brand-Gruwel et al., 2009). Insofern sollten idealerweise alle Probanden über die Suchanfrage reflektiert haben, um möglichst schnell und effizient an die gewünschten Informationen zu gelangen. Dies ist jedoch anhand der Verteilung nicht

erkennbar und ist mit bisheriger Forschung konform. Spink et al. (2001, S. 233) fanden beispielsweise in ihrer Analyse von Suchanfragen heraus, dass der Großteil ihrer Probanden kurze, wenig veränderte Anfragen mit einfacher Struktur verwenden, was mit einer oberflächlichen Suchanfrage gleichzusetzen ist. Aula und Nordhausen (2006) fanden heraus, dass die Studierenden selten Bool'sche Operatoren oder modifizierte Suchanfragen, also spezifische Suchanfragen, nutzen. Weber et al. (2018) berichteten über eine diesbezügliche Interventionsmaßnahme (Workshop) und schlussfolgern, dass selbst nach dem Workshop der Komplexitätsgrad der Suchanfragen immer noch auf niedrigem Niveau ist. Das Ergebnis dieser Forschung spiegelt sich in bereits vorliegender Forschung aus dem Bereich wider.

Kritische Bewertung des Ergebnisses. Das Ergebnis zeigt, dass die Hälfte der Probanden nur unzureichend über die Fragestellung und Suchbegriffe nachdenken, ehe sie eine erste Suche tätigen. Forschende fanden zugleich heraus (Weber et al., 2019), dass die Verwendung von fortgeschrittenen Suchstrategien (d. h. komplexe Suchanfragen, Anpassung der Suchbegriffe) positiv in Zusammenhang mit der akademischen Leistung steht. Die Verwendung spezifischer Suchbegriffe ist daher nicht nur aus Effizienzsicht bedeutsam, sondern auch im Hinblick auf die akademische Performanz. Kritisch zu hinterfragen ist jedoch, ob die berufliche Realität der Probanden auch tatsächlich abgebildet werden konnte. In der hier vorliegenden Studie haben sie 20 Minuten Zeit, die Aufgabe zu lösen. In der realen Arbeitswelt könnten sie jedoch mehr Zeit für eine Internet-Recherche haben. Dies ist pauschal jedoch nicht zu statuieren, da die Berufswelt der angehenden Juristen, Mediziner und Lehrer verschieden ist und andere Informationen und damit verbundene Entscheidungen erfordert.

Mögliche Erklärung. Das Verhalten, oberflächliche Suchbegriffe einzugeben, könnte damit erklärt werden, dass die Probanden eben genau aufgrund des empfundenen Stresses oder Zeitdrucks kognitive Kapazitäten und Zeit sparen wollen und direkt eine Suche starten, ohne darüber nachzudenken. Dies zeigt sich auch anhand der Logdaten, da die Probanden überwiegend zu Beginn und innerhalb der ersten beiden Minuten eine Suche tätigen. Sie lassen sich kaum Zeit, über die Suchbegriffe nachzudenken und starten auch später im Prozess keinen neuen Suchvorgang mit spezifizierten Begriffen.

Nachrecherchieren. Lediglich 44 (= ca. 33 %) Probanden recherchieren nach, was sich mit bisheriger Forschung deckt: Studierende überprüfen generell kaum die Quellen (Walraven et al., 2009) bzw. lesen nur selten ohne gezielte Schulung lateral (Breakstone, Smith, Connors, et al., 2021). *Das Ergebnis ist als kritisch zu bewerten*, denn es zeigt, dass die Probanden die Quellen aufrufen, ohne sich abzusichern oder eine zweite Meinung darüber einzuholen. In Anbetracht der Unmengen an wenig glaubwürdigen oder faktisch inkorrekten Online-

Informationen verdeutlicht das Ergebnis die Notwendigkeit, einen effizienten Umgang mit Online-Informationen gezielt zu schulen.

Mögliche Erklärung. Die Probanden könnten deshalb wenig nachrecherchieren, weil dies kein notwendiger Schritt zu Beantwortung der Aufgabenstellung darstellt. Sie ist ein empfehlenswerter Schritt im Idealprozess zur Lösung eines Online-Informationsproblems (Nagel et al., 2020). Sie wird aber nicht bewusst in der Aufgabenstellung ausgelöst, denn die Studierenden sollten lediglich die Zuverlässigkeit der Informationen aus der Online-Recherche überprüfen. Es ist nicht angemerkt, dass sie dies mithilfe des Internets oder dem Aufrufen weiterer Webseiten tun sollten. Die Begründung aus A1, warum eine Quelle als zuverlässig/glaubwürdig erachtet wird, muss nicht aus einer zusätzlichen Online-Recherche (also Nachrecherche) stammen, sondern kann auch mithilfe der eigenen Fähigkeit oder anhand der Merkmale auf den Webseiten beurteilt werden. Außerdem könnten sie aufgrund der limitierten Zeitdauer von 20 Minuten unter Druck oder Stress stehen. Dieser verursacht, dass sie sich einzig auf die geforderten Aufgabenstellungen konzentrieren und keine zeitliche Kapazität haben, Zusatzschritte zu machen oder andere Prozessschritte als bedeutsam erachten. Das Ergebnis verdeutlicht jedoch auch umso mehr, dass der idealtypische Prozess einer erfolgreichen und effizienten Online-Suche von Berufsanfängern nicht ausreichend vorhanden ist. Sie haben zwar in einer beruflichen Alltagssituation in der Regel keine Zeitbeschränkung und daher auch kein oder ein geringeres Stressempfinden, sodass sie sich möglicherweise intensiver mit ihrem Suchthema beschäftigen können. Dennoch zeigt es, dass der Idealprozess zur optimalen Suche noch nicht ausreichend verinnerlicht ist.

Weiterverlinkung anklicken. Lediglich 35 (= ca. 26 %) Probanden klicken verlinkte Webseiten oder Quellen an. Dass die Weiterverlinkungen angeklickt werden sollten, wird zwar nicht explizit in der Aufgabenstellung gefordert. Dennoch zeigt dieses Ergebnis *unter kritischer Betrachtung* die mangelhafte Recherchefähigkeit der Probanden auf, da idealerweise zur effizienten und kritischen Recherche von Online-Informationen Zitationen bzw. Quellen zurückverfolgt werden sollten (Caulfield, 2019). Das ist in Anbetracht der bisherigen Forschungen jedoch nicht überraschend, führten beispielsweise Walraven et al. (2009) an, dass Studierende nur in unzureichendem Maße Informationen bzw. Quellen überprüfen. Die Weiterverlinkung anzuklicken stellt eine Möglichkeit dar, Informationen zu überprüfen.

Mögliche Erklärung. Das Ergebnis könnte, stets unter der Prämisse, dass die Probanden die Verlinkung überhaupt wahrnehmen, zunächst einmal damit erklärt werden, dass die Probanden schlichtweg nicht daran denken die Links zu verfolgen oder diesen Rechschritt bewusst aufgrund der limitierten Bearbeitungsdauer auslassen. Sie haben zur Bearbeitung lediglich 20 Minuten Zeit, weshalb sie sich möglicherweise nur auf die tatsächliche

Aufgabenstellung fokussieren. Eine weitere mögliche Begründung, welche allerdings direkt ausgeschlossen werden kann, wäre, dass die Webseiten gar keine anklickbaren Verlinkungen enthalten und Probanden deshalb keine Links verfolgen. Diese Begründung kann aufgrund der dargestellten aufgerufenen unquien Webseiten aus Tabelle 15 in Kapitel 5.1.2 falsifiziert werden. Denn von beispielsweise 74 Probanden, die *Quarks* aufrufen, haben lediglich 11 die Weiterverlinkung angeklickt. Die verbliebenen 63 Probanden könnten dies ebenso tun, denn die Möglichkeit ist von Seiten der Webseite gegeben. Wie die Ausführungen und die Tabelle somit bestätigen, ist die Begründung zwar denkbar, jedoch nicht anwendbar.

Korrektheit der Quelleneinschätzung. Wie anhand der deskriptiven Statistik bereits dargestellt, schätzen die Probanden im Schnitt die Quellen eher korrekt ein ($M = 2.947$, $SD = 0.797$). Die Einschätzung der Rater über die korrekte Qualität der Quellen stimmt somit häufig mit der Einschätzung der Probanden überein.

Einordnung in den Forschungskontext und kritische Bewertung. Das Ergebnis, dass Probanden die Quellen generell eher korrekt einschätzen, findet sich auch in der Literatur wieder: Kiili et al. (2022) fanden in ihrer Forschung heraus, dass Probanden die glaubwürdigen Texte von den weniger glaubwürdigen Texten erfolgreich unterscheiden können. Dennoch ergibt die Prämissenüberprüfung, dass die Variable nicht normalverteilt ist, was in ebenfalls nicht normalverteilte Residuen resultiert und einen Hinweis auf eine fehlende Variable darstellen kann. Insofern ist das Ergebnis aus inhaltlicher Perspektive zwar überwiegend als positiv zu bewerten, denn es zeigt, dass die Probanden ihre verwendeten Quellen als korrekt glaubwürdig einschätzen. Dennoch kann die Nicht-Normalverteilung Konsequenzen auf das Ergebnis gehabt haben, sodass das Ergebnis aus methodischer Perspektive mit Vorsicht zu bewerten ist.

Kategorie der Quelleneinschätzung. Die Verteilung der AV systematische Begründung der Quelleneinschätzung zeigt, dass der Großteil der Probanden (= ca. 69 %) mindestens einmal eine Begründung aus einer systematischen Kategorie angibt. Im Umkehrschluss bedeutet es, dass 42 Probanden nur auf Basis von heuristischer Informationsverarbeitung bzw. Heuristiken die Glaubwürdigkeit einer Quelle beurteilen. Zwar ist die Anwendung von Heuristiken keineswegs pauschal schlechter oder führt zu einem unzureichenden Ergebnis (Wirth et al., 2007), dennoch sollten die Probanden gemäß eines effizienten Recherche- und Evaluationsprozess die Quelle idealerweise ausführlich und systematisch beurteilen (Molero et al., 2020; Brand-Gruwel et al., 2009). Es zeigt sich, dass über 30 % dies nicht tun.

Einordnung in den Forschungskontext und kritische Bewertung. Dieses Ergebnis deckt sich auch mit bisheriger Forschung. Kiili et al. (2022) schlussfolgerten, dass die Probanden zwar fähig sind glaubwürdige von unglaubwürdigen Informationen zu unterscheiden, ihre Begründungen aber nicht sonderlich ausgefeilt sind. Die Begründungen sind lediglich oberflächlich

und konzentrieren sich nicht auf spezifische Glaubwürdigkeitsaspekte (Kiili et al., 2022). Das Ergebnis ist kritisch zu bewerten, denn eine ausgefeilte und systematische Herangehensweise zur Begründung der Quellenglaubwürdigkeit ist in Anbetracht der zunehmenden unglaublichen oder faktisch inkorrekten Online-Informationen wichtig. Anhand stichhaltiger Beweise die Quellenglaubwürdigkeit kritisch zu beurteilen, ist im Rahmen einer effizienten Online-Recherche notwendig (Molerov et al., 2020; Brand-Gruwel et al., 2009). Das Ergebnis macht erneut deutlich, dass das Evaluationsverhalten der Probanden ausgebaut und geschult werden sollte.

Mögliche Begründung. Eine mögliche Erklärung findet sich in Anlehnung an die gleiche Forschung (Kiili et al., 2022), denn bei ihnen ist die Aufgabenbearbeitungsdauer der einzig statistisch signifikante Prädiktor. Je mehr Zeit ihre Probanden für die Aufgabe haben, desto qualitativ hochwertiger sind die Begründungen der Glaubwürdigkeit (Kiili et al., 2022). Es ist daher denkbar, dass die Begründungen für die Quellenglaubwürdigkeit wenig ausgeklügelt und oberflächlich (und somit heuristisch-kategorisiert) sind, weil die Probanden in dieser Aufgabe eine verhältnismäßig kurze Bearbeitungsdauer von 20 Minuten haben. Ihnen ist es möglicherweise nicht gelungen, die Quellen ausgiebiger auf Glaubwürdigkeit zu untersuchen und oder die Begründung sogar noch in A1 zu verschriftlichen.

Argumentationsstruktur. Die Verteilung der Argumentationsstruktur zeigt, dass 88 der Probanden eine gute und 55 eine schlechte Argumentationsstruktur i. S. v. von vollständig und gesundheitsbezogen aufweisen. Obwohl das Ergebnis als überwiegend positiv zu bewerten ist, wird in Anbetracht der prozentualen Betrachtung deutlich: Circa 39 % der Probanden argumentieren entweder nicht vollständig oder nicht inhaltsbezogen oder verfehlen die Aufgabenstellung komplett. Dies ist im Einklang mit aktueller Forschung, fanden beispielsweise Münchow et al. (2023, S. 717) zumindest für Bachelorstudierende heraus, dass diese überwiegend unzureichend (nur 17 % korrekte Zuordnungen) die verschiedenen Bestandteile bzw. Struktur eines argumentativen Textes erkennen und zuordnen. Nach einer Intervention kann der Anteil der korrekt zugeordneten Antworten zumindest auf 22 % erhöht werden. Die Bestandteile eines gut strukturierten bzw. vollständigen Argumentes zu erkennen, ist neben weiteren auch eine Voraussetzung dafür, dass dies auch selbst in einem eigens geschriebenen Argument umgesetzt werden kann (Newell et al., 2011). Insofern liefert die Forschung einen Ansatzpunkt, dass das hier gefundene Ergebnis ein aktuelles Problem der Hochschulbildung ist. Dies ist jedoch als kritisch zu betrachten, denn weiterhin wurde herausgefunden, dass zumindest die Evaluation von Argumenten in positivem Zusammenhang mit akademischen Erfolg steht (Münchow et al., 2019). Gut strukturierte Argumente zu erkennen, weiterhin zu bewerten und letztlich selbst zu erfassen, sollte daher eine notwendige Fähigkeit im Umgang mit Online-Informationen und zur kritischen Auseinandersetzung damit sein.

Mögliche Erklärung. Das Vorkommen an schlechter Argumentationsstruktur könnte zum einen unter Zuhilfenahme der Daten mit mangelndem Zeitmanagement erklärt werden. Eine unvollständige Argumentation zu schreiben, war unter Umständen nicht beabsichtigt. Die Probanden sind möglicherweise nicht fertig geworden oder haben gar nicht erst mit dem Schreiben begonnen, weil sie sich in den vorherigen Schritten zu viel Zeit gelassen oder unterschätzt haben, wie viel Zeit für das Verfassen eines vollständigen und gesundheitsbezogenen Argumentes benötigt wird. Sie haben sich möglicherweise ihre Zeit schlecht eingeteilt. Die bereits angefangenen Abschnitte und unvollständigen Sätze aus den Daten A2 belegen zumindest für manche Fälle, dass die schlechte Argumentationsstruktur auf schlechtes Zeitmanagement zurückzuführen sein kann.

Weiterhin könnten die generellen Schreib- und Lesefähigkeiten der Berufsanfänger unzureichend sein, sodass überhaupt ein gut strukturiertes Argument verfasst werden könnte. Denn Schreibfähigkeiten und Lesefähigkeiten hängen miteinander zusammen (Shanahan, 2006) und können sich auch gegenseitig verbessern (Graham & Harris, 2017). Argumentativ schreiben zu können, wird demnach von der Lesefähigkeit bedingt, welche bei den Berufsanfängern möglicherweise generell oder aufgrund der Situation (Zeitdruck, Stress) nicht kompetent genutzt wird.

Letztlich ist denkbar, dass die Berufsanfänger aufgrund ihrer Erfahrungen beim argumentativen Schreiben (Bachelor- und Masterthesis) zwar ausreichend theoretisches Wissen über das Verfassen von gut strukturierten Argumenten besitzen, diese aber in der Situation nicht anwenden können, da eine bisherige Situation (Informationsproblem, 20 Minuten Zeit) nicht (ausreichend) geschult wurde. Diese Diskrepanz zwischen theoretischem Wissen und tatsächlicher praktischer Anwendung ist bereits in der Argumentations-Literatur bekannt (Noroozi et al., 2013, S. 72; Valero Haro et al., 2019, S. 330).

Obwohl keine expliziten Hypothesen mit den soziodemografischen Merkmalen postuliert wurden und das Forschungsinteresse dieser Arbeit übergreifende Personenmerkmale umfasst, sind die Erkenntnisse aus den Zusammenhängen diskussionswürdig und sollten zumindest kurz Erwähnung finden, da auch sie ein Ansatzpunkt liefert perspektivisch die Hochschulbildung zu verbessern. Bisherige Forschungen (z. B. Arcand & Nantel, 2012; Chin et al., 2009) zeigen, dass soziodemografische Merkmale das kritische Online-Denken oder die Informationssuche und -evaluation beeinflussen können, weshalb diese zumindest als Kontrollvariablen mit aufgenommen wurden. Die für die Hochschulbildung signifikanten und relevantesten Ergebnisse, die der Domäne, werden nachstehend diskutiert.

Aus den Modellen II-A (NfC und AV Nachrecherche) und B (IO und AV Nachrecherche) resultiert das Ergebnis, dass Probanden des Lehramtes eine 8-fach bzw. sogar 9-fach größere

Wahrscheinlichkeit im Vergleich zur Referenzgruppe – Mediziner – haben, nachzurecherchieren. Dies deutete sich auch bereits in den korrelativen Ergebnissen an. Sie klicken also z. B. das Impressum an und/oder recherchieren die Webseite nochmals in einer Suchmaschine nach. Eine Intention zur Nachrecherche könnte sein, dass eine Quelle aufgrund der Unsicherheit nochmals gegengeprüft wird (Laterales Lesen; Breakstone, Smith, Connors, et al., 2021; Wineburg & McGrew, 2019). Unter Annahme dieser Intention kann das Ergebnis mithilfe der Logdaten folgendermaßen erklärt werden: Angehende Mediziner suchen verstärkt über wissenschaftliche Datenbanken (z. B. *PubMed*), welche sie im Rahmen ihres Studiums kennenlernen und nutzen. Da die Suchergebnisse wissenschaftliche Fachzeitschriften bzw. Artikel daraus sind, besteht die Vermutung, dass die Probanden implizit eine hohe Glaubwürdigkeit der Quelle annehmen. Für sie gibt es daher keine Notwendigkeit, nachzurecherchieren. Probanden des Lehramtes hingegen suchen überwiegend mit *Google* oder *Google Scholar*, was breitere Suchergebnisse liefert. Diese werden nicht, wie bei *PubMed*, „vorselektiert“, d. h. durch eine Datenbank gefunden. Die angehenden Lehrer sehen sich also mit allgemeineren und vielfältigeren Suchergebnissen konfrontiert, die sie zunächst durch eine Nachrecherche überprüfen müssen. Ihre offenerer, aber breiterer Suchprozess erfordert daher einen Zusatzschritt, den die angehenden Mediziner durch die Suche via *PubMed* bereits implizit tätigen. Dies könnte erklären, weshalb die angehenden Lehrer daher eine höhere Wahrscheinlichkeit haben, nachzurecherchieren.

Aus Modell V-A (NfC, IO und Kategorie der Quelleneinschätzung) und V-B (NfCC und Kategorie der Quelleneinschätzung) ist ersichtlich, dass angehende Lehrkräfte eher Begründungen aus der heuristischen Kategorie angeben. Die korrelativen Ergebnisse zeigten auch hier bereits eine Tendenz an. Das bedeutet, dass ihre Begründungen heuristischer Natur sind, ohne jegliche systematische Auseinandersetzung damit. Dies kann im Hinblick auf die universitäre Lehre und den Veranstaltungen erklärt werden. Obwohl angehende Lehrkräfte Inhalte zur pädagogische Diagnostik, also zur Lernprozessanalyse (Ingenkamp & Lissmann, 2008, S. 13) haben [z. B. an der JGU Mainz in einer Veranstaltung im Bachelor (Johannes Gutenberg-Universität Mainz, 2023a) und in einer weiteren im Master (Johannes Gutenberg-Universität Mainz, 2023b)], wird der Lehre der Analyse- und Diagnosefähigkeit bei angehenden Mediznern eine weitaus größere Rolle zugeschrieben. Sie lernen eine analytische, strukturierte Vorgehensweise zur Beurteilung von Informationen und Ableiten von Diagnosen mithilfe des sogenannten PICO (*Patient – Intervention – Comparison – Outcome*)-Schemas (Richardson et al., 1995), welches die Analyse- und Diagnosefähigkeit fördert. Die Diagnostik als essenzielle Kompetenz ist zudem als ein Lernziel im NKLM (*Nationaler Kompetenzbasierter Lernzielkatalog Medizin*) (Medizinische Fakultätentag, o. J.) verschriftlicht. Dies zeigt, dass die Mediziner bzgl. der Fähigkeiten per se schon besser ausgebildet werden und die medizinische

Ausbildung von vorneherein darauf ausgerichtet ist. Es ist daher davon auszugehen, dass Medizinstudenten sich häufiger mit den Fähigkeiten auseinandersetzen und auch selbst anwenden müssen. Die Fähigkeiten, sich analytisch und genau mit etwas zu beschäftigen, könnte auch in Situationen, wie z. B. der Online-Recherche, angewendet werden. Wenn sich also Medizinstudierende aufgrund ihrer antrainierten Fähigkeiten aus dem Studium tendenziell genauer und somit systematischer mit Informationen auseinandersetzen, so können sie diese Fähigkeit nutzen und die Quellenglaubwürdigkeit auch genauer ergründen, woraus mehr Begründungen aus dieser Kategorie in den Antworttexten resultieren könnten.

Aus den Modellen VI-A bis C (NfC, IO, NfCC und Textglaubwürdigkeit) wird der direkte, aber nicht postulierte Zusammenhang ersichtlich, dass die angehenden Lehrer im Vergleich zur Referenzgruppe der Mediziner weniger glaubwürdige Webseiten bzgl. textlicher Merkmale selektieren, was ebenfalls bereits in den korrelativen Ergebnissen zu erkennen war. Dies ist erneut mithilfe der Logdaten zu erklären. Anhand derer ist zu sehen, dass die angehenden Mediziner überwiegend Fachzeitschriften-Artikel selektieren, die sie zuvor via *PubMed* finden. Da diese Artikel bezüglich textlicher Merkmale eine höhere Glaubwürdigkeit aufweisen, z. B. transparenter, aktueller oder Primärquellen sind, Lehramtsabsolventen gleichzeitig aber weniger dieser Webseiten selektieren, ist damit das Ergebnis erklärbar.

Kritische Einordnung der Ergebnisse. Die Unterschiede zwischen angehenden Lehrkräften und Ärzten zeigt, dass sie unterschiedlich mit Informationen umgehen und möglicherweise durch die Hochschulbildung unterschiedlich darauf vorbereitet werden. Obwohl ihre berufliche Tätigkeit und Arbeitswelt verschieden ist bzw. sein wird und sie sich deshalb unterschiedliche spezifische Fähigkeiten angeeignet haben, ist der kritische Umgang mit Informationen in beiden Arbeitswelten gleichermaßen notwendig. Insofern zeigt das Ergebnis die Diskrepanz auf, die sich zwischen hier exemplarisch zwei Domänen im Hinblick auf die Schulung von kritischem Online-Denken und Umgang mit Online-Informationen ergibt.

7. Limitationen

Die Arbeit unterliegt gewissen Limitationen, welche nachfolgend vorgestellt werden.

Da die Daten aus einem durch Drittmittel geförderten Projekt stammen, gehen damit Restriktionen einher, die wenig Raum für Einflussnahme gewährten. Insofern lag von vorneherein eine Datenabhängigkeit vor, mit der zwar eine theoretisch passende Idee gefunden wurde. Jedoch konnten aufgrund der mangelnden Passung der Daten zu der Idee in allen sechs Modellen fast keine signifikanten Ergebnisse nachgewiesen werden. In diesem Zug sind als erste Reihe von Restriktionen diverse Prämissenverletzungen zu nennen:

Prämissenverletzungen. Die Stichprobengrößen mit $N = 144$ sowie $N = 64$ sind unzureichend für die Anzahl an Parametern, die in den Modellen untersucht wurden. Obwohl die Webseitenmerkmale bereits in drei Blöcken aggregiert wurden, waren die Modelle mit den Interaktionen zu umfangreich. Dies ist ein Grund, warum keine signifikanten Effekte gefunden wurden.

Damit einhergehend befanden sich zu wenige Merkmalskombinationen in der an sich schon kleinen Stichprobe. Obwohl die Personenmerkmale isoliert betrachtet eine ausreichende Streuung zeigten, waren alle Merkmale in Kombination (z. B. NfC, Textglaubwürdigkeit und die Nachrecherche als AV) in zu geringer Zahl und Ausprägung vorhanden. Hinzukommt, dass IO und die Webseitenmerkmale nicht normalverteilt waren. Die Unterschiede zwischen den Probanden konnten aufgrund der wenig gestreuten Merkmalskombinationen, der kleinen Stichprobe und den möglichen Problemen aufgrund der Nicht-Normalverteilung nicht sichtbar werden.

Die unterstellte Prämisse der Linearität war bei manchen Modellen nicht vorhanden. Bergtold et al. (2017) führen an, dass es bei kleinen Stichproben (< 200) problematisch sein könnte, wenn in einer logistischen Regression keine Linearität vorkommt. Die Linearität wurde für die Modelle, bei denen signifikante Modellgüten und Regressionskoeffizienten gefunden wurden, schlichtweg angenommen. Eine gesonderte Prüfung erfolgte nicht.

Entgegen der Annahme, dass die relevantesten Variablen in den Modellen enthalten waren, zeigen die Daten, dass Underfitting (Backhaus, 2021) vorliegen könnte. Die Personenmerkmale haben lediglich in zwei von sechs Modellen einen Einfluss gezeigt. Es scheint daher, dass nicht die richtigen, für das Verhalten verantwortlichen Personenmerkmale in den Modellen integriert wurden. Mögliche Personenmerkmale, die im Zusammenhang mit Such- und Evaluationsstrategie sowie kritischem Denken stehen, sind u. a. Motivation (Chaiken, 2014; Metzger, 2007; Petty & Cacioppo, 1986), Stressempfinden (Bermes, 2021), Erfahrungen mit dem Informationsgegenstand bzw. der Domäne (Experte/Novize) (Brand-Gruwel et al., 2017) oder dem Medium (Flanagin & Metzger, 2000, 2007), epistemische Überzeugungen bezüglich des Internets (Kammerer et al., 2021), Leseverständnis (Hahnel et al., 2018) oder Vorüberzeugungen (Bouali & Kolinsky, 2023) und Wissen (Werth et al., 2020).

Probleme der Konzeptionalisierung. Auch als Restriktion im Projektrahmen zu führen, aber vielmehr der Konzeptionalisierung zuzurechnen, ist die Limitation, dass die Aufgabenstellung möglicherweise nicht richtig verstanden wurde. Die Aufgabe war u. a. „die von Ihnen verwendeten Quellen auf Zuverlässigkeit zu überprüfen“. Hierunter kann einerseits verstanden werden, alle besuchten Quellen, denen man begegnet, auf Glaubwürdigkeit zu überprüfen. Andererseits kann es aber auch dahingehend verstanden werden, dass nur die Quellen überprüft werden sollten, die für die Stellungnahme verwendet werden. Anhand der Logdaten und der

Antworten A1 und A2 ist ersichtlich, dass die Probanden dies teils gemäß ersterem Verständnis, teils letzterem Verständnis beantwortet haben. Sie haben also möglicherweise nicht alle gleichermaßen die Webseiten beurteilt und unterschiedlich viel Zeit für die Ausführlichkeit der Beurteilung und das Verfassen der Antwort aufgebracht.

Weitere konzeptionell-bedingte und unabhängig vom Projektkontext zu betrachtende Limitationen werden nachstehend thematisiert: Wie bereits im Zwischenfazit in Kapitel 4.2.4.11 erwähnt, flossen nicht alle Webseitenmerkmale, die zu Beginn herausgearbeitet wurden, in die Analyse mit ein. Ein als anfänglich holistisches bezeichnetes Modell ist somit nicht gegeben. Lediglich neun von 22 angenommenen Webseitenmerkmalen finden sich im Modell aggregiert in drei Blöcken wieder. Dies ist entweder dem Kontext, den Konzeptualisierungs- oder den Analyseschwierigkeiten geschuldet. Dennoch können auch diese (z. B. Narrative oder Informationsstruktur) in Zusammenhang mit Glaubwürdigkeit gebracht werden.

Hinsichtlich der wahrgenommenen Glaubwürdigkeit wurde lediglich erfasst, ob die Webseiten die entsprechenden Merkmale bereitstellten. Wie bereits zu Beginn des theoretischen Kapitels kurz angeführt, sind laut Prominenz-Interpretations-Theorie (Fogg, 2003) die Prominenz bzw. die Wahrnehmung, und die Interpretation für die Beurteilung der Glaubwürdigkeit essenziell. Diese beiden Schritte wurden jedoch nicht geprüft, sondern lediglich vorausgesetzt. Ob ein Proband ein vorhandenes Merkmal auch tatsächlich wahrgenommen und dann auch noch korrekt interpretiert hat, ist nicht bekannt.

Probleme in der Operationalisierung sind ebenfalls als unabhängig vom Forschungskontext zu betrachten: Für die Integration in die Modelle und zum Zwecke der Reduzierung der Modellkomplexität wurden die einzelnen Webseitenmerkmale in Merkmalsblöcke zusammengefasst. Der durch die Aggregation resultierende Datenverlust erschwerte die zielgerichteten Aussagen und Begründungen für das Verhalten der Probanden. Es war beispielsweise nicht möglich zu statuieren, ob Probanden nachrecherchierten, weil sie verständlichere oder transparentere Webseiten aufgerufen haben. Die beiden Merkmale wurden unter dem Merkmalsblock „Text“ subsumiert und ihr Einfluss zusammen gemessen.

Hinsichtlich des Merkmals Aktualität ist festzustellen, dass die Aktualität pauschal für alle Kohorten gleichsam angenommen wurde. Da die Kohorten aber jeweils im Abstand von circa sechs Monaten erhoben wurden, wäre eine aktuelle Webseite der ersten Kohorte nicht mit einer aktuellen Webseite der dritten Kohorte vergleichbar. Für die Einfachheit der Erhebung wurde pauschal und begründet mit Statistiken die Einteilung in aktuell bzw. nicht aktuell vorgenommen. Noch zuverlässiger wäre jedoch die Anpassung der aktuellen Webseiten gemäß dem Zeitraum der entsprechenden Kohortenerhebung.

Die abhängige Variable der systematischen Begründung wurde nach einem ersten, nicht erfolgreichen Versuch dichotomisiert. Die zuerst ermittelte, metrische Variable enthielt an sich schon viele Probleme, weshalb sich dafür ausgesprochen wurde, sie zu vereinfachen. Die hieraus erfolgte Dichotomisierung führte allerdings zu einem Informationsverlust. Genauere Aussagen wären jedoch mit einer metrischen Variablen, die das Verhältnis der genannten Begründungen zur Gesamtzahl misst, möglich.

Sowohl die Nachrecherche als auch das Anklicken der Weiterverlinkung wurden als dichotome Variable erhoben. Es wurde nicht ermittelt, wie oft ein Proband das Verhalten zeigte. Es kam zwar selten vor, dass Probanden mehr als einmal im Prozess nachrecherchiert haben oder die Links verfolgt haben. Dennoch hätte die quantitative Erfassung dieses Verhaltens zielgerichtete Aussagen und daran schließend weitere Erkenntnisse ermöglicht.

Zur Bildung der abhängigen Variablen der Linkverfolgung wurde nicht differenziert, um welche Art des Links es sich handelt. Bei der Operationalisierung wurde lediglich berücksichtigt, dass es ein anklickbarer Link ist. Es gibt jedoch unterschiedliche Arten von Links, z. B. die tatsächlich referierte Quelle oder weiterführende Links zu ähnlichen Artikeln. Um die genaueren Intentionen und damit einhergehend Begründungen für das Verhalten „Link weiterverfolgen“ zu verstehen, wäre eine Differenzierung der Links notwendig.

Viele Prozessschritte und deren Überlegungen zur Intention (warum wurde etwas getan) beruhen auf Überlegungen der Forschenden selbst, nicht jedoch der Probanden. Anhand der Logdaten konnten die Begründungen zwar eingegrenzt werden, dennoch sind treffsichere und zuverlässige Aussagen nicht möglich. Es konnte nicht nachvollzogen werden, warum ein Proband mit einem gewissen Personenmerkmal beim Aufrufen einer Webseite mit gewissen Merkmalen nachrecherchiert hat. Beispielsweise könnte ein Proband die entsprechende Webseite gekannt haben, was er aber nicht in seine Antwort A1 mit aufnahm, oder er hatte nicht mehr ausreichend Zeit und deshalb nicht nachrecherchiert. Das hieraus resultierende Verhalten, in dem Fall nämlich nicht nachzurecherchieren, wurde möglicherweise falsch den Webseiten- oder Personenmerkmalen attribuiert.

In dieser Forschungsarbeit wurde lediglich das Aufrufen der Webseiten getrackt. Obwohl zunächst scrolled-up/down-Einträge detailliert erfasst wurden, wurde auch dies verworfen, da manche Webseiten so kompakt gestaltet waren, dass sie ohne zu Scrollen lesbar/anschaulich gewesen wären. Ob die Webseite dann aber tatsächlich gelesen bzw. angeschaut wurde oder nur in einem Tab im Hintergrund geöffnet wurde, war hieraus nicht erkennbar. Die Wahrnehmung eben der Webseitenmerkmale wird lediglich unterstellt und als gegeben angenommen, ohne hierüber tatsächliches Wissen zu haben.

Letztlich ist ein methodisches Problem zu nennen, das ebenfalls unabhängig vom Projektkontext ist: Da der Großteil der abhängigen Variablen dichotom war, wurden logistische Regressionen berechnet. Für geringe Stichproben ist dies aber nur bis zu einer gewissen Anzahl ratsam. Bessere Methoden wären beispielsweise Bootstrapping, Random Forest Klassifikation oder Bayes' Statistik. Das war für vorliegende Stichprobe nicht anwendbar, da die Merkmalskombinationen ohnehin kaum streuten.

8. Implikationen für Forschung und Praxis

Trotz der soeben genannten Limitationen können zukünftige Forschungen von dieser Arbeit profitieren. Der Großteil der Hypothesen konnte nicht als signifikant verifiziert werden. Dennoch bietet die Arbeit ein Ansatzpunkt für weitere Forschende, die eben genannten Limitationen zu berücksichtigen und darauf aufbauend neue Experimente zu konstruieren, die zur Annahme der Hypothesen führen könnten.

Für mögliche neue Experimente können unter Berücksichtigung der Limitationen sowie Erkenntnisse folgende Ansatzpunkte für die *Bildungsforschung* genannt werden: Da die Idee dieser Arbeit ein Forschungsdesiderat trifft und somit als innovativ zu betrachten ist, sollte sie zukünftig weiterverfolgt werden. Das setzt voraus, dass neue Daten erhoben werden, die eine ausreichende Passung zur Idee gewährleisten. Dies könnte in einem ersten Schritt dadurch gelingen, dass die Stichprobe erhöht wird. Dem hier vorliegenden Underfitting sowie der teilweise mangelnden Linearität könnte dem somit entgegenwirken. Weiterhin hätte eine größere Stichprobenanzahl den Vorteil, dass die Merkmalskombinationen möglicherweise gestreuter und zu einer größeren Anzahl vorliegen. So könnte an der grundsätzlichen Idee der vielen Interaktionen zwischen Personen- und Webseitenmerkmalen festgehalten werden. Um zu gewährleisten, dass die Aufgabenstellung nicht missverstanden wird, könnte diese zum einen noch eindeutiger formuliert und weiterhin im Vorfeld getestet werden (Pretest). Dadurch können Ungereimtheiten bei der Formulierung direkt gelöst und verbessert werden. Werden die Probanden die Aufgabenstellung dann trotzdem nicht richtig gelöst oder falsch verstanden haben, wird dies auf die Fähigkeit der Probanden und nicht auf die Aufgabenkonzeption zurückzuführen sein.

Alle möglichen Personenmerkmale, die einen Einfluss auf den Informationsverarbeitungsprozess haben können, sollten integriert werden, was ebenfalls dem Problem des Unterfittings begegnen würde. Dies würde wesentlich zum Verständnis des gesamten Prozesses beitragen sowie weiterhin ein noch holistischeres Modell abbilden. Aus konzeptioneller Sicht sollten von vornherein die Forschenden, welche auf dem Gebiet der narrativen und computerlinguistische Analyse ausgebildet sind, integriert werden, sodass die Konzeptionalisierung und Operationalisierung aller Webseitenmerkmale ermöglicht werden können. Dies würde ebenso zu

einem ganzheitlichen Bild beitragen. Daran anknüpfend wird vorgeschlagen, dass die Personenmerkmale auch situativ und die Webseitenmerkmale wiederum generell erhoben werden sollten. Gemäß der Trait-State-Theorie (Steyer et al., 1999) stehen diese in einem wechselseitigen Verhältnis und beschreiben zwei unterschiedliche Betrachtungsebenen. Um diese beiden unterschiedlichen Ebenen aneinander anzunähern, könnten daher zum einen Personenmerkmale situativ erfasst werden. Das bedeutet, dass z. B. NfC direkt vor der Aufgabenbearbeitung sowie direkt im Anschluss daran erhoben wird. Somit könnte der Disposition auch ein situativer Charakter verliehen werden. Zum anderen könnten die Probanden nach der Aufgabenbearbeitung gefragt werden, auf welche Merkmale sie generell bzw. üblicherweise bei der Online-Recherche achten. Auch wenn diese übliche bzw. habituelle Wahrnehmung von Webseitenmerkmalen nicht gleichzusetzen ist mit einer Disposition bzw. dispositionellem Verhalten, würde dies der State-Variablen (Webseitenmerkmale) zumindest einen allgemeineren Charakter geben.

Da Stressempfinden (Bermes, 2021) und Zeitdruck (Maule et al., 2000) im Zusammenhang mit der Informationsverarbeitung stehen, könnte für das Stressempfinden und den empfundenen Zeitdruck kontrolliert werden, indem die Variablen kurz nach der Aufgabenstellung abgefragt werden. Zudem könnten zwei Gruppen gebildet werden, wobei die Kontrollgruppe eine Aufgabenstellung ohne Zeitlimit erhält. Die Experimentalgruppe löst dieselbe Aufgabe innerhalb von z. B. 20 Minuten. Bei ausreichender Stichprobengröße können dadurch Unterschiede zwischen den Gruppen erkennbar werden, die zwar auf z. B. Personenmerkmale zurückzuführen sein könnten, aber von empfundenem Stress oder Zeitdruck moderiert werden könnten. Zukünftige Forschungen könnten aus solchen möglichen Erkenntnissen nicht nur weitere Implikationen für die Einflussfaktoren, sondern auch für die Aufgabenkonstruktion ableiten.

Die konzeptionellen Vorüberlegungen der Webseitenmerkmale basierte auf Forschungen, die teilweise allgemeine Kontexte aber auch sehr spezifische Themenbereiche abdeckten. Insofern waren manche Webseitenmerkmale zwar theoretisch fundiert hergeleitet, aber letztlich aufgrund kontextbedingter Probleme nicht realisierbar. Zukünftige Forschung könnte diesem Problem entgegentreten, indem sie zum einen unterschiedliche Aufgabenstellungen zu verschiedenen Themenbereichen erstellt. Dadurch können die spezifischen Erkenntnisse dieser Arbeit auf andere Kontexte transferiert und eine Generalisierung erzielt werden. Zum anderen könnten zukünftige Forschende versuchen, die Webseitenmerkmale so allgemeingültig zu formulieren und weiterhin zu operationalisieren, sodass sie über die verschiedenen Aufgaben und Kontexte hinweg Anwendung finden können. Es gilt hier eine Ausgewogenheit zwischen einer möglichst akkuraten und detaillierten aber auch einer ausreichend breiten und verallgemeinerbaren Konzeptionalisierung bzw. Operationalisierung zu finden. Eine weitere Limitation aus der Operationalisierung und thematisch daran anknüpfend zeigt einen Anknüpfungspunkt

für zukünftige Forschung auf. Das Webseitenmerkmal der Aktualität ist sehr aufgaben- bzw. kontextspezifisch und kann sich bei unterschiedlichen Erhebungszeiträumen verändern. Da unterschiedliche Erhebungszeiträume aufgrund unvorhergesehener geringer Teilnehmeranzahl nicht verhinderbar sind, sollte ein anderer Proxy für die Aktualität herangezogen werden, welcher nicht in direktem Zusammenhang mit der Datumsangabe auf der Webseite steht.

Da mit der Dichotomisierung mancher Prozess- und Performanzdaten Informationsverluste einhergingen, könnten zukünftige Forschende versuchen, metrische Variablen aus den Daten zu generieren. Hierdurch könnte ein korrekteres Abbild des tatsächlichen Verhaltens geschaffen werden. Konkret könnte z. B. die Variable der Nachrecherche unter Berücksichtigung der Zeit gemessen werden. Von den angegebenen 20 Minuten hat dann ein Proband beispielsweise XX Minuten dafür aufgewendet, um Quellen durch Anklicken des Impressums oder weiteres Suchen in einer Suchmaschine zu überprüfen. Für die Weiterverlinkung könnte zunächst erfasst werden, welche Art von Link angeklickt wurde. Diese Information kann später genutzt werden, um mithilfe eines Think-Aloud-Interviews die konkrete Intention für das Verfolgen des Links zu erfragen. In einem nächsten Schritt könnte dann berechnet werden, wie häufig ein Link verfolgt wurde im Verhältnis zu der möglichen maximalen Anzahl der gegebenen Links auf den besuchten Seiten. So könnten Probanden hinsichtlich der metrisch gebildeten Variablen besser verglichen werden im Vergleich zur hier getätigten dichotomen Operationalisierung, die nur erfasst, ob ein Link verfolgt wurde. Für eine detaillierte Operationalisierung der Argumentationsstruktur könnte nicht nur auf die Vollständigkeit bzw. den Inhaltsbezug geachtet werden, sondern auch die unterschiedlichen Argumente gezählt sowie deren Qualität (stark/schwach) und Positionierung (pro/kontra) erfasst werden. Eine derart detaillierte Aufschlüsselung der Argumentationsstruktur bzw. der Stellungnahme der Probanden ermöglicht tiefere Einblicke in ihre Gedankenprozesse und somit ihren Reasoning-Fähigkeiten. Durch diese Aufhebung der Dichotomisierung könnten daher zukünftige Forschungen quantitative Unterschiede messbar machen, welche in einem nächsten Schritt auch die Auswertungsmethoden vereinfachen würden.

Denn mit der logistischen Regression einer non-linearen Variablen inklusive von Interaktionen war die Auswertung sehr komplex. Zukünftige Forschungen könnten daher unter Annahme einer ausreichend großen Stichprobe andere Auswertungsmethoden, wie z. B. Strukturgleichungsmodelle, Random Forest, Bayes oder Mehrebenenmodelle anwenden. Sollten die einzelnen Prozess- und Performanzvariablen als metrische Variablen vorliegen, reduziert sich automatisch die Komplexität und weitere, einfachere Analysemethoden könnten herangezogen werden. Daran anknüpfend sollten die Methoden der narrativen Analyse sowie der (computer-)linguistischen Auswertung beachtet werden, da diese abgesehen von der Analyse der Webseitenmerkmale auch tiefergehende Einblicke in die Probandenantworten und darüber

hinaus in die damit verbundenen Informationsverarbeitungsprozesse liefern kann. So könnten beispielsweise die Narrative auf den besuchten oder selektierten Webseiten analysiert werden und weiterhin überprüft werden, inwieweit sich die Probanden bei der eigenen Stellungnahme von ihnen haben leiten lassen. Weiterhin könnte untersucht werden, ob Webseiten mit gewissen Narrativen häufiger selektiert wurden als andere. Bezüglich der linguistischen Analyse könnten beispielsweise strukturierende Elemente auf den Webseiten zu Rate gezogen werden und ermittelt werden, wie häufig diese Webseite mit mehr strukturierenden Elementen selektiert wurden.

Letztlich sollte das Set an Datenerhebungs- und auswertungsmethoden mit dem Einsatz von Think-Aloud-Interviews und Eye-Tracking (deutsch: Blickbewegungsaufzeichnung) komplettiert werden. Think-Aloud-Interviews ermöglichen die effektive Aufdeckung kognitiver Prozesse (McGrew, 2021b), ohne dass durch das Verbalisieren die kognitiven Prozesse und Strukturen stark verändert werden (Ericsson & Simon, 1980). Sie können folglich als valide Datenerhebungsmethode angesehen werden. Eye-Tracking ermöglichen Forschenden aufgrund der Blickbewegungen auf kognitive Prozesse innerhalb des Individuums zu schließen (Just & Carpenter, 1980). Think-Aloud und Eye-Tracking werden auch bereits in Kombination nachweislich erfolgreich eingesetzt (Neuert & Lenzner, 2019). Beide Methoden, ob separat oder kombiniert, haben daher das Potenzial tiefergehende Einblicke in die komplexen Denkweisen von Individuen zu geben.

In der Limitation noch als Interpretation der Forschenden selbst ausgelegt, könnte die Intention zur Nachrecherche oder zum Anklicken der Weiterverlinkung dank Think-Aloud-Interviews direkt von den Probanden erfragt werden. Dies würde mehr Aufschluss über die tatsächlichen Beweggründe für das Informationsverhaltensverhalten liefern.

Mithilfe des Eye-Trackings könnte valide zurückverfolgt werden, welche Webseiten tatsächlich angeschaut und welche lediglich als Tab im Hintergrund geöffnet wurden. Zudem könnte die bislang lediglich theoretisch angenommene Verbindung zwischen Prominenz, Wahrnehmung und Interpretation der Webseitenmerkmale aufgezeigt werden. In Kombination mit den Think-Aloud-Interviews könnten daher Probanden konkret danach gefragt werden, warum sie eine Webseite aufgerufen oder ein konkretes Merkmal angeschaut haben. Es bietet sich zukünftigen Forschungen weiterhin an, dass die Probanden nach der Aufgabenbearbeitung gefragt werden, auf welche Merkmale sie in der Situation zuvor geachtet hatten. Wenn anschließend das Think-Aloud in Kombination mit der Blickaufzeichnung durchgeführt wird, können die Probanden mit ihrem tatsächlichen Verhalten konfrontiert werden. Hieraus ergeben sich möglicherweise zusätzliche Erkenntnisse zur Selbsteinschätzung bzw. Diskrepanz des eingeschätzten und tatsächlichen Verhaltens. Wenn dann, wie an vorheriger Stelle bereits angedeutet, die

habituelle Glaubwürdigkeitsbeurteilung bzw. hierfür zu Grunde liegenden Merkmale damit verglichen werden, können weitere Erkenntnisse daraus gewonnen werden.

Das im September 2023 gestartete Forschungsprojekt CORE (CORE, 2023; Zlatkin-Troitschanskaia et al., 2023) adressiert beispielsweise bereits viele dieser Implikationen und kann zukünftig insbesondere von den Implikationen zur Datenauswertung bzw. Operationalisierung profitieren. Das Projekt hat unter anderem zum Ziel zu verstehen, wie Studierende Informationen aus zuverlässigen Online-Quellen erkennen und nutzen und für die Bearbeitung von studienbezogenen Aufgaben einsetzen (CORE, 2023; Zlatkin-Troitschanskaia et al., 2023). Folgende Eigenschaften des Projektes in Anlehnung an Zlatkin-Troitschanskaia et al. (2023) können angeführt werden, welche die von dieser Forschungsarbeit herausgestellten Implikationen bereits als Ansatzpunkt aufgreifen:

Mit der planmäßigen Erhebung von über 1.000 Studierenden ermöglicht das Forschungsprojekt eine weitaus größere Datenbasis als es in vorliegender Forschung der Fall war (S. 6). Es werden weiterhin Personenmerkmale, wie Vorwissen, Lesefähigkeit oder epistemologische Überzeugen integriert (S. 6). In Kombination mit einer quantitativen Inhaltsanalyse der Webseite (S. 26) sowie der Integration der Forschungsgebiete der Narrative (S. 27) und der Linguistik (S. 24) wird ein vollumfängliches und detailliertes Bild der möglichen Webseitenmerkmale dargestellt. Die Betrachtung der vielen Webseiten- sowie Personenmerkmale schafft ein holistischeres Modell, welches Aufschluss über mögliche Einflussfaktoren der Informationsverarbeitung geben könnte. Zudem ergänzen Eye-Tracking-Daten sowie Think-Aloud-Interviews (S. 29) die Datenerhebung bzw. -auswertung, sodass zielgerichtete Aussagen über das Verhalten und der Intention der Studierenden möglich sein werden. Aufgrund der Anwendung von verschiedenen GEN-COR bzw. DOM-COR-Aufgaben (S. 6) wird der hier vorliegenden Problematik der Aufgaben- bzw. Kontextspezifität Einhalt geboten und ermöglicht, die Erkenntnisse über verschiedene Aufgaben und Kontexte hinweg zu verallgemeinern.

Bildungspraxis. Obwohl die vorliegende Forschungsarbeit nicht untersucht hatte, ob und inwieweit die Hochschulbildung das kritische Online-Denken beeinflusst, zeigen sich anhand der Ergebnisse dennoch Defizite hinsichtlich COR am Ende der Hochschulbildung. Ob die Defizite allein auf die Hochschulbildung zurückzuführen sind, oder schon ein Ergebnis mangelnder allgemeiner Schulbildung darstellen, bleibt fraglich. Dennoch möchte die vorliegende Forschungsarbeit ihrem zu Beginn formulierten perspektivischen Ziel, die Erkenntnisse rund um das kritische Online-Denken und die von den Personen- und Webseitenmerkmalen abhängige Fehleranfälligkeit in die Curricula zu integrieren, gerecht werden. Hierzu werden nachfolgend konkrete Umsetzungsmöglichkeiten für die Bildungspraxis vorgeschlagen.

Die Hochschulbildung kann erstens die Erkenntnisse um die signifikanten Effekte der Personenmerkmale nutzen, um Lernende für ihre Disposition und den damit verbundenen Stolperfallen zu sensibilisieren. Gerade zu Beginn des Studiums kann es z. B. in Einstiegsveranstaltungen oder -kursen sinnvoll sein, die Studierenden über ihre Dispositionen und gängigen Probleme im Umgang mit Online-Informationen aufzuklären. Da die Nutzung von digitalen Medien im Studium, wie bereits erläutert, nicht mehr wegzudenken ist, werden Studierende zwangsläufig mit Online-Recherchen, dem kritischen Einschätzen von Informationen oder dem Erkennen von unglaublichen Informationen konfrontiert. Die vorliegende Forschung kann dabei helfen zu verstehen, mit welchen dispositionellen Voraussetzungen welche Schwierigkeiten beim Online-Problemlöseprozess entstehen. So können beispielsweise Personen, die schnell IO empfinden, dafür sensibilisiert werden, wie sie Informationen effizienter einschätzen und dadurch kognitive Kapazitäten sparen können.

Zweitens kann aufgrund der Erkenntnisse aus dem Vorhaben verstanden werden, an welcher Stelle im IPS-I bzw. COR Handlungsbedarf besteht. Die isolierte Betrachtung der Häufigkeiten der abhängigen Variablen zeigt auf, dass die Berufseinsteiger in unzureichendem Maße wesentliche Schritte zur Überprüfung von Informationen tätigen. Lediglich ca. die Hälfte recherchiert zusätzliche Informationen nach oder nutzt spezifische Suchanfragen. Gerade von absolvierten Masterstudierenden kann erwartet werden, dass sie durch das Hochschulstudium ausreichend befähigt sind, Informationen generell aber auch Online-Informationen im Besonderen kritisch zu hinterfragen und effizient zu nutzen (OECD, 2022). Eine jüngste Forschung zeigt (Weber et al. 2019), dass fortgeschrittene und spezifische Suchstrategien auch die akademische Performanz verbessern können. Insofern sollte bereits im Rahmen des Studiums eine entsprechende Schulung und Integration von COR stattfinden. Dies hätte nicht nur Auswirkungen auf die Noten im Studium, sondern könnte die Studierenden ausreichend auf einen guten und effizienten Umgang mit Online-Informationen im zukünftigen Berufsleben ausbilden.

Die Ausführungen zeigen, dass die Erkenntnisse essenziell für die Hochschulbildung sind. Es werden daher zwei konkrete Maßnahmen zum Transfer der Implikationen in die Hochschulbildung vorgeschlagen: ein studiumsübergreifender sowie ein studiumsbezogener Kurs.

Wie ein Literaturreview aufzeigt, gibt es neben Kurzzeitrainings (einmalig, 20 Minuten bis mehrere Stunden) bereits eine Vielzahl von Forschungen zu mindestens mehrwöchigen Interventions-Trainings zum ganzheitlichen Informationsproblemlöse-Prozess und Glaubwürdigkeitsbeurteilungen (Kammerer & Brand-Gruwel, 2020). Der Großteil der Langzeitrainings ist jedoch für Schülerinnen und Schüler ausgelegt (Kammerer & Brand-Gruwel, 2020), was aus lehr-lerntheoretischer Sicht zunächst Sinn macht. Denn jüngere Menschen, wie Schülerinnen und Schüler, lernen im Vergleich zu Erwachsenen, wie Studierende, effizienter (Frank et al.,

2022). Die vorliegende Forschungsarbeit plädiert jedoch (zusätzlich zu in der Schule bereits gelehrt Inhalten) dafür, zu Beginn bzw. während des Studiums Inhalte zum Informationsproblemlöse-Prozess und Glaubwürdigkeitsbeurteilung zu lehren. Denn, wie die Arbeit verdeutlicht, haben Berufsanfänger, welche im Laufe des Studiums eigentlich ihre Kompetenz hinsichtlich der Online-Informationsbeschaffung bzw. kritischem Denken aufgebaut oder verbessert haben sollten, zum Studienende jedoch nur unzureichende Fähigkeiten in COR.

Studiumsübergreifender Kurs. Alle Studierenden könnten zu Beginn ihres Studiums und ungeachtet ihres Studiengangs einen einheitlichen, verpflichtenden Kurs zum effizienten Umgang mit Online-Informationen erhalten. Bestandteil dieses Kurses könnte sein, welche kognitiven Personenmerkmale bei der Informationsverarbeitung und den damit einhergehenden möglichen Problemen in der Verarbeitung beteiligt sind. Die Studierenden könnten mithilfe eines Selbst-Testes ermitteln, wie die Kognitionen bei ihnen ausgeprägt sind und sich dafür sensibilisieren, dass sie für gewisse Prozesse im Verarbeitungsprozess anfälliger sind. Weiterhin könnten sie eine Schulung zu den 22 Webseitenmerkmalen erhalten und anhand realer Webseiten erarbeiten, ob sie die Merkmale erkennen. Zudem könnten sie in einem weiteren Selbst-Test mithilfe von Eye-Tracking erproben, auf welche Webseitenmerkmale sie in einer spontanen und realen Umgebung achten. Dieser Ansatz könnte die Studierenden umso mehr dafür sensibilisieren, beim Lösen eines echten, beispielsweise studienbezogenen Informationsproblems auf gewisse Webseitenmerkmale besonders zu achten. Denn wie die Autoren des Literaturreviews kritisch anmerken, wurden bisherige Interventionen in einer wenig authentischen und beschränkten Umgebung und bei Anwesenheit der Forschenden selbst getestet (Kammerer & Brand-Gruwel, 2020), weshalb die Effektivität der Interventionen möglicherweise nicht die tatsächliche Realität widerspiegelt. Mit einem Selbst-Test und einer realen Umgebung könnte dieses Problem umgangen werden. Weiterhin könnte ein weiteres Modul innerhalb des Kurses die spezifischen COR-bezogenen Schritte abdecken: Studierende könnten lernen, wann eine allgemeine oder spezifische Suche sinnvoll eingesetzt werden sollten, wie sie ihre Suche optimieren können (Operatoren, Datenbanken) und welche notwendigen Schritte es zur Überprüfung von Informationen bedarf (z. B. Nachrecherchieren, Links verfolgen, Quellenüberprüfung etc.). Letztlich sollte darin auch enthalten sein, wie sie die gefundenen Informationen in einem gut strukturierten Argument präsentieren können. Denn auch hier besteht noch Entwicklungsbedarf, wie die Arbeit aufzeigt.

Studiumsbezogener Kurs. Im Laufe oder gegen Ende ihres Studiums könnten die Studierenden einen Auffrischkurs zum Umgang mit Online-Informationen erhalten. Dies würde auch dem Kritikpunkt der Autoren des Literaturreviews gerecht werden, dass die Vielzahl der bisherigen Forschungen den Erfolg der Interventionen nicht auf lange Sicht messen (Kammerer & Brand-Gruwel, 2020). Mit der Auffrischung der Inhalte könnte ein Vergleich der

individuellen Fähigkeiten zum ersten Zeitpunkt (studiumsübergreifender Kurs) erfolgen. Inhaltlich unterscheide sich der studiumsbezogene vom studiumsübergreifenden Kurs dahingehend, als dass er auf den entsprechenden Studiengang zugeschnitten ist. Denn insbesondere inhaltliche und informationsbezogene Webseitenmerkmale, z. B. Argumentqualität oder Vollständigkeit, sollten im Zusammenhang mit denen im spezifischen Studiengang genutzten Informationen anhand von praktischen, domänenspezifischen Themen herausgearbeitet werden. Wie die vorliegende Arbeit aufgezeigt hat, gibt es zudem domänenspezifische Unterschiede in der Suchstrategie, da z. B. die angehenden Mediziner gezielter via *PubMed* suchen. Diese Best-Practices können sich im Laufe des Studiums entwickeln, daher sollte das ebenso studiumsbezogen berücksichtigt und nochmals darauf hingewiesen werden, dass sich für bestimmte Domänen oder Studiengänge unterschiedliche erste Anlaufstellen zur Suche (wie über fachspezifische Datenbanken) besser eignen. Konkret könnten sie erneut einen Selbst-Test zur Bestimmung ihrer Kognitionen vornehmen und sich nochmals neu dafür sensibilisieren, auf welche möglichen Hürden sie damit einhergehend stoßen. Weiterhin könnten anhand realer Webseiten alle Webseitenmerkmale aufgefrischt, die inhaltlichen Webseitenmerkmale vertieft sowie auf die studiumsbezogenen Anlaufstellen zur Suche (*PubMed* für Mediziner, *nwb*⁶² für Juristen etc.) hingewiesen werden. Letztlich könnten die spezifischen COR-Schritte erneut abgefragt und bei Bedarf über digitale Selbstlerneinheiten vertieft werden. Digitale Selbstlerneinheiten, wie z. B. Medienpakete mit Erklärvideos, bieten den Vorteil, dass sie in eigenem Tempo, ortsunabhängig und je nach individuellem Bedarf des Lernenden bearbeitet werden können (Brückner et al., 2021). Die Auffrischung am Ende des Studiums ermöglicht den baldigen Absolventen ihre Erkenntnisse auch noch nachhaltiger über den Studienabschluss hinaus in ihren Berufsalltag zu transferieren und anzuwenden.

9. Ausblick und Fazit

Zusammenfassung. Die vorliegende Forschungsarbeit untersucht den Einfluss von Personen- und Webseitenmerkmalen auf den Such- und Evaluationsprozess von Berufsanfängern, welche hierzu ein allgemeines Informationsproblem mithilfe des Internets lösen. Damit einhergehend hat das Vorhaben das Ziel, die konkreten Einflussfaktoren von NfC, IO, NfCC sowie der neun Webseitenmerkmale zu ermitteln und gleichzeitig als holistisches Modell abzubilden.

In Bezug auf das Ergebnis, die Hypothesenüberprüfung in Verbindung mit der Forschungsfrage, zeigt die Arbeit, dass erstens die Berufsanfänger in gesamter Betrachtung nur unzureichend die empfohlenen Schritte zum effizienten Umgang mit Online-Informationen durchführen. Sie werden dabei teilweise von ihren eigenen Dispositionen (z. B. IO) sowie von den

⁶² *nwb* (*nwb*, o. J.) ist eine fachspezifische, juristische Datenbank, mithilfe derer juristische Texte und Aufsätze unter anderem aus den verlageigenen Fachzeitschriften gesucht werden können.

zu beurteilenden Webseiten und ihren Merkmalen (z. B. Textglaubwürdigkeit) beeinflusst. Zweitens verifiziert die Arbeit damit, dass am Such- und Evaluationsprozess zum Umgang mit Online-Informationen Personen- und Webseitenmerkmale beteiligt sind, die Zusammenhänge aber nicht bei allen linear sind. Dadurch wird deutlich, dass das Verständnis des Online-Informationsverarbeitungsprozesses sehr komplex ist und Handlungsempfehlungen zur Verbesserung notwendig sind, dies jedoch noch weitere Forschungsarbeit erfordert.

Mehrwert. In Bezug auf den Prozess, der zu dem Ergebnis geführt hat, hat die Arbeit einen weitaus größeren Mehrwert:

Die grundsätzliche Idee, Personen- und Webseitenmerkmale in einem holistischen Modell abzubilden, scheint aufgrund der Wechselwirkung zwischen Personen als Traits und Webseiten als States gemäß der Theorie (Steyer et al., 1999) unabdingbar. Dennoch gibt es bislang nicht viele Forschungen, die beides gleichzeitig untersuchen (Zlatkin-Troitschanskaia, Hartig, et al., 2021) und dabei zusätzlich, wie in Kapitel 2.7 dargestellt, eine derartige Fülle an Webseitenmerkmalen und Personenmerkmalen betrachten. Die Forschungsarbeit ist dabei nicht nur dem Desiderat nachgekommen. Es ist darüber hinaus gelungen, trotz der Komplexität darüber hinaus auch beide unterschiedlichen Ebenen auf einer Vergleichsebene zu vereinen, ohne sich dabei bspw. eines Mehrebenenmodells zu bedienen. Dies ermöglichte aus methodischer Sicht die Schaffung einer einheitlichen Datenbasis, welche zwar aggregiert vorlag, was aber dennoch separate Aussagen zu jeweils Personen- oder Webseitenmerkmalsblöcken zuließ. Zudem mussten aufgrund der dichotomen Operationalisierung fast aller abhängiger Variablen logistische Regressionen gerechnet werden, was die an sich schon komplexen Modellberechnungen weiterhin verkomplizierten. Dennoch erwies sich diese Vorgehensweise als machbar und zielführend. Aufgrund der theoretischen Herleitung beider Einflussfaktoren sowie der methodischen Umsetzung dieser in einem holistischen Modell kann die Arbeit daher als ein Belegexemplar angeführt werden, dass derartige Ideen auch unabhängig eines Mehrebenenmodells, mithilfe logistischer Regressionen und trotz der Vergleichsproblematik aufgrund zweier Bezugsobjekte erfolgreich Anwendung finden können.

Mit der Beschreibung der allgemeinen Informationsverarbeitung sowie der Kombination dieser mit dem HSM (Chen & Chaiken, 1999), dem ISP-I (Brand-Gruwel et al., 2009) sowie dem COR-Konstrukt (Molero et al., 2020) wurde eine fundierte Grundlage geschaffen, die die Prozesse zur Informationsverarbeitung im Online-Kontext ganzheitlich darstellt. Es wurden dabei grundsätzliche Verarbeitungsprozesse (z. B. Prozesse und Kapazität im Arbeitsgedächtnis etc.) mit den zwei Wegen zur Verarbeitung (HSM) berücksichtigt. Dies wurde dann mit den Modellen des Online-Kontextes verknüpft. Resultierend hieraus ergab sich eine für den Online-Kontext anwendbare Prozessdarstellung, welche alle zuvor genannten Theorien bzw. Modelle

vereint. Diese Prozessdarstellung wurde anschließend auf den Prozess im Forschungskontext transferiert, welcher durch die ausführliche Darstellung der Prozessschritte und Berücksichtigung der vielen Theorien ausreichend theoretisch erfasst wurde. Die Arbeit liefert daher eine theoretische Basis zur Beschreibung der relevanten Prozessschritte, die tief in den Theorien der allgemeinen Informationsverarbeitung verankert ist und gleichzeitig die Unterschiedlichkeit der personenbedingten Verarbeitung (heuristisch/systematisch) sowie den digitalen Kontext aufgreift.

Wie die herausgearbeitete Lücke zeigt, gibt es bislang keine Forschung, die ein realistisches Szenario zu einem Online-Informationsproblem in Kombination mit derartig vielen Webseitenmerkmalen und gleichzeitig Personenmerkmalen abdeckt. Die angewandten Webseitenmerkmale wurden auf Basis eines ausgiebigen Literaturreviews ermittelt und herausgearbeitet, um erstens eine große Anzahl an Einflussfaktoren abzudecken und weiterhin bestmögliche Operationalisierungsstrategien zu finden. Als Ergebnis ist ein umfangreicher Review entstanden, der für weitere Forschende, wie z. B. auch das Forschungsprojekt CORE, als theoretische Grundlage und Übersicht dienlich sein kann. Im Zusammenhang damit ist ebenfalls herauszustellen, dass die Operationalisierung der Webseiten eine differentielle Taxonomie ist, die detailliert und akkurat anhand der im Real-Szenario aufgerufenen Webseiten herausgearbeitet wurde. Die Operationalisierung basiert daher auf einem innovativen Vorgehen, bei dem zwar echte, aufgaben- und kontextspezifische Webseiten zu Grunde liegen, die aufgrund der großen Diversität der Webseiten dennoch ausreichend verallgemeinerbar sind. Die Vorgehensweise zur Operationalisierung ist daher nicht nur sinnvoll für den vorliegenden Kontext, sondern kann auch auf andere Kontexte transferiert werden. Zukünftige Forschende können sich daher dieser Vorgehensweise bedienen.

Aufgrund der Abbildung eines realistischen Szenarios ergaben sich nicht nur eine große Diversität und Anzahl an Webseiten, sondern auch an Prozessdaten in Form von Logdaten. Mithilfe der genutzten Literatur (Molero et al., 2020, Brand-Gruwel et al., 2009) wurden die Logdaten sinnvoll taxonomisiert. Trotz der Komplexität der generellen Such-, Selektions- bzw. Evaluationsprozesse konnten hierdurch Indikatoren für die einzelnen Prozessschritte heruntergebrochen, abgeleitet und trennscharf herausgearbeitet werden. Dies stellt einen innovativen Ansatz dar, da es bislang kein allgemeines Vorgehen zur Operationalisierung bzw. Ableitung von Indikatoren in einem solch komplexen Prozess gibt. Mit dieser Forschungsarbeit wird eine Blaupause geschaffen, an welche sich Forschungen, wie beispielsweise das bereits mehrfach thematisierte Projekt CORE, bei Erhebung von Logdaten und zur Beschreibung von Prozessschritten bei der Suche von Online-Informationen orientieren können.

Weiterhin wird durch diese Forschungsarbeit eine gute Möglichkeit verdeutlicht, wie erfolgreich Mixed-Methods-Ansätze (Schreier & Odağ, 2020) zur Erhebung bzw. Analyse von verschiedenen Datentypen angewendet werden können. Trotz der Komplexität der verschiedenen vorliegenden Daten wurde durch deren Kombination ein optimaler Weg gefunden, Sinn aus den vielen unterschiedlichen Daten zu schaffen und darauf aufbauend derartige Ergebnisse und Implikationen abzuleiten. Die Anwendung des Mixed-Methods-Ansatzes ermöglichte es, ein vollständiges Bild des Forschungsgegenstandes zu schaffen (T. Lund, 2012). Diese Forschungsarbeit ist daher als ein positives Beispiel herauszustellen, welche durch die Integration verschiedener Datentypen ihre Möglichkeiten mit den vorliegenden Daten voll ausgeschöpft hat. Damit zusammenhängend ist anzumerken, dass die Arbeit einen Schritt darüber hinaus geht, da die Analyse der Zusammenhänge (logistische Regression mit non-linearen Variablen) selbst eine erneute Komplexität mit sich bringt. Insofern ist nicht nur die Verwendung des Mixed-Methods-Ansatzes zur Erhebung der Daten herauszustellen. Vielmehr ist die Kombination des Ansatzes mit unüblichen Zusammenhängen bzw. Analysen derer als neu zu bezeichnen und kann den Blickwinkel von Forschenden zur Betrachtung ihrer Daten und den damit verbundenen Möglichkeiten zur Analyse erweitern.

Fazit. Auf inhaltlicher Ebene sprechen die Ergebnisse somit dafür, dass das kritische Online-Denken bei vielen Berufsanfängern nur unzureichend ausgebildet ist und weiterhin, dass die klassische Annahme des linearen Zusammenhangs zwischen Glaubwürdigkeit und Informationsverarbeitungsverhalten nicht immer Anwendung findet. Die Ergebnisse sind aufgrund der genannten Limitationen jedoch als vorläufig zu betrachten. Es benötigt weitere Studien, wie in den Implikationen dargestellt, die die Zusammenhänge in ähnlichen Situationen, weiteren Aufgaben und unter Verwendung zusätzlicher Erhebungsmethoden (z. B. Think-Aloud) erneut untersuchen. Das Forschungsprojekt CORE, wie bereits mehrfach verdeutlicht, trifft genau den Kern dieser Limitationen und kann mithilfe des Untersuchungsdesigns, den weiteren geplanten Erhebungen sowie Auswertungen die von dieser Arbeit als vorläufig zu betrachtenden Zusammenhänge herausfinden und ggf. verifizieren.

Was die Arbeit jedoch vielmehr verdeutlicht ist, dass weniger vom Ergebnis an sich, sondern vom Weg dahin zu lernen ist. Der Prozess der Forschungsarbeit, der Datenaufbereitung sowie -analyse liefert mehr Erkenntnisse als die tatsächliche Überprüfung der Hypothesen. Der bewusste Vergleich zweier unterschiedlicher Bezugsobjekte füllt nicht nur die Forschungslücke, sondern ermöglicht eine wahrheitsgetreuere Abbildung der tatsächlich beteiligten Einflussfaktoren auf die Informationsverarbeitung bzw. dem damit verbundenen Verhalten. Die Forschungsarbeit liefert eine umfassende theoretische Grundlage unter Berücksichtigung der sowohl klassischen als auch digitalen Informationsverarbeitungsprozesse. Die ausführliche Recherche und Ausarbeitung der vielen Webseitenmerkmale gewährt eine überblicksartige und

gute Grundlage für eine Vielzahl an online-bedingter Einflussfaktoren. Damit einhergehend ist die Operationalisierung der ausgewählten Webseitenmerkmalen als Vorlage dienlich, um sie auf neue Situationen und Aufgaben zu transferieren. Die Taxonomisierung der Logdaten ist innovativ und ermöglicht eine trennscharfe Ableitung von prozessbasierten Indikatoren. Die Arbeit zeigt durch die Verwendung des Mixed-Methods-Ansatzes auf, welche Potenziale komplexe Daten haben und wie diese durch den Ansatz sinnstiftend und erfolgreich genutzt werden können.

Mit diesen ausgeführten Punkten hat die Arbeit aufgrund ihres Schöpfungsprozesses einen Mehrwert bzw. Werkzeugkoffer für zukünftige Forschungen, wie CORE (CORE, 2023) geschaffen, um von den detailliert beschriebenen Vorgehensweisen zu profitieren und durch Anwendung dieser letztlich hypothesenkonforme Ergebnisse zu erzielen. Das Ziel, den Einfluss der Personen- und Webseitenmerkmale auf die Prozess- und Performanzschritte beim Lösen eines Online-Informationsproblems zu ermitteln und in einem holistischen Modell abzubilden, ist in Anbetracht der Ergebnisse erreicht. Ein Bewusstsein für den Einfluss und das Zusammenspiel beider Betrachtungsebenen wurde geschaffen, sodass weitere Forschungen aufbauend auf den Erkenntnissen aus dem Zusammenspiel eine effizientere und weniger fehleranfällige Informationsverarbeitungsstrategie entwickeln könnten. Das perspektivische Ziel, die von den Personen- und Webseitenmerkmalen abhängige Fehleranfälligkeit in die Curricula zu integrieren, wurde ebenfalls in Ansätzen bei den Implikationen thematisiert. Online-Informationen kritisch einschätzen zu können und basierend auf dieser Einschätzung Entscheidungen zu treffen, ist in Anbetracht der aktuellen Entwicklungen eine der essenziellsten Fähigkeiten des 21. Jahrhunderts (González-Pérez & Ramírez-Montoya, 2022; van Laar et al., 2017). Diese Arbeit kann im Bereich der Bildungsforschung einen wichtigen Beitrag leisten, diesen Prozess des kritischen Online-Denkens besser zu verstehen.

Literaturverzeichnis

- Abdekhoda, M., Dehnad, A., & Yousefi, M. (2016). Effectiveness of training intervention to improve medical student's information literacy skills. *Korean Journal of Medical Education*, 28(4), 391–395. <https://doi.org/10.3946/kjme.2016.44>
- Accelerator, A. (o. J.). *Academic Accelerator | Accelerate Your Scholarly Research*. Academic Accelerator. Abgerufen 18. Dezember 2023, von <https://academic-accelerator>.
- Aljazzaf, Z. M., Perry, M., & Capretz, M. A. M. (2010). Online trust: Definition and principles. *2010 Fifth International Multi-conference on Computing in the Global Information Technology*, 163–168. <https://doi.org/10.1109/ICCGI.2010.17>
- Allen, M., & Preiss, R. W. (1997). Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, 14(2), 125–131. <https://doi.org/10.1080/08824099709388654>
- Al-Samarraie, H., Eldenfria, A., & Dawoud, H. (2017). The impact of personality traits on users' information-seeking behavior. *Information Processing & Management*, 53(1), 237–247. <https://doi.org/10.1016/j.ipm.2016.08.004>
- Amelang, M., & Bartussek, D. (1997). *Differentielle Psychologie und Persönlichkeitsforschung* (4. Aufl.). Kohlhammer.
- American Psychological Association. (o. J.). Objectivity. In *APA Dictionary of Psychology*. Abgerufen 26. März 2024, von <https://dictionary.apa.org/objectivity>
- Amichai-Hamburger, Y., Fine, A., & Goldstein, A. (2004). The impact of internet interactivity and need for closure on consumer preference. *Computers in Human Behavior*, 20(1), 103–117. [https://doi.org/10.1016/S0747-5632\(03\)00041-4](https://doi.org/10.1016/S0747-5632(03)00041-4)
- Amstad, T. (1978). *Wie verständlich sind unsere Zeitungen?* [Dissertation, Universität Zürich]. WorldCat. <https://search.worldcat.org/de/title/wie-verstandlich-sind-unsere-zeitungen/oclc/32497323>
- Anderson, J. R. (2013). *Kognitive Psychologie* (J. Funke, Hrsg.; 7. Aufl.). Springer VS.
- Andreassen, R., & Bråten, I. (2013). Teachers' source evaluation self-efficacy predicts their use of relevant source features when evaluating the trustworthiness of web sources on special education. *British Journal of Educational Technology*, 44(5), 821–836. <https://doi.org/10.1111/j.1467-8535.2012.01366.x>

- Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, 93(1), 59–79.
<https://doi.org/10.1177/1077699015606057>
- Arcand, M., & Nantel, J. (2012). Uncovering the nature of information processing of men and women online: The comparison of two models using the think-aloud method. *Journal of Theoretical & Applied Electronic Commerce Research*, 7(2), 106–120. <https://doi.org/10.4067/S0718-18762012000200010>
- Argelagós, E., & Pifarré, M. (2012). Improving information problem solving skills in secondary education through embedded instruction. *Computers in Human Behavior*, 28(2), 515–526.
<https://doi.org/10.1016/j.chb.2011.10.024>
- Artz, N., & Tybout, A. M. (1999). The moderating impact of quantitative information on the relationship between source credibility and persuasion: A persuasion knowledge model interpretation. *Marketing Letters*, 10(1), 51–62. <https://doi.org/10.1023/A:1008035107314>
- Aufderheide, P. (1997). Media literacy: A report of the national leadership conference on media literacy. In R. W. Kubey (Hrsg.), *Media literacy in the information age: Current perspectives* (S. 79–86). Routledge.
- Aula, A., & Nordhausen, K. (2006). Modeling successful performance in Web searching. *Journal of the American Society for Information Science and Technology*, 57(12), 1678–1693.
<https://doi.org/10.1002/asi.20340>
- Backhaus, K., Erichson, B., Gensler, S., Weiber, R., & Weiber, T. (2021). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (16. Aufl.). Springer Fachmedien Wiesbaden.
<https://doi.org/10.1007/978-3-658-32425-4>
- Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2006). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (11. Aufl.). Springer. <https://doi.org/10.1007/978-3-658-32425-4>
- Badke, W. (2021). *Research strategies: Finding your way through the information fog* (7. Aufl.). iUniverse.
- Banerjee, M., Zlatkin-Troitschanskaia, O., & Roeper, J. (2020). Narratives and their impact on students' information seeking and critical online reasoning in higher education economics and medicine. *Frontiers in Education*, 5, 570625. <https://doi.org/10.3389/feduc.2020.570625>

- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Barzilai, S., Thomm, E., & Shlomi-Elouz, T. (2020). Dealing with disagreement: The roles of topic familiarity and disagreement explanation in evaluation of conflicting expert claims and sources. *Learning and Instruction*, 69, 101367. <https://doi.org/10.1016/j.learninstruc.2020.101367>
- Bäsler, S.-A. (2021). Medienbildung in der ersten Phase der Lehrkräftebildung und die Bedeutung des medialen Habitus von Lehramtsstudierenden. In *Digitalisierung in Studium und Lehre gemeinsam gestalten*. (S. 431–451). Springer VS.
- Bawden, D., Holtham, C., & Courtney, N. (1999). Perspectives on information overload. *Aslib Proceedings*, 51(8), 249–255. <https://doi.org/10.1108/EUM00000000006984>
- Bawden, D., & Robinson, L. (2015). *Introduction to information science*. Facet Publishing.
- Bazrafkan, L., Hayat, A. A., Abbasi, K., Bazrafkan, A., Rohalamini, A., & Fardid, M. (2017). Evaluation of information literacy status among medical students at Shiraz University of Medical Sciences. *Journal of Advances in Medical Education and Professionalism*, 5(1), 42–48.
- Beck, S., & Gergel, R. (2014). *Contrasting english and german grammar: An introduction to syntax and semantics*. Walter de Gruyter GmbH & Co KG.
- Beißert, H., Köhler, M., Rempel, M., & Beierlein, C. (2014). *Eine deutschsprachige Kurzskala zur Messung des Konstrukts Need for Cognition: Die Need for Cognition Kurzskala (NFC-K)* (GESIS-Working Papers, 2014/32). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-403157>
- Belabbes, M. A., Ruthven, I., Moshfeghi, Y., & Rasmussen Pennington, D. (2023). Information overload: A concept analysis. *Journal of Documentation*, 79(1), 144–159. <https://doi.org/10.1108/JD-06-2021-0118>
- Belsley, D. A., Kuh, E., & Welsch, R. A. (1980). Detecting and assessing collinearity. In *Regression diagnostics: Identifying influential data and sources of collinearity* (S. 85–191). John Wiley & Sons, Ltd. <https://doi.org/10.1002/0471725153.ch3>
- Benninghaus, H. (Hrsg.). (2007). Die Beschreibung der Beziehung zwischen einer nominalen und einer metrischen Variablen. In *Deskriptive Statistik: Eine Einführung für Sozialwissenschaftler*

- (S. 228–249). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90739-0_8
- Bergtold, J. S., Yeager, E. A., & Featherstone, A. M. (2017). Inferences from logistic regression models in the presence of small samples, rare events, nonlinearity, and multicollinearity with observational data. *Journal of Applied Statistics*, 45(3), 528–546. <https://doi.org/10.1080/02664763.2017.1282441>
- Bermes, A. (2021). Information overload and fake news sharing: A transactional stress perspective exploring the mitigating role of consumers' resilience during COVID-19. *Journal of Retailing and Consumer Services*, 61, 102555. <https://doi.org/10.1016/j.jretconser.2021.102555>
- Bevölkerung nach Nationalität und Geschlecht (Quartalszahlen)*. (o. J.). Destatis. Statistisches Bundesamt. Abgerufen 14. Juni 2023, von <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/liste-zensus-geschlecht-staatsangehoerigkeit.html>
- Blanz, M. (2015). *Forschungsmethoden und Statistik für die Soziale Arbeit. Grundlagen und Anwendungen*. 1. Aufl. (1. Aufl.). Kohlhammer.
- Borah, P. (2014). The hyperlinked world: A look at how the interactions of news frames and hyperlinks influence news credibility and willingness to seek information. *Journal of Computer-Mediated Communication*, 19(3), 576–590. <https://doi.org/10.1111/jcc4.12060>
- Bouali, H., & Kolinsky, R. (2023). Source evaluation: Components and impacts. *Thinking Skills and Creativity*, 47, 101250. <https://doi.org/10.1016/j.tsc.2023.101250>
- Boukes, M., & LaMarre, H. L. (2021). Narrative persuasion by corporate CSR messages: The impact of narrative richness on attitudes and behavioral intentions via character identification, transportation, and message credibility. *Public Relations Review*, 47(5), 102107. <https://doi.org/10.1016/j.pubrev.2021.102107>
- Bourne, J. E., Sauchelli, S., Perry, R., Page, A., Leary, S., England, C., & Cooper, A. R. (2018). Health benefits of electrically-assisted cycling: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 15(1), 116–130. <https://doi.org/10.1186/s12966-018-0751-8>
- Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics*, 4(4), 531–550. <https://doi.org/10.1080/00401706.1962.10490038>

- Braasch, J. L. G., Bråten, I., & McCrudden, M. T. (Hrsg.). (2018). *Handbook of multiple source use*. Routledge. <https://doi.org/10.4324/9781315627496>
- Brand-Gruwel, S., Aarnoutse, C. A. J., & Van Den Bos, K. P. (1998). Improving text comprehension strategies in reading and listening settings. *Learning and Instruction*, 8(1), 63–81. [https://doi.org/10.1016/S0959-4752\(97\)00002-9](https://doi.org/10.1016/S0959-4752(97)00002-9)
- Brand-Gruwel, S., Kammerer, Y., van Meeuwen, L., & van Gog, T. (2017). Source evaluation of domain experts and novices during Web search. *Journal of Computer Assisted Learning*, 33(3), 234–251. <https://doi.org/10.1111/jcal.12162>
- Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior*, 21(3), 487–508. <https://doi.org/10.1016/j.chb.2004.10.005>
- Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, 53(4), 1207–1217. <https://doi.org/10.1016/j.compedu.2009.06.004>
- Breakstone, J., Smith, M., Connors, P., Ortega, T., Kerr, D., & Wineburg, S. (2021). Lateral reading: College students learn to critically evaluate internet sources in an online course. *The Harvard Kennedy School Misinformation Review*, 2(1). <https://doi.org/10.37016/mr-2020-56>
- Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M., & Saavedra, A. (2021). Students' civic online reasoning: A national portrait. *Educational Researcher*, 50(8), 505–515. <https://doi.org/10.3102/0013189X211017495>
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Brennan, K., Kelly, D., & Arguello, J. (2014). The effect of cognitive abilities on information search for tasks of varying levels of complexity. *Proceedings of the 5th Information Interaction in Context Symposium*, 165–174. <https://doi.org/10.1145/2637002.2637022>
- Broadbent, D. E. (1958). *Perception and communication* (1. Aufl.). Pergamon Press.
- Brückner, S., Saas, H., Reichert-Schlax, J., Zlatkin-Troitschanskaia, O., & Kuhn, C. (2021). Digitale Medienpakete zur Förderung handlungsnaher Unterrichtskompetenzen. *bwp@ Berufs- und Wirtschaftspädagogik – online*, 40, 1–25. https://www.bwpat.de/ausgabe40/brueckner_et_al_bwpat40.pdf

- Brünken, R., Münzer, S., & Spinath, B. (2019). *Pädagogische Psychologie—Lernen und Lehren*. Hogrefe Verlag GmbH & Company KG.
- Bullock, O. M., Shulman, H. C., & Huskey, R. (2021). Narratives are persuasive because they are easier to understand: Examining processing fluency as a mechanism of narrative persuasion. *Frontiers in Communication*, 6. <https://doi.org/10.3389/fcomm.2021.719615>
- Burgoon, M., Denning, V. P., & Roberts, L. (2002). Language expectancy theory. In J. P. Dillard & L. Shen (Hrsg.), *The persuasion handbook: Developments in theory and practice* (2. Aufl., S. 117–136). SAGE Publications.
- Burgoon, M., Jones, S. B., & Stewart, D. (1975). Toward a message-centered theory of persuasion: Three empirical investigations of language intensity. *Human Communication Research*, 1(3), 240–256. <https://doi.org/10.1111/j.1468-2958.1975.tb00271.x>
- Busse, D. (1992). *Textkohärenz und Textverstehen. Kriterien für Textualität in semantischer Perspektive* [Unveröffentlichtes Dokument]. Abgerufen am 10.02.2024, von https://www.germanistik.hhu.de/fileadmin/redaktion/Fakultaeten/Philosophische_Fakultaet/Germanistik/Germanistische_Sprachwissenschaft/Dateien/Busse/Text/Busse-UP-1992.pdf
- Busse, D. (2015). *Sprachverstehen und Textinterpretation: Grundzüge einer verstehenstheoretisch reflektierten interpretativen Semantik*. Springer-Verlag.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, B. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119(2), 197–253. <https://doi.org/10.1037/0033-2909.119.2.197>
- Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, 45(4), 805–818. <https://doi.org/10.1037/0022-3514.45.4.805>
- Caulfield, M. (2019, Juni 19). SIFT (The four moves). *Hapgood*. <https://hapgood.us/2019/06/19/sift-the-four-moves/>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>

- Chaiken, S. (2014). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Hrsg.), *Social influence: The ontario symposium, volume 5* (1. Aufl., S. 3–39). Psychology Press.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. Guilford Press.
- Chan, S. Y. (2001). The use of graphs as decision aids in relation to information overload and managerial decision quality. *Journal of Information Science*, 27(6), 417–425.
<https://doi.org/10.1177/016555150102700607>
- ChatGPT. (o. J.). [KI-Textgenerator]. ChatGPT. Abgerufen 21. Juli 2023, von <https://chat.openai.com>
- Chen, C.-Y., Pedersen, S., & Murphy, K. L. (2012). The influence of perceived information overload on student participation and knowledge construction in computer-mediated communication. *Instructional Science*, 40, 325–349. <https://doi.org/10.1007/s11251-011-9179-0>
- Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In *Dual-process theories in social psychology* (S. 73–96). The Guilford Press.
- Chervany, N. L., & Dickson, G. W. (1974). An experimental evaluation of information overload in a production environment. *Management Science*, 20(10), 1335–1344.
- Cheung, C. M.-Y., Sia, C.-L., & Kuan, K. K. Y. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association for Information Systems*, 13(8). <https://doi.org/10.17705/1jais.00305>
- Chewning, E. G., & Harrell, A. M. (1990). The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task. *Accounting, Organizations and Society*, 15(6), 527–542. [https://doi.org/10.1016/0361-3682\(90\)90033-Q](https://doi.org/10.1016/0361-3682(90)90033-Q)
- Chin, J., Fu, W.-T., & Kannampallil, T. (2009). Adaptive information search: Age-dependent interactions between cognitive profiles and strategies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1683–1692. <https://doi.org/10.1145/1518701.1518961>
- Choi, W. (2015). *New framework of web credibility assessment and an exploratory study of older adults' information behavior on the web* [Dissertation, Florida State University]. Diginole.
http://purl.flvc.org/fsu/fd/FSU_migr_etd-9571
- Clementson, D. E. (2020). Narrative persuasion, identification, attitudes, and trustworthiness in crisis communication. *Public Relations Review*, 46(2), 101889.
<https://doi.org/10.1016/j.pubrev.2020.101889>

- Cohen, A. R. (1957). Need for cognition and order of communication as determinants of opinion change. In C. I. Hovland (Hrsg.), *The order of presentation in persuasion* (S. 79–97).
- Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology*, 51(2), 291–294.
<https://doi.org/10.1037/h0042761>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. Aufl.). Lawrence Erlbaum Associates.
- Coiro, J. (2003). Exploring literacy on the internet: Reading comprehension on the internet: Expanding our understanding of reading comprehension to encompass new literacies. *The Reading Teacher*, 56(5), 458–464.
- Coleman, A. (2020, August 12). „Hundreds dead“ because of Covid-19 misinformation. BBC News.
<https://www.bbc.com/news/world-53755067>
- Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15–18. <https://doi.org/10.1080/00401706.1977.10489493>
- Coppola, V., Girandola, F., & Camus, O. (2020). A strong argumentative orientation makes the cognitive closure easier: The case for a persuasive health message. *Studia Psychologica*, 62(1), 5–22. <https://doi.org/10.31577/sp.2020.01.788>
- CORE – Critical Online Reasoning in Higher Education. (2023, November 28). Johannes Gutenberg University Mainz. <https://core.uni-mainz.de/>
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98–104. <https://doi.org/10.1037/0021-9010.78.1.98>
- Coutinho, S. A. (2006). The relationship between the need for cognition, metacognition, and intellectual task performance. *Educational Research and Reviews*, 1(5), 162–164.
- Coutinho, S. A., Wiemer-Hastings, K., Skowronski, J. J., & Britt, M. A. (2005). Metacognition, need for cognition and use of explanations during ongoing learning and problem solving. *Learning and Individual Differences*, 15(4), 321–337. <https://doi.org/10.1016/j.lindif.2005.06.001>
- Craig, T. Y., & Blankenship, K. L. (2011). Language and persuasion: Linguistic extremity influences message processing and behavioral intentions. *Journal of Language and Social Psychology*, 30(3), 290–310. <https://doi.org/10.1177/0261927X11407167>

- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Curry, A. L., & Stroud, N. J. (2021). The effects of journalistic transparency on credibility assessments and engagement intentions. *Journalism*, 22(4), 901–918. <https://doi.org/10.1177/1464884919850387>
- Dattalo, P. (2013). *Analysis of multiple dependent variables*. Oxford University Press.
- de Meo, A. (2012). How credible is a non-native speaker? Prosody and surroundings. In M. G. Busà & A. Stella (Hrsg.), *Methodological perspectives on second language prosody. Papers from ML2P* (S. 3–9). Cleup. <https://hdl.handle.net/11574/40089>
- Deci, E. L., & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223–238. <https://doi.org/10.25656/01:11173>
- Deepfakes—Gefahren und Gegenmaßnahmen*. (o. J.). Bundesamt für Sicherheit in der Informationstechnik. Abgerufen 18. Mai 2023, von https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/Deepfakes/deepfakes_node.html
- DeepL Übersetzer: Der präziseste Übersetzer der Welt*. (o. J.). Abgerufen 18. Dezember 2023, von <https://www.deepl.com/translator>
- Denner, N., Brocke, J., & Joeckel, S. (2016). Same result, different effect. The credibility attribution to public opinion research results. *Studies in Communication and Media*, 5(2), 173–196. <https://doi.org/10.5771/2192-4007-2016-2-173>
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, 70, 80–90. <https://doi.org/10.1037/h0039515>
- Dict.cc | Wörterbuch Englisch-Deutsch*. (o. J.). Abgerufen 18. Dezember 2023, von <https://www.dict.cc/>
- Diehl, V. A., & Wyrick, M. (2015). The relationships between need for cognition, boredom proneness, task engagement, and test performance. *SAGE Open*, 5(2), 2158244015585606. <https://doi.org/10.1177/2158244015585606>

- Dijksterhuis, A., van Knippenberg, A., Kruglanski, A. W., & Schaper, C. (1996). Motivated social cognition: Need for closure effects on memory and judgment. *Journal of Experimental Social Psychology*, 32(3), 254–270. <https://doi.org/10.1006/jesp.1996.0012>
- Dillard, J. P., & Marshall, L. J. (2003). Persuasion as a social skill. In J. O. Greene & B. R. Burleson (Hrsg.), *Handbook of communication and social interaction skills* (S. 479–513). L. Erlbaum Associates.
- Dissen, A., Qadiri, Q., & Middleton, C. J. (2022). I read it online: Understanding how undergraduate students assess the accuracy of online sources of health information. *American Journal of Lifestyle Medicine*, 16(5), 641–654. <https://doi.org/10.1177/1559827621990574>
- Donsbach, W. (1991). *Medienwirkung trotz Selektion: Einflussfaktoren auf die Zuwendung zu Zeitungsinhalten*. Böhlau.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Dorsch Lexikon der Psychologie. (o. J.). Trait. In *dorsch.hogrefe.com* Online Lexikon. Abgerufen am 15. März 2024, von <https://dorsch.hogrefe.com/stichwort/trait>
- Ebel, T., & Meyermann, A. (2015). Hinweise zur Anonymisierung von quantitativen Daten. *Frankfurt am Main : DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation. Forschungsdaten Bildung informiert Nr. 3*. <https://doi.org/10.25656/01:21970>
- E-Bike & Fitness: Mehr Power durch Motivation*. (2020, Juni 3). Stromer. <https://www.stromer-bike.com/de/e-bike-fitness-gesundheit>
- E-Bike fahren fördert Fitness und Gesundheit – schon nach vier Wochen*. (2018, Juli 26). Universität Basel. <https://www.unibas.ch/de/Aktuell/News/Uni-Research/E-Bike-fahren-foerdert-Fitness-und-Gesundheit-schon-nach-vier-Wochen.html>
- E-Bikes: Gesund und gefährlich zugleich*. (2018, August 7). Deutsches Ärzteblatt. <https://www.aerzteblatt.de/nachrichten/96943/E-Bikes-gesund-und-gefaehrlich-zugleich>
- Ecosia – Die Suchmaschine, die Bäume pflanzt*. (o. J.). Abgerufen 18. Dezember 2023, von <https://www.ecosia.org/>
- Edwards, C., Stoll, B., Faculak, N., & Karman, S. (2015). Social presence on LinkedIn: Perceived credibility and interpersonal attractiveness based on user profile picture. *Online Journal of Communication and Media Technologies*, 5(4), 102–115.

- Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden: Lehrbuch. Mit On-line-Material*. Beltz.
- Eisenberg, M. B., & Berkowitz, R. E. (1990). *Information problem solving: The big six skills approach to library & information skills instruction*. Ablex Publishing Corporation.
- Eisenegger, M. (2015). Identität, Image und Reputation – Eine kommunikationssoziologische Begriffsarchitektur. In R. Fröhlich, P. Szyszka, & G. Bentele (Hrsg.), *Handbuch der Public Relations*. (S. 431–460). Springer Fachmedien. https://doi.org/10.1007/978-3-531-18917-8_27
- Elliot, A. J., & Maier, M. A. (2007). Color and psychological functioning. *Current Directions in Psychological Science*, 16(5), 250–254. <https://doi.org/10.1111/j.1467-8721.2007.00514.x>
- Enge, S., Fleischhauer, M., Brocke, B., & Strobel, A. (2008). Neurophysiological measures of involuntary and voluntary attention allocation and dispositional differences in need for cognition. *Personality and Social Psychology Bulletin*, 34(6), 862–874. <https://doi.org/10.1177/0146167208315556>
- Eppler, M. J., & Mengis, J. (2004). The concept of information overload—A review of literature from organization science, accounting, marketing, MIS, and related disciplines (2004). In M. Meckel & B. F. Schmid (Hrsg.), *Kommunikationsmanagement im Wandel* (S. 271–305). Gabler. https://doi.org/10.1007/978-3-8349-9772-2_15
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215–251. <https://doi.org/10.1037/0033-295X.87.3.215>
- Eveland, W. P. Jr. (2002). News information processing as mediator of the relationship between motivations and political knowledge. *Journalism & Mass Communication Quarterly*, 79(1), 26–40. <https://doi.org/10.1177/107769900207900103>
- Eysenck, M. W. (2001). *Principles of cognitive psychology* (2. Aufl.). Psychology Press.
- Farin, T. (2023, Mai 2). *Gesunde E-Bikes: Der Motor macht Millionen fit*. Envivas. <https://www.envivas.de/magazin/im-alltag/gesund-dank-e-bikes/>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4. Aufl.). SAGE Publications Ltd.
- Fisher, W. R. (1984). Narration as a human communication paradigm: The case of public moral argument. *Communication Monographs*, 51(1), 1–22. <https://doi.org/10.1080/03637758409390180>

- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515–540.
<https://doi.org/10.1177/107769900007700304>
- Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9(2), 319–342. <https://doi.org/10.1177/1461444807075015>
- Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A., & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin*, 36(1), 82–96. <https://doi.org/10.1177/0146167209351886>
- Flesch, R. (1952). „Simplification of Flesch Reading Ease formula“: Reply. *Journal of Applied Psychology*, 36(1), 54–55. <https://doi.org/10.1037/h0051965>
- Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 722–723.
<https://doi.org/10.1145/765891.765951>
- Fogg, B. J., Marshall, J., Kameda, T., Solomon, J., Rangnekar, A., Boyd, J., & Brown, B. (2001). Web credibility research: A method for online experiments and early study results. *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, 295–296.
<https://doi.org/10.1145/634067.634242>
- Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. (2001). What makes web sites credible? A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 61–68. <https://doi.org/10.1145/365024.365037>
- Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). How do users evaluate the credibility of Web sites? A study with over 2,500 participants. *Proceedings of the 2003 conference on Designing for user experiences*, 1–15.
<https://doi.org/10.1145/997078.997097>
- Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. *CHI '99: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 80–87.
<https://doi.org/10.1145/302979.303001>

- Fördert e-Bike fahren die Gesundheit?* (2020, August 20). e-motion. Die E-Bike Experten. <https://emotion-technologies.de/e-bike-infos/vorteile/gesundheit/>
- Forgas, J. P., & Laham, S. M. (2016). Halo effects. In R. F. Pohl (Hrsg.), *Cognitive Illusions: Intriguing phenomena in judgement, thinking and memory* (2. Aufl., S. 276–290). Routledge.
- Fortier, A. (2016). Influence of need for cognition and need for cognitive closure on information behaviour: Qualitative results. *Proceedings of the Annual Conference of CAIS / Actes Du Congrès Annuel de l'ACSI*. Annual Conference of the Canadian Association for Information Science. <https://doi.org/10.29173/cais850>
- Fortier, A., & Burkell, J. (2014). Influence of need for cognition and need for cognitive closure on three information behavior orientations. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1–8. <https://doi.org/10.1002/meet.2014.14505101066>
- Forzani, E. (2018). How well can students evaluate online science information? Contributions of prior knowledge, gender, socioeconomic status, and offline reading ability. *Reading Research Quarterly*, 53(4), 385–390. <https://doi.org/10.1002/rq.218>
- Frank, S. M., Becker, M., Qi, A., Geiger, P., Frank, U. I., Rosedahl, L. A., Malloni, W. M., Sasaki, Y., Greenlee, M. W., & Watanabe, T. (2022). Efficient learning in children with rapid GABA boosting during and after training. *Current Biology*, 32(23), 5022–5030.e7. <https://doi.org/10.1016/j.cub.2022.10.021>
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42. <https://doi.org/10.1257/089533005775196732>
- Freling, T. H., Yang, Z., Saini, R., Itani, O. S., & Rashad Abualsamh, R. (2020). When poignant stories outweigh cold hard facts: A meta-analysis of the anecdotal bias. *Organizational Behavior and Human Decision Processes*, 160, 51–67. <https://doi.org/10.1016/j.obhdp.2020.01.006>
- Frerejean, J., van Strien, J. L. H., Kirschner, P. A., & Brand-Gruwel, S. (2016). Completion strategy or emphasis manipulation? Task support for teaching information problem solving. *Computers in Human Behavior*, 62, 90–104. <https://doi.org/10.1016/j.chb.2016.03.048>
- Freund, T., Kruglanski, A. W., & Shpitajzen, A. (1985). The freezing and unfreezing of impressional primacy: Effects of the need for structure and the fear of invalidity. *Personality and Social Psychology Bulletin*, 11(4), 479–487. <https://doi.org/10.1177/0146167285114013>

- Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is Chat GPT a blessing or a curse? *Frontiers in Education*, 8, 1166682.
<https://doi.org/10.3389/feduc.2023.1166682>
- Gawronski, B., & Creighton, L. A. (2013). Dual process theories. In D. E. Carlston (Hrsg.), *The Oxford handbook of social cognition* (S. 282–312). Oxford University Press.
- Gerjets, P., Kammerer, Y., & Werner, B. (2011). Measuring spontaneous and instructed evaluation processes during Web search: Integrating concurrent thinking-aloud protocols and eye-tracking data. *Learning and Instruction*, 21(2), 220–231. <https://doi.org/10.1016/j.learninstruc.2010.02.005>
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Gigerenzer, G., Todd, P., & ABC Research Group. (1999). *Simple heuristic that makes us smart*. Oxford University Press.
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, 16(1), 15–31.
<https://doi.org/10.1177/1094428112452151>
- Goldman, S. R., & Brand-Gruwel, S. (2018). Learning from multiple sources in a digital society. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Hrsg.), *International Handbook of the Learning Sciences* (S. 86–95). Routledge.
- González-Pérez, L. I., & Ramírez-Montoya, M. S. (2022). Components of education 4.0 in 21st century skills frameworks: Systematic review. *Sustainability*, 14(3), Article 3.
<https://doi.org/10.3390/su14031493>
- Google. (o. J.). Abgerufen 18. Dezember 2023, von <https://www.google.de/>
- Graf, L. K. M., Mayer, S., & Landwehr, J. R. (2017). Measuring processing fluency: One versus five items. *Journal of Consumer Psychology*, 28(3), 393–411. <https://doi.org/10.1002/jcpy.1021>
- Graham, S., & Harris, K. R. (2017). Reading and writing connections: How writing can build better readers (and vice versa). In C. Ng & B. Bartlett (Hrsg.), *Improving reading and reading engagement in the 21st century: International research and innovation* (S. 333–350). Springer.
https://doi.org/10.1007/978-981-10-4331-4_15

- Haas, A., & Unkel, J. (2017). Ranking versus reputation: Perception and effects of search result credibility. *Behaviour & Information Technology*, 36(12), 1285–1298.
<https://doi.org/10.1080/0144929X.2017.1381166>
- Hahnel, C., Goldhammer, F., Kröhne, U., & Naumann, J. (2018). The role of reading skills in the evaluation of online information gathered from search engine environments. *Computers in Human Behavior*, 78, 223–234. <https://doi.org/10.1016/j.chb.2017.10.004>
- Hamilton, M., & Hunter, J. (1998). The effect of language intensity on receiver evaluations of message, source, and topic.. In M. Allen & R. W. Preiss (Hrsg.), *Persuasion: Advances through meta-analysis* (1. Aufl., S. 99–138). Hampton Press. <https://doi.org/10.13140/2.1.1894.9925>
- Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, 44(3), 687–691. <https://doi.org/10.1016/j.jesp.2007.04.005>
- Hart, W., Adams, J. M., Alex Burton, K., Shreves, W., & Hamilton, J. C. (2012). Shaping reality vs. hiding from reality: Reconsidering the effects of trait need for closure on information search. *Journal of Research in Personality*, 46(5), 489–496. <https://doi.org/10.1016/j.jrp.2012.05.004>
- Hayes, A. F. (2022a). *Hacking PROCESS for estimation and probing of linear moderation of quadratic effects and quadratic moderation of linear effects* [Unveröffentlichtes White Paper]. Haskayne School of Business. Abgerufen am 10.02.2024, von https://haskayne.ucalgary.ca/sites/default/files/teams/106/CCRAM_TR_022_02.pdf.
- Hayes, A. F. (2022b). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3. Aufl.). Guilford Publications. <http://ebookcentral.proquest.com/lib/umainz/detail.action?docID=6809031>
- Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods*, 41(3), 924–936. <https://doi.org/10.3758/BRM.41.3.924>
- Heijne-Penninga, M., Kuks, J. B. M., Hofman, W. H. A., & Cohen-Schotanus, J. (2010). Influences of deep learning, need for cognition and preparation time on open- and closed-book test performance. *Medical Education*, 44(9), 884–891. <https://doi.org/10.1111/j.1365-2923.2010.03732.x>
- Henke, J., Leissner, L., & Möhring, W. (2019). How can journalists promote news credibility? Effects of evidences on trust and credibility. *Journalism Practice*, 14(3), 299–318.
<https://doi.org/10.1080/17512786.2019.1605839>

- Hentschel, E. (2023). Syntax. In E. Hentschel (Hrsg.), *Deutsch typologisch: Eine Einführung* (S. 147–229). J.B. Metzler. https://doi.org/10.1007/978-3-476-05945-1_3
- Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, 44(4), 1467–1484. <https://doi.org/10.1016/j.ipm.2007.10.001>
- Hobbs, R. (2006). Multiple visions of multimedia literacy: Emerging areas of synthesis. In M. C. McKenna, L. D. Labbo, R. D. Kieffer, & D. Reinking (Hrsg.), *International handbook of literacy and technology: Volume II* (1. Aufl., S. 15–28). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780203929131-3/multiple-visions-multimedia-literacy-renee-hobbs>
- Hong, H., & Kim, H. J. (2020). Antecedents and consequences of information overload in the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, 17(24), Article 24. <https://doi.org/10.3390/ijerph17249305>
- Hong, T. (2006). The influence of structural and message features on web site credibility. *Journal of the American Society for Information Science and Technology*, 57(1), 114–127. <https://doi.org/10.1002/asi.20258>
- Hope, S. M. (2007). *Searching the web: A think aloud investigation into college students' online behavior* (Publikation Nr. 3291169) [Dissertation, Purdue University]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/docview/304822866/abstract/9A54B1D1A7964161PQ/1>
- Hornikx, J. (2005). A review of experimental research on the relative persuasiveness of anecdotal, statistical, causal, and expert evidence. *Studies in Communication Sciences*, 5(1), 205–216.
- Houghton, D. C., & Grewal, R. (2000). Please, let's get an answer—any answer: Need for consumer cognitive closure. *Psychology & Marketing*, 17(11), 911–934. [https://doi.org/10.1002/1520-6793\(200011\)17:11<911::AID-MAR1>3.0.CO;2-4](https://doi.org/10.1002/1520-6793(200011)17:11<911::AID-MAR1>3.0.CO;2-4)
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. Yale University Press.
- Huang, Y., & Wang, W. (2020). When a story contradicts: Correcting health misinformation on social media through different message formats and mechanisms. *Information, Communication & Society*, 25(8), 1192–1209. <https://doi.org/10.1080/1369118X.2020.1851390>

- Huber, F., Meyer, F., & Lenzen, M. (2014). *Grundlagen der Varianzanalyse: Konzeption - Durchführung - Auswertung*. Springer-Verlag.
- Hutterer, C. (2019, September 1). *E-Bikes und der Einfluss auf das Gesundheitsverhalten von Risikogruppen*. Deutsche Zeitschrift für Sportmedizin. <https://www.zeitschrift-sportmedizin.de/e-bikes-und-der-einfluss-auf-das-gesundheitsverhalten-von-risikogruppen/>
- Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science*, 25(3), 213–218. <https://doi.org/10.1177/016555159902500305>
- Ingenkamp, K., & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik* (6. Aufl.). Beltz.
- Ji, Q., Ha, L., & Sypher, U. (2014). The role of news media use and demographic characteristics in the possibility of information overload prediction. *International Journal of Communication*, 8, 699–714.
- Johannes Gutenberg-Universität Mainz. (2023a, März 28). *Modulhandbuch Bachelor of Science Wirtschaftspädagogik*. https://download.uni-mainz.de/fb03-studienbuero/downloads/WiWi/WiP%C3%A4d_Bachelor/Studiengang/Modulhandbuch/Modulhandbuch_B.Sc.%20Wip%C3%A4d%20ab%20SoSe2023.pdf
- Johannes Gutenberg-Universität Mainz. (2023b, Oktober 10). *Modulhandbuch Master of Education Wirtschaftspädagogik*. https://download.uni-mainz.de/fb03-studienbuero/downloads/WiWi/WiP%C3%A4d_Master%20of%20Education/Studiengang/Modulhandbuch%20M%20Ed%20%20Wipaed_WiSe2324.pdf
- Johnson, F., Sbaffi, L., & Rowley, J. (2016). Students' approaches to the evaluation of digital information: Insights from their trust judgments. *British Journal of Educational Technology*, 47(6), 1243–1258. <https://doi.org/10.1111/bjet.12306>
- Johnson, K. A., & Wiedenbeck, S. (2009). Enhancing perceived credibility of citizen journalism web sites. *Journalism & Mass Communication Quarterly*, 86(2), 332–348. <https://doi.org/10.1177/107769900908600205>
- Joiner, T. A., Leveson, L., & Langfield-Smith, K. (2002). Technical language, advice understandability, and perceptions of expertise and trustworthiness: The case of the financial planner. *Australian Journal of Management*, 27(1), 25–43. <https://doi.org/10.1177/031289620202700102>

- Jonas, M., Kerwer, M., Chasiotis, A., & Rosman, T. (2024). Indicators of trustworthiness in lay-friendly research summaries: Scientificness surpasses easiness. *Public Understanding of Science*, 33(1), 37–57. <https://doi.org/10.1177/09636625231176377>
- Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development*, 45, 65–94. <https://doi.org/10.1007/BF02299613>
- Jung, J. M., & Kellaris, J. J. (2004). Cross-national differences in proneness to scarcity effects: The moderating roles of familiarity, uncertainty avoidance, and need for cognitive closure. *Psychology & Marketing*, 21(9), 739–753. <https://doi.org/10.1002/mar.20027>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kahneman, D. (2012). *Thinking, fast and slow* (1. Aufl.). Penguin Books Ltd.
- Kalisch, M., & Meier, L. (2021). Das logistische Regressionsmodell. In M. Kalisch & L. Meier (Hrsg.), *Logistische Regression: Eine anwendungsorientierte Einführung mit R* (S. 13–29). Springer Spektrum. https://doi.org/10.1007/978-3-658-34225-8_3
- Kammerer, Y., Amann, D. G., & Gerjets, P. (2015). When adults without university education search the internet for health information: The roles of internet-specific epistemic beliefs and a source evaluation intervention. *Computers in Human Behavior*, 48, 297–309. <https://doi.org/10.1016/j.chb.2015.01.045>
- Kammerer, Y., & Brand-Gruwel, S. (2020). Trainings and tools to foster source credibility evaluation during web search. In W. T. Fu & H. van Oostendorp (Hrsg.), *Understanding and improving information search. A conceptual framework* (S. 213–243). Springer. https://doi.org/10.1007/978-3-030-38825-6_11
- Kammerer, Y., Gottschling, S., & Bråten, I. (2021). The role of internet-specific justification beliefs in source evaluation and corroboration during web search on an unsettled socio-scientific issue. *Journal of Educational Computing Research*, 59(2), 342–378. <https://doi.org/10.1177/0735633120952731>
- Karlsson, M. (2010). Rituals of transparency. *Journalism Studies*, 11(4), 535–545. <https://doi.org/10.1080/14616701003638400>

- Karlsson, M., & Clerwall, C. (2018). Transparency to the rescue? *Journalism Studies*, 19(13), 1923–1933. <https://doi.org/10.1080/1461670X.2018.1492882>
- Karr-Wisniewski, P., & Lu, Y. (2010). When more is too much: Operationalizing technology overload and exploring its impact on knowledge worker productivity. *Computers in Human Behavior*, 26(5), 1061–1072. <https://doi.org/10.1016/j.chb.2010.03.008>
- Kattenbeck, M., & Elsweller, D. (2019). Understanding credibility judgements for web search snippets. *Aslib Journal of Information Management*, 71(3), 368–391. <https://doi.org/10.1108/AJIM-07-2018-0181>
- Kaufman, D. Q., Stasson, M. F., & Hart, J. W. (1999). Are the tabloids always wrong or is that just what we think? Need for cognition and perceptions of articles in print media. *Journal of Applied Social Psychology*, 29(9), 1984–2000. <https://doi.org/10.1111/j.1559-1816.1999.tb00160.x>
- Kendall, M. G., & Smith, B. B. (1939). The problem of m rankings. *The Annals of Mathematical Statistics*, 10(3), 275–287.
- Kiili, C., Bråten, I., Strømsø, H., Hagerman, M., Räikkönen, E., & Jyrkiäinen, A. (2022). Adolescents' credibility justifications when evaluating online texts. *Education and Information Technologies*, 27, 7421–7450. <https://doi.org/10.1007/s10639-022-10907-x>
- Kiili, C., Leu, D. J., Utriainen, J., Coiro, J., Kanniainen, L., Tolvanen, A., Lohvansuu, K., & Leppänen, P. H. T. (2018). Reading to learn from online information: Modeling the factor structure. *Journal of Literacy Research*, 50(3), 304–334. <https://doi.org/10.1177/1086296X18784640>
- Kiili, C., Räikkönen, E., Bråten, I., Strømsø, H., & Hagerman, M. (2023). Examining the structure of credibility evaluation when sixth graders read online texts. *Journal of Computer Assisted Learning*, 39(3), 954–969. <https://doi.org/10.1111/jcal.12779>
- Kim, K.-S., & Allen, B. (2002). Cognitive and task influences on web searching behavior. *Journal of the American Society for Information Science and Technology*, 53(2), 109–119. <https://doi.org/10.1002/asi.10014>
- Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review*, 1(6), 602–609. <https://doi.org/10.1037/h0068244>

- Klein, C. T. F., & Webster, D. M. (2000). Individual differences in argument scrutiny as motivated by need for cognitive closure. *Basic and Applied Social Psychology*, 22(2), 119–129.
https://doi.org/10.1207/S15324834BASP2202_5
- Koetsenruijter, A. W. M. (2011). Using numbers in news increases story credibility. *Newspaper Research Journal*, 32(2), 74–82. <https://doi.org/10.1177/073953291103200207>
- Koh, Y. J., & Sundar, S. S. (2010). Effects of specialization in computers, web sites, and web agents on e-commerce trust. *International Journal of Human-Computer Studies*, 68(12), 899–912.
<https://doi.org/10.1016/j.ijhcs.2010.08.002>
- Koltay, T. (2017). Information overload in a data-intensive world. In A. J. Schuster (Hrsg.), *Understanding information: From the Big Bang to Big Data* (S. 197–217). Springer.
https://doi.org/10.1007/978-3-319-59090-5_10
- König, L., & Jucks, R. (2019a). Hot topics in science communication: Aggressive language decreases trustworthiness and credibility in scientific debates. *Public Understanding of Science*, 28(4), 401–416. <https://doi.org/10.1177/0963662519833903>
- König, L., & Jucks, R. (2019b). Influence of enthusiastic language on the credibility of health information and the trustworthiness of science communicators: Insights from a between-subject web-based experiment. *Interactive Journal of Medical Research*, 8(3), e13619.
<https://doi.org/10.2196/13619>
- Kosic, A., Kruglanski, A. W., Pierro, A., & Mannetti, L. (2004). The social cognition of immigrants' acculturation: Effects of the need for closure and the reference group at entry. *Journal of Personality and Social Psychology*, 86(6), 796–813. <https://doi.org/10.1037/0022-3514.86.6.796>
- Kossowska, M. (2007). Motivation towards closure and cognitive processes: An individual differences approach. *Personality and Individual Differences*, 43(8), 2149–2158.
<https://doi.org/10.1016/j.paid.2007.06.027>
- Kossowska, M., & Bar-Tal, Y. (2013). Need for closure and heuristic information processing: The moderating role of the ability to achieve the need for closure. *British Journal of Psychology*, 104(4), 457–480. <https://doi.org/10.1111/bjop.12001>
- Kramp, L., & Weichert, S. (2022). *Wie resilient sind wir im Netz? Abschlussbericht zur Grundlagenstudie „Digitale Resilienz in der Mediennutzung“* (S. 38). VOCER Institut für Digitale Resilienz.

- https://digitale-resilienz.org/wp-content/uploads/2023/02/Abschlussbericht_Digitale-Resilienz-Mediennutzung_website.pdf
- Kroeber-Riel, W., & Esch, F.-R. (2015). *Strategie und Technik der Werbung: Verhaltenswissenschaftliche und neurowissenschaftliche Erkenntnisse* (8. Aufl.). Kohlhammer.
- Kruglanski, A. W. (2013). *The psychology of closed mindedness*. Psychology Press.
- Kruglanski, A. W., & Fishman, S. (2009). The need for cognitive closure. In M. R. Leary & R. H. Hoyle, *Handbook of individual differences in social behavior* (S. 343–353). The Guilford Press.
- Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19(5), 448–468. [https://doi.org/10.1016/0022-1031\(83\)90022-7](https://doi.org/10.1016/0022-1031(83)90022-7)
- Kruglanski, A. W., & Webster, D. M. (1991). Group members' reactions to opinion deviates and conformists at varying degrees of proximity to decision deadline and of environmental noise. *Journal of Personality and Social Psychology*, 61(2), 212–225. <https://doi.org/10.1037/0022-3514.61.2.212>
- Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: „Seizing“ and „freezing.“ *Psychological Review*, 103(2), 263–283. <https://doi.org/10.1037/0033-295X.103.2.263>
- Kubbe, I. (2016). Gütekriterien experimenteller Forschung. In I. Kubbe (Hrsg.), *Experimente in der Politikwissenschaft: Eine methodische Einführung* (S. 61–76). Springer VS.
https://doi.org/10.1007/978-3-658-09424-9_4
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361–371.
[https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5<361::AID-ASI6>3.0.CO;2-%23](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-%23)
- Kuhlthau, C. C. (2004). *Seeking meaning: A process approach to library and information services*. (2. Aufl.). Libraries Unlimited. <http://informationr.net/ir/reviews/revs129.html>
- Kuhn, S., Müller, N., Kirchgässner, E., Ulzheimer, L., & Deutsch, K. L. (2020). Digital skills for medical students – qualitative evaluation of the curriculum 4.0 “Medicine in the digital age”. *GMS Journal for Medical Education*, 37(6), Doc60. <https://doi.org/10.3205/zma001353>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>

- Landwehr, J. R., Labroo, A. A., & Herrmann, A. (2011). Gut liking for the ordinary: Incorporating design fluency improves automobile sales forecasts. *Marketing Science*, 30(3), 416–429.
<https://doi.org/10.1287/mksc.1110.0633>
- Landwehr, J. R., Wentzel, D., & Herrmann, A. (2013). Product design for the long run: Consumer responses to typical and atypical designs at different stages of exposure. *Journal of Marketing*, 77(5), 92–107. <https://doi.org/10.1509/jm.11.0286>
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, 50(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
- Lauterbach, B. (2020). *Konzeption und Entwicklung eines Feedbacksystems als Teil eines digitalen Lehr-Lernarrangements zur Förderung von Critical Online Reasoning (COR)* [Unveröffentlichte Masterthesis].
- le Roux, D. B., & Parry, D. A. (2017). In-lecture media use and academic performance: Does subject area matter? *Computers in Human Behavior*, 77, 86–94.
<https://doi.org/10.1016/j.chb.2017.08.030>
- Lei, P.-L., Lin, S. S. J., & Sun, C.-T. (2013). Effect of reading ability and internet experience on keyword-based image search. *Journal of Educational Technology & Society*, 16(2), 151–162.
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093–1096.
<https://doi.org/10.1016/j.jesp.2010.05.025>
- Lipowsky, F. (2020). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–118). Springer. https://doi.org/10.1007/978-3-662-61403-7_4
- Liu, Z. (2004). Perceptions of credibility of scholarly information on the web. *Information Processing & Management*, 40(6), 1027–1038. [https://doi.org/10.1016/S0306-4573\(03\)00064-5](https://doi.org/10.1016/S0306-4573(03)00064-5)
- Liu, Z., & Huang, X. (2005). Evaluating the credibility of scholarly information on the web: A cross cultural study. *The International Information & Library Review*, 37(2), 99–106.
<https://doi.org/10.1016/j.iilr.2005.05.004>
- López-González, H., Sosa, L., Sánchez, L., & Faure-Carvallo, A. (2023). Media and information literacy and critical thinking: A systematic review. *Revista Latina de Comunicación Social, English Ed.*, 81, 399–422. <https://doi.org/10.4185/RLCS-2023-1939>

- Lorenz-Spreen, P., Mønsted, B. M., Hövel, P., & Lehmann, S. (2019). Accelerating dynamics of collective attention. *Nature Communications*, 10, 1759. <https://doi.org/10.1038/s41467-019-09311-w>
- Lucassen, T., Muilwijk, R., Noordzij, M. L., & Schraagen, J. M. (2013). Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology*, 64(2), 254–264. <https://doi.org/10.1002/asi.22743>
- Lucassen, T., & Schraagen, J. M. (2011). Factual accuracy and trust in information: The role of expertise. *Journal of the American Society for Information Science and Technology*, 62(7), 1232–1242. <https://doi.org/10.1002/asi.21545>
- Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*, 40(3), 26–29. <https://doi.org/10.1108/LHTN-01-2023-0009>
- Lund, T. (2012). Combining qualitative and quantitative approaches: Some arguments for mixed methods research. *Scandinavian Journal of Educational Research*, 56(2), 155–165. <https://doi.org/10.1080/00313831.2011.568674>
- Mack, L. (2021). Methoden auswählen und entwickeln. Wie Sie von der Leistungsbeschreibung zu den Methoden gelangen. In J. Glunde, L. Mack, M. Mohaupt, & G. Schüller (Hrsg.), *Projektmanagement für Promovierende: Werkzeuge und Methoden für eine erfolgreiche Doktorarbeit* (S. 93–101). Springer Gabler. https://doi.org/10.1007/978-3-662-63732-6_10
- Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, 61(1), 13–25. <https://doi.org/10.1037/0022-3514.61.1.13>
- Maule, A. J., Hockey, G. R. J., & Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: Changes in affective state and information processing strategy. *Acta Psychologica*, 104(3), 283–301. [https://doi.org/10.1016/S0001-6918\(00\)00033-0](https://doi.org/10.1016/S0001-6918(00)00033-0)
- Maurer, M., Quiring, O., & Schemer, C. (2018). Media effects on positive and negative learning. In O. Zlatkin-Troitschanskaia, G. Wittum, & A. Dengel (Hrsg.), *Positive learning in the age of information: A blessing or a curse?* (S. 197–208). Springer VS. https://doi.org/10.1007/978-3-658-19567-0_11
- Maurer, M., Schemer, C., Zlatkin-Troitschanskaia, O., & Jitomirski, J. (2020). Positive and negative media effects on university students' learning: Preliminary findings and a research program. In

- O. Zlatkin-Troitschanskaia (Hrsg.), *Frontiers and advances in positive learning in the age of information (PLATO)* (S. 109–119). Springer. https://doi.org/10.1007/978-3-030-26578-6_8
- Mayring, P., & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 633–648). Springer VS.
https://doi.org/10.1007/978-3-658-21308-4_42
- McCroskey, J. C., & Mehrley, R. S. (1969). The effects of disorganization and nonfluency on attitude change and source credibility. *Speech Monographs*, 36(1), 13–21.
<https://doi.org/10.1080/03637756909375604>
- McEvily, W. J., Weber, R. A., Bicchieri, C., & Ho, V. T. (2002). *Can groups be trusted? An experimental study of collective trust* (Carnegie Mellon Behavioral Decision Research Working Paper No. 308). <http://dx.doi.org/10.2139/ssrn.323223>
- McGrew, S. (2020). Learning to evaluate: An intervention in civic online reasoning. *Computers & Education*, 145, 103711. <https://doi.org/10.1016/j.compedu.2019.103711>
- McGrew, S. (2021a). Internet or archive? Expertise in searching for digital sources on a contentious historical question. *Cognition and Instruction*, 40(4), 488–516.
<https://doi.org/10.1080/07370008.2021.1908288>
- McGrew, S. (2021b). Skipping the source and checking the contents: An in-depth look at students' approaches to web evaluation. *Computers in the Schools*, 38(2), 75–97.
<https://doi.org/10.1080/07380569.2021.1912541>
- McGrew, S., Breakstone, J., Ortega, T., Smith, M., & Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & Research in Social Education*, 46(2), 165–193. <https://doi.org/10.1080/00933104.2017.1416320>
- McKnight, D. H., & Kacmar, C. J. (2007). Factors and effects of information credibility. *Proceedings of the ninth international conference on Electronic commerce*, 423–432.
<https://doi.org/10.1145/1282100.1282180>
- Media Bias/Fact Check News. (2023, Juli 12). Media Bias/Fact Check. <https://mediabiasfactcheck.com/>
- Medienpolitik: Gatekeeper. (o. J.). Bundeszentrale für politische Bildung. Abgerufen 19. Oktober 2023, von <https://www.bpb.de/themen/medien-journalismus/medienpolitik/500676/gatekeeper/>

- Medienpolitik: Web 2.0.* (o. J.). Bundeszentrale für politische Bildung. Abgerufen 22. Juli 2023, von <https://www.bpb.de/themen/medien-journalismus/medienpolitik/500735/web-2-0/>
- Medizinische Fakultätentag. (o. J.). *VIII.1. Medizinisch-wissenschaftliche Fertigkeiten*. Nationaler Kompetenzbasierter Lernzielkatalog Medizin. Version 2.0. Abgerufen 26. Oktober 2023, von <https://nklm.de/zend/objective/list/orderBy/@objectivePosition/studiengang/PF2/zeitsemester/2021/fachsemester/VIII.1.%20Medizinisch-wissenschaftliche%20Fertigkeiten>
- Meijnders, A. L., Midden, C. J. H., & Wilke, H. A. M. (2001). Communications about environmental risks and risk-reducing behavior: The impact of fear on information processing. *Journal of Applied Social Psychology*, 31(4), 754–777. <https://doi.org/10.1111/j.1559-1816.2001.tb01412.x>
- Meinert, J., & Krämer, N. C. (2022). How the expertise heuristic accelerates decision-making and credibility judgments in social media by means of effort reduction. *PLOS ONE*, 17(3), e0264428. <https://doi.org/10.1371/journal.pone.0264428>
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078–2091. <https://doi.org/10.1002/asi.20672>
- Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220. <https://doi.org/10.1016/j.pragma.2013.07.012>
- Metzger, M. J., & Flanagin, A. J. (2015). Psychological approaches to credibility assessment online. In S. S. Sundar (Hrsg.), *The handbook of the psychology of communication technology* (1. Aufl., S. 445–466). Wiley. <https://doi.org/10.1002/9781118426456.ch20>
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & Mccann, R. M. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. *Annals of the International Communication Association*, 27(1), 293–335. <https://doi.org/10.1080/23808985.2003.11679029>
- Metzger, M. J., Flanagin, A. J., Markov, A., Grossman, R., & Bulger, M. (2015). Believing the unbelievable: Understanding young people's information literacy beliefs and practices in the United States. *Journal of Children and Media*, 9(3), 325–348. <https://doi.org/10.1080/17482798.2015.1056817>

- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413–439. <https://doi.org/10.1111/j.1460-2466.2010.01488.x>
- Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student web use, perceptions of information credibility, and verification behavior. *Computers & Education*, 41(3), 271–290. [https://doi.org/10.1016/S0360-1315\(03\)00049-6](https://doi.org/10.1016/S0360-1315(03)00049-6)
- Meyer, L. (2019, Juni 24). *E-Bikes: So profitiert Ihre Gesundheit*. Focus Gesundheit. <https://focus-gesundheit.de/magazin/gesundheitstipps/e-bikes-so-profitiert-ihre-gesundheit>
- Middaugh, E. (2018). Civic media literacy in a transmedia world: Balancing personal experience, factual accuracy and emotional appeal as media consumers and circulators. *Journal of Media Literacy Education*, 10(2), 33–52.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <https://doi.org/10.1037/h0043158>
- Miller, M. D., & Levine, T. R. (2019). Persuasion. In D. W. Stacks, M. B. Salwen, & K. C. Eichhorn (Hrsg.), *An integrated approach to communication theory and research* (3. Aufl., S. 261–275). Routledge.
- Molero, D., Zlatkin-Troitschanskaia, O., Nagel, M.-T., Brückner, S., Schmidt, S., & Shavelson, R. J. (2020). Assessing university students' critical online reasoning ability: A conceptual and assessment framework with preliminary evidence. *Frontiers in Education*, 5, 557843. <https://doi.org/10.18452/22320>
- Moosbrugger, H., & Kelava, A. (2008). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 7–26). Springer. https://doi.org/10.1007/978-3-540-71635-8_2
- Moschner, B., & Anschütz, A. (2010). Kombination und Integration von qualitativen und quantitativen Forschungsmethoden in einem interdisziplinären Forschungsprojekt. *Didaktik der Informatik - Möglichkeiten empirischer Forschungsmethoden und Perspektiven der Fachdidaktik*, 11–20. <https://dl.gi.de/handle/20.500.12116/19633>
- Münchow, H., Richter, T., von der Mühlen, S., & Schmid, S. (2019). The ability to evaluate arguments in scientific texts: Measurement, cognitive processes, nomological network, and relevance for

- academic success at the university. *British Journal of Educational Psychology*, 89(3), 501–523. <https://doi.org/10.1111/bjep.12298>
- Münchow, H., Tiffin-Richards, S. P., Fleischmann, L., Pieschl, S., & Richter, T. (2023). Promoting students' argument comprehension and evaluation skills: Implementation of two training interventions in higher education. *Zeitschrift Für Erziehungswissenschaft*, 26, 703–725. <https://doi.org/10.1007/s11618-023-01147-x>
- Murphy, C., Klotz, A. C., & Kreiner, G. E. (2017). Blue skies and black boxes: The promise (and practice) of grounded theory in human resource management research. *Human Resource Management Review*, 27(2), 291–305. <https://doi.org/10.1016/j.hrmr.2016.08.006>
- Nagel, M.-T., Schäfer, S., Zlatkin-Troitschanskaia, O., Schemer, C., Maurer, M., Molerov, D., Schmidt, S., & Brückner, S. (2020). How do university students' web search behavior, website characteristics, and the interaction of both influence students' critical online reasoning? *Frontiers in Education*, 5, 565062. <https://doi.org/10.3389/educ.2020.565062>
- Nagel, M.-T., Zlatkin-Troitschanskaia, O., & Fischer, J. (2022). Validation of newly developed tasks for the assessment of generic critical online reasoning (COR) of university students and graduates. *Frontiers in Education*, 7, 914857. <https://doi.org/10.3389/educ.2022.914857>
- Nawratil, U. (1999). Glaubwürdigkeit als Faktor im Prozess medialer Kommunikation. In P. Rössler & W. Wirth (Hrsg.), *Glaubwürdigkeit im Internet. Fragestellungen, Modelle, empirische Befunde*. (Bd. 32, S. 15–32). Reinhard Fischer.
- Neuberger, C. (2007). Das Ende des »Gatekeeper«-Zeitalters. In K. Lehmann & M. Schetsche (Hrsg.), *Die Google-Gesellschaft – Vom digitalen Wandel des Wissens* (S. 205–212). transcript. <https://mediarep.org/handle/doc/12792>
- Neuert, C. E., & Lenzner, T. (2019). Die Ergänzung kognitiver Interviews um Eye Tracking. In N. Me-nold & T. Wolbring (Hrsg.), *Qualitätssicherung sozialwissenschaftlicher Erhebungsinstrumente* (S. 195–224). Springer VS. https://doi.org/10.1007/978-3-658-24517-7_7
- Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly*, 46(3), 273–304. <https://doi.org/10.1598/RRQ.46.3.4>

- Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers & Education*, 61, 59–76. <https://doi.org/10.1016/j.compedu.2012.08.013>
- Novin, A., & Meyers, E. (2017). Making sense of conflicting science information: Exploring bias in the search engine result page. *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, 175–184. <https://doi.org/10.1145/3020165.3020185>
- Nussbaum, E. M. (2005). The effect of goal instructions and need for cognition on interactive argumentation. *Contemporary Educational Psychology*, 30(3), 286–313. <https://doi.org/10.1016/j.cedpsych.2004.11.002>
- Nwb Datenbank. (o. J.). Abgerufen 31. Januar 2024, von <https://datenbank.nwb.de/>
- O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5), 673–690. <https://doi.org/10.1007/s11135-006-9018-6>
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1609406919899220. <https://doi.org/10.1177/1609406919899220>
- OECD. (2022). *Does higher education teach students to think critically?* (D. van Damme & D. Zahner, Hrsg.). <https://www.oecd-ilibrary.org/content/publication/cc9fa6aa-en>
- Online Umfrage erstellen. (o. J.). [Startseite]. Unipark. Abgerufen 25. Oktober 2023, von <https://www.unipark.com/>
- Oswald, B. (2013). Vom Produkt zum Prozess. In L. Kramp, L. Novy, D. Ballwieser, & K. Wenzlaff (Hrsg.), *Journalismus in der digitalen Moderne: Einsichten—Ansichten—Aussichten* (S. 63–79). Springer VS. https://doi.org/10.1007/978-3-658-01144-4_4
- Oswald, M., & Grosjean, S. (2004). Confirmation bias. In R. F. Pohl (Hrsg.), *Cognitive illusions. A handbook on fallacies and biases in thinking, judgement and memory* (S. 79–96). Psychology Press.
- Otto, P., Müller-Brehm, J., & Puntschuh, M. (2020). *Digitalisierung* (344; Informationen zur politischen Bildung / izpb, S. 84). Bundeszentrale für politische Bildung/bpb. <https://www.bpb.de/shop/zeitschriften/izpb/digitalisierung-344/>

- Oxford Learner's Dictionaries. (o.J. - a). Abstract. In *oxfordlearnersdictionaries.com* Online-Wörterbuch. Abgerufen am 08. Februar 2024, von https://www.oxfordlearnersdictionaries.com/definition/english/abstract_1?q=abstract
- Oxford Learner's Dictionaries. (o.J. - b). Credibility. In *oxfordlearnersdictionaries.com* Online-Wörterbuch. Abgerufen am 08. Februar 2024, von <https://www.oxfordlearnersdictionaries.com/definition/english/credibility?q=credibility>
- Oxford Learner's Dictionaries. (o.J. - c). Trustworthiness. In *oxfordlearnersdictionaries.com* Online-Wörterbuch. Abgerufen am 08. Februar 2024, von <https://www.oxfordlearnersdictionaries.com/definition/english/trustworthiness?q=trustworthiness>
- Paivio, A. (1991). *Images in mind: The evolution of a theory*. Harvester Wheatsheaf.
- Pardoe, I. (2021). *Applied regression modeling* (3. Aufl.). John Wiley & Sons.
- Paus, I., & Börsch-Supan, J. (2020). *Die Jugend in der Infodemie. Eine repräsentative Befragung zum Umgang junger Menschen in Deutschland mit Falschnachrichten während der Coronakrise* (S. 16). Vodafone Stiftung Deutschland. <https://www.vodafone-stiftung.de/wp-content/uploads/2020/12/Studie-Vodafone-Stiftung-Umgang-mit-Falschnachrichten.pdf>
- Peng, D., & Yu, Z. (2022). A literature review of digital literacy over two decades. *Education Research International*, 2022, e2533413. <https://doi.org/10.1155/2022/2533413>
- Petty, R. E., Briñol, P., Loersch, C., & McCaslin, M. J. (2009). The need for cognition. In M. R. Leary & R. H. Hoyle (Hrsg.), *Handbook of individual differences in social behavior* (S. 318–329). The Guilford Press.
- Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, 46(1), 69–81. <https://doi.org/10.1037/0022-3514.46.1.69>
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In R. E. Petty & J. T. Cacioppo (Hrsg.), *Communication and persuasion: Central and peripheral routes to attitude change* (S. 123–205). Springer. https://doi.org/10.1007/978-1-4612-4964-1_1
- Petty, R. E., & Cacioppo, J. T. (1996). *Attitudes and persuasion: Classic and contemporary approaches* (1. Aufl.). Westview Press. <https://doi.org/10.4324/9780429502156>

- Pravossoudovitch, K., Cury, F., Young, S. G., & Elliot, A. J. (2014). Is red the colour of danger? Testing an implicit red-danger association. *Ergonomics*, 57(4).
<https://doi.org/10.1080/00140139.2014.889220>
- PubMed. (o. J.). PubMed. Abgerufen 31. Januar 2024, von <https://pubmed.ncbi.nlm.nih.gov/>
- Qadir, J. (2023). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. *2023 IEEE Global Engineering Education Conference (EDUCON)*, 1–9.
<https://doi.org/10.1109/EDUCON54358.2023.10125121>
- Raes, A., Schellens, T., De Wever, B., & Vanderhoven, E. (2012). Scaffolding information problem solving in web-based collaborative inquiry learning. *Computers & Education*, 59(1), 82–94.
<https://doi.org/10.1016/j.compedu.2011.11.010>
- Rains, S. A., & Karmikel, C. D. (2009). Health information-seeking and perceptions of website credibility: Examining web-use orientation, message characteristics, and structural features of websites. *Computers in Human Behavior*, 25(2), 544–553.
<https://doi.org/10.1016/j.chb.2008.11.005>
- Rainsberger, L. (2022). Der iii-Mensch. In L. Rainsberger (Hrsg.), *Der moderne Kunde – das PHANTOM: Kunden auf der Flucht: Wie der Vertrieb auf ein radikal neues Kaufverhalten reagieren muss* (S. 21–38). Springer Gabler. https://doi.org/10.1007/978-3-658-36358-1_2
- Rasul, T., Nair, S., Kalendra, D., Robin, M., Santini, F., Ladeira, W., Sun, M., Day, I., Rather, A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching*, 6(1).
<https://doi.org/10.37074/jalt.2023.6.1.29>
- Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8(3), 338–342. <https://doi.org/10.1006/ccog.1999.0386>
- Reidl, A. (2018, August 6). *Pedelec: Ist Pedelec fahren gesünder als Radeln ohne Motor?* Die Zeit.
<https://www.zeit.de/mobilitaet/2018-08/pedelec-gesundheit-alter-bewegung-radfahren>
- Renkl, A. (2020). Wissenserwerb. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 3–24). Springer. https://doi.org/10.1007/978-3-662-61403-7_1
- Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. A. (1995). The well-built clinical question: A key to evidence-based decisions. *ACP Journal Club*, 123(3), A12-3.
<https://doi.org/10.7326/ACPJC-1995-123-3-A12>

- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the Web. *Journal of the American Society for Information Science and Technology*, 53(2), 145–161.
<https://doi.org/10.1002/asi.10017>
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, 41(1), 307–364.
<https://doi.org/10.1002/aris.2007.1440410114>
- Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing & Management*, 44(1), 386–399. <https://doi.org/10.1016/j.ipm.2007.02.003>
- Roetzel, P. G. (2019). Information overload in the information age: A review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research*, 12(2), 479–522.
- Rossi, R. J. (2022). *Applied biostatistics for the health sciences* (2. Aufl.). John Wiley & Sons.
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23(3), 393–404.
<https://doi.org/10.5465/amr.1998.926617>
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5(6), 359–371. <https://doi.org/10.1111/j.1751-9004.2011.00355.x>
- Saleh, H. F. (2016). Developing new media credibility scale: A multidimensional perspective. *International Journal of Humanities and Social Sciences*, 10(4), 1351–1364.
<https://doi.org/10.5281/zenodo.1124519>
- Sanders, J., Gass, R., Wiseman, R., & Bruschke, J. (1992). Ethnic comparison and measurement of argumentativeness, verbal aggressiveness, and need for cognition. *Communication Reports*, 5(1), 50–56. <https://doi.org/10.1080/08934219209367543>
- Schaewitz, L., Kluck, J. P., Klösters, L., & Krämer, N. C. (2020). When is disinformation (in)credible? Experimental findings on message characteristics and individual differences. *Mass Communication and Society*, 23(4), 484–509. <https://doi.org/10.1080/15205436.2020.1716983>
- Scharrer, L., Britt, M. A., Stadtler, M., & Bromme, R. (2013). Easy to understand but difficult to decide: Information comprehensibility and controversiality affect laypeople's science-based decisions. *Discourse Processes*, 50(6), 361–387. <https://doi.org/10.1080/0163853X.2013.813835>

- Scharrer, L., Bromme, R., & Stadtler, M. (2021). Information easiness affects non-experts' evaluation of scientific claims about which they hold prior beliefs. *Frontiers in Psychology*, 12, 678313. <https://doi.org/10.3389/fpsyg.2021.678313>
- Scharrer, L., Pape, V., & Stadtler, M. (2022). Watch out: Fake! How warning labels affect laypeople's evaluation of simplified scientific misinformation. *Discourse Processes*, 59(8), 575–590. <https://doi.org/10.1080/0163853X.2022.2096364>
- Scheiter, K., Gerjets, P., Vollmann, B., & Catrambone, R. (2009). The impact of learner characteristics on information utilization strategies, cognitive load experienced, and performance in hypermedia learning. *Learning and Instruction*, 19(5), 387–401. <https://doi.org/10.1016/j.learninstruc.2009.02.004>
- Schlink, S. (2009). *Persönlichkeit entscheidet: Der Einfluss des Bedürfnisses nach kognitiver Geschlossenheit auf Entscheidungen unter Unsicherheit* [Dissertation, Universität Trier]. OPUS. <https://doi.org/10.25353/ubtr-xxxx-8761-1ba8/>
- Schlink, S. A., & Walther, E. (2007). Kurz und gut: Eine deutsche Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit. *Zeitschrift für Sozialpsychologie*, 38(3), 153–161. <https://doi.org/10.1024/0044-3514.38.3.153>
- Schmeck, R. R., & Geisler-Brenstein, E. (1989). Individual differences that affect the way students approach learning. *Learning and Individual Differences*, 1(1), 85–124. [https://doi.org/10.1016/1041-6080\(89\)90011-3](https://doi.org/10.1016/1041-6080(89)90011-3)
- Schmidt, C. (2020). Empirische Befunde zum formativen Assessment. In C. Schmidt, *Formatives Assessment in der Grundschule. Konzepte, Einschätzungen der Lehrkräfte und Zusammenhänge* (S. 63–81). Springer VS. https://doi.org/10.1007/978-3-658-26921-0_4
- Schmitt, M. (2022). State. In *Dorsch Lexikon der Psychologie* (zuletzt geändert am 07.10.2022). <https://dorsch.hogrefe.com/stichwort/state>
- Schneider, S. C. (1987). Information overload: Causes and consequences. *Human Systems Management*, 7(2), 143–153. <https://doi.org/10.3233/HSM-1987-7207>
- Schöll, P. (2023). *Die Flesch-Formel. Gelesen ist nicht verstanden*. Fleschindex. <https://fleschindex.de/formel/>
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112(3), 610–628. <https://doi.org/10.1037/0033-295X.112.3.610>

- Schreier, M., & Odağ, Ö. (2020). Mixed Methods. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie: Band 2: Designs und Verfahren* (2. Aufl., S. 159–184). Springer Fachmedien. https://doi.org/10.1007/978-3-658-26887-9_22
- Schwandner, M. (2019, August 2). *So gesund ist das E-Bike wirklich*. Quarks. <https://www.quarks.de/technik/mobilitaet/so-gesund-ist-das-e-bike-wirklich/>
- Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida, & J. A. Bargh (Hrsg.), *APA handbook of personality and social psychology: Attitudes and social cognition* (1. Aufl., S. 203–229). APA. <https://doi.org/10.1037/14341-006>
- Self, C. C., & Roberts, C. (2019). Credibility. In D. Stacks, M. Salwen, & K. Eichhorn (Hrsg.), *An integrated approach to communication theory and research* (3. Aufl., S. 435–446). Routledge.
- Shah, J. Y., Kruglanski, A. W., & Thompson, E. P. (1998). Membership has its (epistemic) rewards: Need for closure effects on in-group bias. *Journal of Personality and Social Psychology*, 75(2), 383–393. <https://doi.org/10.1037/0022-3514.75.2.383>
- Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Hrsg.), *Handbook of writing research* (S. 171–183). The Guilford Press.
- Sharit, J., Hernández, M. A., Czaja, S. J., & Pirolli, P. (2008). Investigating the roles of knowledge and cognitive abilities in older adult information seeking on the web. *ACM Transactions on Computer-Human Interaction*, 15(1), 1–25. <https://doi.org/10.1145/1352782.1352785>
- Sharit, J., Taha, J., Berkowsky, R. W., Profita, H., & Czaja, S. J. (2015). Online information search performance and search strategies in a health problem-solving scenario. *Journal of Cognitive Engineering and Decision Making*, 9(3), 211–228. <https://doi.org/10.1177/1555343415583747>
- Sharp, H. Jr., & McClung, T. (1966). Effects of organization on the speaker's ethos. *Speech Monographs*, 33(2), 182–183. <https://doi.org/10.1080/03637756609375495>
- Shenk, D. (1997). *Data smog: Surviving the information glut*. GB Gardners Books.
- Sillence, E., Briggs, P., Harris, P. R., & Fishwick, L. (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine*, 64(9), 1853–1862. <https://doi.org/10.1016/j.socscimed.2007.01.012>
- Simons, R. (1988). *Using management control systems to shape strategy* (Working Paper).

- So gut ist E-Bike fahren für Deine Gesundheit!* (2019, Dezember 16). EBike24. <https://www.elektrofahrrad24.de/blog/e-biken-fuer-die-gesundheit>
- Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F. S., & Natalicio, D. S. (1971). The state-trait anxiety inventory. *Interamerican Journal of Psychology*, 5(3, 4). <https://doi.org/10.30849/rip/ijp.v5i3%20&%204.620>
- Spink, A., Wolfram, D., Jansen, M. B. J., & Saracevic, T. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology*, 52(3), 226–234. [https://doi.org/10.1002/1097-4571\(2000\)9999:9999<::AID-ASI1591>3.0.CO;2-R](https://doi.org/10.1002/1097-4571(2000)9999:9999<::AID-ASI1591>3.0.CO;2-R)
- Steindl, C., Jonas, E., Sittenthaler, S., Traut-Mattausch, E., & Greenberg, J. (2015). Understanding psychological reactance. *Zeitschrift für Psychologie*, 223(4), 205–214. <https://doi.org/10.1027/2151-2604/a000222>
- Steyer, R., Mayer, A., Geiser, C., & Cole, D. A. (2015). A theory of states and traits—Revised. *Annual Review of Clinical Psychology*, 11(1), 71–98. <https://doi.org/10.1146/annurev-clinpsy-032813-153719>
- Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality*, 13(5), 389–408. [https://doi.org/10.1002/\(SICI\)1099-0984\(199909/10\)13:5<389::AID-PER361>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<389::AID-PER361>3.0.CO;2-A)
- Stiff, J. B., & Mongeau, P. A. (2003). *Persuasive communication* (2. Aufl.). The Guilford Press.
- Stoltenberg, C. D., & Davis, C. S. (1988). Career and study skills information: Who says what can alter message processing. *Journal of Social and Clinical Psychology*, 6(1), 38–52. <https://doi.org/10.1521/jscp.1988.6.1.38>
- Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine*, 18(10), 1099–1104. <https://doi.org/10.1111/j.1553-2712.2011.01185.x>
- Stripling, B. K., & Pitts, J. M. (1988). *Brainstorms and blueprints: Teaching library research as a thinking process*. Libraries Unlimited.
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Hrsg.), *Digital media, youth, and credibility* (S. 73–100). The MIT Press.
- Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication*, 51(1), 52–72. <https://doi.org/10.1111/j.1460-2466.2001.tb02872.x>

- Sussman, S. A., & Zhang, W. (2008). Capitalizing on content: Information adoption in two online communities. *Journal of the Association for Information Systems*, 9(2).
<https://doi.org/10.17705/1jais.00149>
- Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10(3), 321–351. [https://doi.org/10.1016/S0953-5438\(98\)00013-7](https://doi.org/10.1016/S0953-5438(98)00013-7)
- Sutherland-Smith, W. (2002). Weaving the literacy web: Changes in reading from page to screen. *The Reading Teacher*, 55(7), 662–669.
- Swain, M. R., & Haka, S. F. (2000). Effects of information load on capital budgeting decisions. *Behavioral Research in Accounting*, 12, 171–198.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Syarief, A., & Hibino, H. (2007). Looking into the credibility of appearance: Exploring the role of color in interface aesthetics and how it affects our perception on system's credibility. *ITB Journal of Visual Art and Design*, 1(1), 1–9. <https://doi.org/10.5614/itbj.vad.2007.1.1.1>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5. Aufl.). Pearson Education.
- Taboada, M. (2016). Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics*, 2, 325–347. <https://doi.org/10.1146/annurev-linguistics-011415-040518>
- Tarchi, C., & Mason, L. (2020). Effects of critical thinking on multiple-document comprehension. *European Journal of Psychology of Education*, 35(2), 289–313. <https://doi.org/10.1007/s10212-019-00426-8>
- Tate, M. A. (2019). *Web wisdom: How to evaluate and create information quality on the web* (3. Aufl.). Taylor & Francis Group.
- Thomm, E., & Bromme, R. (2012). “It should at least seem scientific!” Textual features of “scientificness” and their impact on lay assessments of online information. *Science Education*, 96(2), 187–211. <https://doi.org/10.1002/sce.20480>
- Thon, F. M., & Jucks, R. (2017). Believing in expertise: How authors' credentials and language use influence the credibility of online health information. *Health Communication*, 32(7), 828–836. <https://doi.org/10.1080/10410236.2016.1172296>
- Toulmin, S. E. (2003). *The uses of argument* (2. Aufl.). Cambridge University Press.

- Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, 3(6), 449–459. [https://doi.org/10.1016/S0022-5371\(64\)80015-3](https://doi.org/10.1016/S0022-5371(64)80015-3)
- Trepte, S., Reinecke, L., Schäwel, J., Leplow, B., & Salisch, M. von. (2021). *Medienpsychologie*. Kohlhammer Verlag. <http://ebookcentral.proquest.com/lib/umainz/detail.action?docID=6579036>
- Trumbo, C. W. (1999). Heuristic-systematic information processing and risk judgment. *Risk Analysis*, 19(3), 391–400. <https://doi.org/10.1111/j.1539-6924.1999.tb00415.x>
- Tseng, S., & Fogg, B. J. (1999). Credibility and computing technology. *Communications of the ACM*, 42(5), 39–44. <https://doi.org/10.1145/301353.301402>
- Unkel, J. (2019). *Informationsselektion mit Suchmaschinen: Wahrnehmung und Auswahl von Suchresultaten* (1. Aufl.). Nomos Verlag.
- Uslu, N. A., & Durak, H. Y. (2022). The relationships between university students' information-seeking strategies, social-media specific epistemological beliefs, information literacy, and personality traits. *Library & Information Science Research*, 44(2), 101155. <https://doi.org/10.1016/j.lisr.2022.101155>
- Vakkari, P. (2001). A theory of the task-based information retrieval process: A summary and generalisation of a longitudinal study. *Journal of Documentation*, 57(1), 44–60. <https://doi.org/10.1108/EUM0000000007075>
- Valavičienė, N., Penkauskienė, D., Pivorienė, J., Railienė, A., Merfeldaitė, O., Sadauskas, J., Jegelevičienė, V., & Indrašienė, V. (2020). Critical thinking embeddedness in higher education programmes. *The Journal of Education, Culture, and Society*, 11(2), 121–132.
- Valero Haro, A., Noroozi, O., Biemans, H., & Mulder, M. (2019). First- and second-order scaffolding of argumentation competence and domain-specific knowledge acquisition: A systematic review. *Technology, Pedagogy and Education*, 28(3), 329–345. <https://doi.org/10.1080/1475939X.2019.1612772>
- van Laar, E., van Deursen, A. J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577–588. <https://doi.org/10.1016/j.chb.2017.03.010>
- van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical approach to modelling cognitive processes*. Academic Press.

- Verplanken, B., Hazenberg, P. T., & Palen wen, G. R. (1992). Need for cognition and external information search effort. *Journal of Research in Personality*, 26(2), 128–136.
[https://doi.org/10.1016/0092-6566\(92\)90049-A](https://doi.org/10.1016/0092-6566(92)90049-A)
- Voltaire, F. M. A. de. (1786). Libert  de penser. In *Dictionnaire philosophique. Oeuvres compl tes de Voltaire—Band 41* (Bd. 5). Jean-Jaques Tourneisen.
- vom Hofe, R. (2010). Residual plot. In N. J. Salkind (Hrsg.), *Encyclopedia of research design* (3. Aufl., S. 1262–1267). SAGE Publications, Inc.
- von der M hlen, S., Richter, T., Schmid, S., Schmidt, E. M., & Berthold, K. (2016). The use of source-related strategies in evaluating multiple psychology texts: A student–scientist comparison. *Reading and Writing*, 29, 1677–1698. <https://doi.org/10.1007/s11145-015-9601-0>
- Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2009). How students evaluate information and sources when searching the world wide web for information. *Computers & Education*, 52(1), 234–246. <https://doi.org/10.1016/j.compedu.2008.08.003>
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273–281. <https://doi.org/10.1080/14640746808400161>
- Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, 53(2), 134–144.
<https://doi.org/10.1002/asi.10016>
- Weber, H., Becker, D., & Hillmert, S. (2019). Information-seeking behaviour and academic success in higher education: Which search strategies matter for grade differences among university students and how does this relevance differ by field of study? *Higher Education*, 77, 657–678.
<https://doi.org/10.1007/s10734-018-0296-4>
- Weber, H., Hillmert, S., & Rott, K. J. (2018). Can digital information literacy among undergraduates be improved? Evidence from an experimental study. *Teaching in Higher Education*, 23(8), 909–926. <https://doi.org/10.1080/13562517.2018.1449740>
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), 1049–1062. <https://doi.org/10.1037/0022-3514.67.6.1049>
- Wells, J. A., LaFon, D. S., & Gratian, M. (2021). Assessing the credibility of cyber adversaries. *International Journal of Cybersecurity Intelligence & Cybercrime*, 4(1), 3–24.

- Werth, L., Denzler, M., & Mayer, J. (2020). *Sozialpsychologie – Das Individuum im sozialen Kontext. Wahrnehmen – Denken – Fühlen* (2. Aufl.). Springer.
- WHO. (2021). *Infodemic*. Infodemic. https://www.who.int/health-topics/infodemic#tab=tab_1
- Wie gesund ist E-Bike fahren? (2022, Oktober 14). E-Bike Gesundheit. <https://www.ebike-gesundheit.de/>
- Wierzbicki, A. (2018). *Web content credibility*. Springer. https://doi.org/10.1007/978-3-319-77794-8_2
- Wikipedia, die freie Enzyklopädie. (o. J.). Abgerufen 18. Dezember 2023, von <https://www.wikipedia.de/>
- Wiley, J., Goldman, S. R., Graesser, A. C., Sanchez, C. A., Ash, I. K., & Hemmerich, J. A. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal December*, 46(4), 1060–1106. <https://doi.org/10.3102/0002831209333183>
- Wineburg, S., Breakstone, J., McGrew, S., & Ortega, T. (2018). Why google can't save us. In O. Zlatkin-Troitschanskaia, G. Wittum, & A. Dengel (Hrsg.), *Positive learning in the age of information: A blessing or a curse?* (S. 221–228). Springer VS. https://doi.org/10.1007/978-3-658-19567-0_13
- Wineburg, S., Breakstone, J., Ziv, N., & Smith, M. (2020). *Educating for misunderstanding: How approaches to teaching digital literacy make students susceptible to scammers, rogues, bad actors, and hate mongers* (Working Paper A-21322). Stanford History Education Group. <https://purl.stanford.edu/mf412bt5333>
- Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, 121(11), 1–40. <https://doi.org/10.1177/016146811912101102>
- Winkel, O. (1999). Die Förderung von Vertrauen, Glaubwürdigkeit und Verlässlichkeit. In P. Rössler & W. Wirth (Hrsg.), *Glaubwürdigkeit im Internet. Fragestellungen, Modelle, empirische Befunde* (Bd. 32, S. 193–208). Reinhard Fischer.
- Winter, S., & Krämer, N. C. (2012). Selecting science information in web 2.0: How source cues, message sidedness, and need for cognition influence users' exposure to blog posts. *Journal of Computer-Mediated Communication*, 18(1), 80–96. <https://doi.org/10.1111/j.1083-6101.2012.01596.x>

- Wirth, W. (1999). Methodologische und konzeptionelle Aspekte der Glaubwürdigkeitsforschung. In P. Rössler & W. Wirth (Hrsg.), *Glaubwürdigkeit im Internet: Fragestellungen, Modelle, empirische Befunde* (S. 47–66). Reinhard Fischer.
- Wirth, W., Böcking, T., Karnowski, V., & Von Pape, T. (2007). Heuristic and systematic use of search engines. *Journal of Computer-Mediated Communication*, 12(3), 778–800.
<https://doi.org/10.1111/j.1083-6101.2007.00350.x>
- Wirth, W., & Brecht, M. (1998). Medial und personal induzierte Selektionsentscheidungen bei der Nutzung des World Wide Web. In P. Rössler (Hrsg.), *Online-Kommunikation: Beiträge zu Nutzung und Wirkung* (S. 147–168). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-87318-7_8
- Wirth, W., & Schweiger, W. (1999). *Selektion neu betrachtet: Empirische Analysen zu einem Schlüsselkonzept* (W. Wirth & W. Schweiger, Hrsg.). Westdeutscher Verlag.
https://doi.org/10.1007/978-3-322-91623-5_2
- Wopereis, I., Brand-Gruwel, S., & Vermetten, Y. (2008). The effect of embedded instruction on solving information problems. *Computers in Human Behavior*, 24(3), 738–752.
<https://doi.org/10.1016/j.chb.2007.01.024>
- Wu, W.-C., Kelly, D., & Sud, A. (2014). Using information scent and need for cognition to understand online search behavior. *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval*, 557–566.
<https://doi.org/10.1145/2600428.2609626>
- Wuyckens, G., Landry, N., & Fastrez, P. (2022). Untangling media literacy, information literacy, and digital literacy: A systematic meta-review of core concepts in media education. *Journal of Media Literacy Education*, 14(1), 168–182. <https://doi.org/10.23860/JMLE-2022-14-1-12>
- Yao, N., & Wang, Q. (2023). Technostress from smartphone use and its impact on university students' sleep quality and academic performance. *The Asia-Pacific Education Researcher*, 32, 317–326. <https://doi.org/10.1007/s40299-022-00654-5>
- Zebregs, S., van den Putte, B., Neijens, P., & de Graaf, A. (2015). The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: A meta-analysis. *Health Communication*, 30(3), 282–289. <https://doi.org/10.1080/10410236.2013.842528>

- Zhang, K. Z. K., Zhao, S. J., Cheung, C. M. K., & Lee, M. K. O. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, 67, 78–89. <https://doi.org/10.1016/j.dss.2014.08.005>
- Zhang, S., Duke, N. K., & Jiménez, L. M. (2011). The WWWDOT approach to improving students' critical evaluation of websites. *The Reading Teacher*, 65(2), 150–158. <https://doi.org/10.1002/TRTR.01016>
- Zhao, X., Lynch, J., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *The Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>
- Zhou, M., & Lam, K. K. L. (2019). Metacognitive scaffolding for online information search in K-12 and higher education settings: A systematic review. *Educational Technology Research and Development*, 67, 1353–1384. <https://doi.org/10.1007/s11423-019-09646-7>
- ZIV Zweirad-Industrie-Verband. *Marktdaten Fahrräder und E-Bikes 2022* (S. 41). (2023). [Pressekonferenz]. Zweirad-Industrie-Verband. https://www.ziv-zweirad.de/wp-content/uploads/2023/09/ZIV_Marktdatenpraesentation_2023_fuer_GJ_2022.pdf
- Zlatkin-Troitschanskaia, O., Brückner, S., Nagel, M.-T., Bültmann, A.-K., Fischer, J., Schmidt, S., & Molerov, D. (2021). Performance assessment and digital training framework for young professionals' generic and domain-specific online reasoning in law, medicine, and teacher practice. *Journal of Supranational Policies of Education (JoSPoE)*, 13, 9–36. <https://doi.org/10.15366/JOSPOE2021.13.001>
- Zlatkin-Troitschanskaia, O., Fischer, J. & Molerov, D. (2023). *Critical online reasoning in higher education (CORE) – overview of the research unit: Objectives, framework, structure, and research projects* (CORE Working Papers No. 1). Johannes Gutenberg Universität. https://core.uni-mainz.de/files/2023/11/CORE_Working_Paper_No.1_final.pdf
- Zlatkin-Troitschanskaia, O., Hartig, J., Goldhammer, F., & Krstev, J. (2021). Students' online information use and learning progress in higher education – A critical literature review. *Studies in Higher Education*, 46(10), 1996–2021. <https://doi.org/10.1080/03075079.2021.1953336>

Anhang

A 1: Deskriptive Statistiken der Webseitenmerkmale	267
Tabelle A1.1: Aktualität.	267
Tabelle A1.2: Verständlichkeit.	267
Tabelle A1.3: Transparenz.	267
Tabelle A1.4: Quellenherkunft.	268
Tabelle A1.5: Verständlichkeit.	268
Tabelle A1.6: Kommerzielles Interesse.	268
Tabelle A1.7: Genre.	269
Tabelle A1.8: Bekanntheit.	269
Tabelle A1.9: Angabe des Autors.	269
A 2: Variablengüte der Personenmerkmale	270
Tabelle A2.1: Reliabilitätsstatistik von NfC.	270
Tabelle A2.2: Item-Skala Statistik von NfC.	270
Tabelle A2.3: Reliabilitätsstatistik von NfCC.	271
Tabelle A2.4: Item-Skala Statistik von NfCC.	271
Tabelle A2.5: Reliabilitätsstatistik von IO.	273
Tabelle A2.6: Item-Skala Statistik von IO.	273
A 3: Ergänzende Tabellen aus der Regressionsberechnungen für die Modelle	274
Tabelle A3.1: Vollständige Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-A NfC und Nachrecherche.	274
Tabelle A3.2: Vollständige Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-B IO und Nachrecherche.	275
Tabelle A3.3: Vollständige Ergebnisse der Berechnung der logistischen Regression von Modell II-C NfCC und Nachrecherche.	276
Tabelle A3.4: Vollständige Ergebnisse der schrittweisen Berechnung von Modell III-A NfC und Weiterverlinkung.	277
Tabelle A3.5: Vollständige Ergebnisse der schrittweisen Berechnung von Modell III-B IO und Weiterverlinkung.	278
Tabelle A3.6: Vollständige Ergebnisse der schrittweisen Modellberechnung III-C NfCC und Weiterverlinkung anklicken.	279
Tabelle A3.7: Vollständige Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-A NfC, IO und systematische Begründung.	280
Tabelle A3.8: Vollständige Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-B NfCC und systematische Begründung.	280

Anhang A 1: Deskriptive Statistiken der Webseitenmerkmale

Tabelle A1.1: *Aktualität*.

Aktualität		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0	16	13
1	89	71
2	34	21
k. A.	8	5
<i>M</i>	1.129	1.076
<i>SD</i>	0.588	0.564

Anmerkung.

Acht Webseiten erzielten für diese Kategorie kein bewertbares Urteil, weil sechs davon nicht zu öffnen waren (E-Book Sonnenschmidt, Tretwerk Artikel 2, Efahrer.chip Artikel 5, Elektrobike Online Artikel 5, Uni Biberach, Emotion Ebikes Artikel 2). Eine weitere Webseite war der Aufruf zur Teilnahme an der Studie (Ebike Gesundheit Artikel Aufruf) sowie das bereits erwähnte Forum, welches unterschiedliche Datums-Einträge enthielt (Forum Öffentlicher Dienst).

Tabelle A1.2: *Verständlichkeit*.

Verständlichkeit		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: Schwer (Index von 0-30)	25	17
1: Schwierig (Index von 30-50)	45	36
2: Anspruchsvoll (Index von 50-60)	43	30
3: Normal (Index von 60-70)	25	20
4: Einfach (Index von 70-80)	1	1
5: Leicht (Index von 80-90)	0	0
6: Sehr leicht (Index von 90-100)	0	0
k. A.	8	6
<i>M</i>	1.507	1.538
<i>SD</i>	1.015	1.014

Anmerkung.

Acht Webseiten konnten nicht beurteilt werden, da sechs davon nicht zu öffnen waren, eines das bereits bekannte Forum war und zuletzt sich weiterhin wieder der Aufruf zu Teilnahme der Studie darunter befand.

Tabelle A1.3: *Transparenz*.

Transparenz		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: Intransparent	44	31
1: Transparent	95	72
k. A.	8	7
<i>M</i>	0.683	0.700
<i>SD</i>	0.465	0.459

Anmerkung.

Acht Webseiten erzielten für diese Kategorie kein bewertbares Urteil, weil 6 davon nicht zu öffnen waren (E-Book Sonnenschmidt, Tretwerk Artikel 2, Efahrer.chip Artikel 5, Elektrobike Online Artikel 5, Uni Biberach, Emotion Ebikes Artikel 2). Wiederum eine war der Aufruf zu Teilnahme an der Studie (Ebike Gesundheit Artikel Aufruf) sowie das bereits erwähnte Forum (Forum Öffentlicher Dienst).

Tabelle A1.4: *Quellenherkunft*.

Quellenherkunft		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: Sekundärquelle	87	66
1: Primärquelle	53	39
k. A.	7	5
<i>M</i>	0.379	0.371
<i>SD</i>	0.485	0.483

Anmerkung.

Sieben Webseiten konnten nicht beurteilt werden, da sechs nicht zu öffnen waren und eines der Aufruf zur Teilnahme an der E-Bike Studie war.

Tabelle A1.5: *Nachweisqualität*

Nachweisqualität		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: keine Angabe	15	13
1: allgemeine, nicht inhaltsrelevante Links	8	7
2: Expertenmeinung im Text	20	12
3: Quellen und/oder Links im Text und/oder unterhalb des Textes	96	73
k. A.	8	5
<i>M</i>	2.428	2.380
<i>SD</i>	1.005	1.055

Anmerkung.

Acht Webseiten erzielten für diese Kategorie kein bewertbares Urteil, weil sechs davon nicht zu öffnen waren (E-Book Sonnenschmidt, Tretwerk Artikel 2, Efahrr.chip Artikel 5, Elektrobike Online Artikel 5, Uni Biberach, Emotion Ebikes Artikel 2). Wiederum eines war der Aufruf zu Teilnahme an der Studie (Ebike Gesundheit Artikel Aufruf) sowie das bereits erwähnte Forum (Forum Öffentlicher Dienst).

Tabelle A1.6: *Kommerzielles Interesse*.

Kommerzielles Interesse		
	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
<i>n</i>	147	110
0	33	24
1	72	52
2	40	33
k. A.	4	1
<i>M</i>	1.063	1.083
<i>SD</i>	0.704	0.718

Anmerkung.

Vier aufgerufene Webseiten erzielten für diese Kategorie kein bewertbares Urteil, weil zwei davon nicht zu öffnen waren (E-Book Sonnenschmidt und Uni Biberach). Eine Webseite weiterhin war der Aufruf zur Teilnahme an der Studie und eine letzte Webseite war ein Forum, weshalb diese jeweils keine Bewertung erhielten.

Tabelle A1.7: *Genre*.

Genre		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: Online-Shop (<i>n</i> = 24), World-Wide-Web (<i>n</i> = 10)	34	25
1: Enzyklopädie (<i>n</i> = 0)	0	0
2: Nachrichtenseite (<i>n</i> = 29), Verband (<i>n</i> = 11)	40	29
3: Fachmagazin/Fachverlag (<i>n</i> = 21), Google Books (<i>n</i> = 1)	22	18
4: (Fach-)Datenbanken (<i>n</i> = 16), Lehrmaterialien (<i>n</i> = 4), Regierungsinstanz (<i>n</i> = 4), Universitärer Online-Katalog (<i>n</i> = 2), Wissenschaftliche Fachzeitschrift/Fachbuch (<i>n</i> = 22), Wissenschaftliches Forschungsinstitut (<i>n</i> = 3)	51	38
<i>M</i>	2.385	2.400
<i>SD</i>	1.518	1.515

Tabelle A1.8: *Bekanntheit*.

Bekanntheit		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: Unbekannt	101	76
1: Bekannt	36	29
k. A.	9	5
<i>M</i>	0.268	0.276
<i>SD</i>	0.443	0.449

Anmerkung.

Neun Webseiten konnten nicht beurteilt werden, da sie bei der ersten Beurteilung nicht mehr offenbar waren (z. B. Uni Biberach, E-Book Sonnenschmidt) und beim Versuch einer nachträglichen Beurteilung war es seitens des Tools (similarweb.com) nicht mehr möglich, den exakt gleichen Zeitraum, wie bei der ersten Bewertungsrunde einzustellen.

Tabelle A1.9: *Angabe des Autors*

Angabe des Autors		
Ausprägung	Aufgerufene Webseiten; <i>n</i> = 147	Selektierte Webseiten; <i>n</i> = 110
0: keine Angabe des Autors/der Autorin	43	34
1: Angabe existiert	97	71
k. A.	7	5
<i>M</i>	0.691	0.676
<i>SD</i>	0.462	0.468

Anmerkung.

Sieben Webseiten konnten nicht beurteilt werden, da sechs nicht zu öffnen waren und eines erneut der Aufruf zur Teilnahme an der E-Bike Studie war.

Anhang A 2: Variablenlänge der Personenmerkmale

Tabelle A2.1: Reliabilitätsstatistik von NfC.

Cronbach's α	Cronbach's α für standardisierte Items	Anzahl der Items
.672	.658	4

Tabelle A2.2: Item-Skala Statistik von NfC.

Items	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbach's α , wenn Items weggelassen	Min / Max	M (SD)
1: Es genügt mir, die Antwort zu kennen, ohne die Gründe für die Antwort eines Problems zu verstehen (<i>inv</i>)	11.625	8.474	0.305	.689	2 / 6	4.993 (0.964)
2: Ich habe es gern, wenn mein Leben voller kniffliger Aufgaben ist, die ich lösen muss.	13.319	5.492	0.665	.441	1 / 6	3.299 (1.280)
3: Ich würde kompliziertere Probleme einfacheren vorziehen.	13.451	5.592	0.564	.524	1 / 6	3.167 (1.369)
4: In erster Linie denke ich, weil ich muss (<i>inv</i>).	11.458	8.194	0.313	.687	1 / 6	5.160 (1.036)

Tabelle A2.3: *Reliabilitätsstatistik von NfCC.*

Cronbach's α	Cronbach's α für standardisierte Items	Anzahl der Items
.752	.764	16

Tabelle A2.4: *Item-Skala Statistik von NfCC.*

Items	Skalen- mittel- wert, wenn I- tem weg- gelassen	Skalen- varianz, wenn I- tem weg- gelassen	Korri- gierte I- tem- Skala- Korrela- tion	Cron- bach's α , wenn Items weggelas- sen	Min / Max	M (SD)
1: Ich mag es nicht, wenn die Aussage einer Person mehrdeutig ist.	35.594	24.182	0.153	.759	1 / 4	2.906 (0.791)
2: Ich finde, nachdem ich eine Lösung für ein Problem gefunden habe, ist es Zeitverschwendung, weitere mögliche Lösungen in Betracht zu ziehen	36.531	22.475	0.488	.727	1 / 4	1.969 (0.666)
3: Ich mag keine unvorhersehbaren Situationen.	35.781	24.015	0.274	.745	1 / 4	2.719 (0.603)
4: Ich finde es spannend nicht zu wissen, was das Leben einem bringen wird (inv).	36.125	22.968	0.320	.742	1 / 4	2.375 (0.787)
5: Ein Problem aus verschiedenen Blickwinkeln zu betrachten, führt nur zu Verwirrung.	37.203	23.593	0.491	.733	1 / 2	1.297 (0.461)
6: Im Allgemeinen suche ich nicht nach Alternativlösungen für Probleme, für welche ich schon eine Lösung parat habe.	36.125	24.270	0.169	.756	1 / 4	2.375 (0.724)
7: Ich bevorzuge die Gesellschaft guter Freunde, weil ich weiß, was ich von ihnen zu erwarten habe.	35.391	24.337	0.209	.751	2 / 4	3.109 (0.620)
8: Ich fühle mich unbehaglich, wenn ich es nicht schaffe eine schnelle Antwort auf Probleme zu geben, denen ich gegenüber stehe	35.844	23.626	0.280	.745	1 / 4	2.656 (0.695)
9: Ich bevorzuge Tätigkeiten, bei denen stets klar ist, was getan und wie es getan werden muss.	35.609	22.877	0.399	.734	2 / 4	2.891 (0.693)
10: Wenn ich ein Problem lösen muss, verschwende ich im Allgemeinen keine Zeit damit, die unterschiedlichen Standpunkte dazu zu erwägen	36.688	22.853	0.448	.731	1 / 3	1.813 (0.640)
11: Ich mag Aufgaben, bei denen noch unklar ist, wie der genaue Lösungsweg aussieht (inv).	36.109	22.035	0.547	.721	1 / 4	2.391 (0.681)

12: Ich liebe die Ungewissheit und die Überraschung, die oft im Alltäglichen steckt (inv).	35.953	21.506	0.523	.721	1 / 4	2.547 (0.795)
13: Jedwede Lösung eines Problems ist besser, als in einem Zustand der Ungewissheit zu verharren.	35.703	23.704	0.222	.752	1 / 4	2.797 (0.780)
14: Ich ziehe Dinge, die ich gewohnt bin, solchen vor, die ich nicht kenne und die ich nicht vorhersagen kann.	35.781	23.253	0.368	.738	1 / 4	2.719 (0.654)
15: Im Allgemeinen vermeide ich es, mich an Diskussionen über uneindeutige und umstrittene Themen zu beteiligen.	36.375	23.444	0.254	.749	1 / 4	2.125 (0.787)
16: Ich bevorzuge es, mich für die erst-mögliche Lösung zu entscheiden, anstelle lange darüber nachzudenken, was für eine Entscheidung ich treffen sollte	36.688	22.536	0.530	.724	1 / 4	1.813 (0.614)

Tabelle A2.5: *Reliabilitätsstatistik von IO.*

Cronbach's α	Cronbach's α für standardisierte Items	Anzahl der Items
0.855	.858	3

Tabelle A2.6: *Item-Skala Statistik von IO.*

Items	Skalenmittel- wert, wenn I- tem wegge- lassen	Skalen- varianz, wenn I- tem wegge- lassen	Korrigierte Item-Skala- Korrelation	Cronbach's α, wenn Items weg- gelassen	Min / Max	M (SD)
1: Ich bin oftmals abgelenkt von der übermäßigen Fülle an Informationen in den Medien.	6.813	7.118	0.742	.786	1 / 6	3.153 (1.334)
2: Ich fühle mich überfordert von der Menge an Informationen aus den Medien, mit denen ich täglich konfrontiert werde.	6.868	6.465	0.798	.728	1 / 6	3.097 (1.416)
3: Meistens ist mein Problem, dass ich zu viele Informationen verarbeiten muss und nicht, dass ich zu wenig Informationen habe, um mir eine Meinung zu bilden	6.250	6.748	0.651	.874	1 / 6	3.715 (1.522)

Anhang A 3: *Ergänzende Tabellen aus der Regressionsberechnungen für die Modelle.*Tabelle A3.1: *Vollständige Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-A NfC und Nachrecherche*

Logistische Regression mit AV Nachrecherche und UV NFC							
Mo- dellnr.	Variable	Regressi- onskoeffi- zientB	Stan- dard- fehler	Likelihood-Quotienten- Test		Signif- ikanz	Exp(B)
				Chi²	df		
1 ^a	Alter	-.045	0.064	0.503	1	.478	0.956
	HZB-Note	-.337	0.373	0.83	1	.362	0.714
	Weiblich	-.268	0.442	0.365	1	.546	0.765
	Divers	-1.264	0.919	2.118	1	.146	0.283
	Lehramt	1.825	0.588	10.567	1	.001	6.205
	Jura	.843	0.591	2.01	1	.156	2.322
	Konstante	.491	1.749	0.079	1	.778	1.633
2 ^b	Alter	-.055	.066	.718	1	.397	.946
	HZB-Note	-.281	.390	.525	1	.469	.755
	Weiblich	-.071	.487	.021	1	.884	.931
	Divers	-.978	.964	1.109	1	.292	.376
	Lehramt	1.887	.638	9.682	1	.002	6.599
	Jura	.675	.620	1.172	1	.279	1.965
	NfC	.322	.217	2.277	1	.131	1.379
	Textglaubwürdigkeit	-.508	.357	2.079	1	.149	.602
	Textglaubwürdigkeit ²	.028	.223	.016	1	.900	1.028
	Medienglaubwürdigkeit	.227	.302	.569	1	.451	1.254
	Medienglaubwürdigkeit ²	-.108	.207	.280	1	.597	.898
	Autorenglaubwürdigkeit	.412	.312	1.771	1	.183	1.510
	Autorenglaubwürdigkeit ²	.096	.211	.203	1	.652	1.101
	Konstante	.496	1.825	0.074	1	.785	1.642
3 ^c	Alter	-.053	.069	.584	1	.445	.949
	HZB-Note	-.434	.446	.960	1	.327	.648
	Weiblich	-.158	.529	.089	1	.765	.854
	Divers	-1.062	1.058	1.103	1	.294	.346
	Lehramt	2.041	.714	9.198	1	.002	7.699
	Jura	.674	.666	1.016	1	.313	1.962
	NfC	.585	.384	2.488	1	.115	1.795
	Textglaubwürdigkeit	-.715	.388	3.635	1	.057	.489
	Textglaubwürdigkeit ²	.538	.329	2.774	1	.096	1.712
	Medienglaubwürdigkeit	-.214	.351	.375	1	.540	.807
	Medienglaubwürdigkeit	.073	.247	.087	1	.768	1.076
	Autorenglaubwürdigkeit	.798	.371	5.052	1	.025	2.220
	Autorenglaubwürdigkeit ²	-.046	.246	.036	1	.850	.955
	NfC*Textglaubwürdigkeit	.240	.423	.325	1	.569	1.272
	NfC*	-.586	.301	4.096	1	.043	.557
	Textglaubwürdigkeit ²						
	NfC*Medieng- laubwürdigkeit	-.008	.249	4.229	1	.040	.992
	NfC* Medieng- laubwürdigkeit ²	.727	.363	0.001	1	.974	2.070
	NfC* Auto- renglaubwürdigkeit	-.625	.418	2.371	1	.124	.535
	NfC* Auto- renglaubwürdigkeit ²	.492	.362	1.935	1	.164	1.636
	Konstante	.175	1.969	0.008	1	.929	1.191

Tabelle A3.2 Vollständige Ergebnisse der schrittweisen Berechnung der logistischen Regression von Modell II-B IO und Nachrecherche.

Logistische Regression mit AV Nachrecherche und UV IO							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
1 ^a	Alter	-.045	.064	.503	1	.478	.956
	HZB-Note	-.337	.373	.830	1	.362	.714
	Weiblich	-.268	.442	.365	1	.546	.765
	Divers	-1.264	.919	2.118	1	.146	.283
	Lehramt	1.825	.588	10.567	1	.001	6.205
	Jura	.843	.591	2.010	1	.156	2.322
	Konstante	.491	1.749	0.079	1	.778	1.633
2 ^b	Alter	-.053	.068	.623	1	.430	.949
	HZB-Note	-.343	.389	.789	1	.374	.709
	Weiblich	-.213	.481	.195	1	.659	.808
	Divers	-1.278	.969	1.921	1	.166	.278
	Lehramt	1.938	.646	10.005	1	.002	6.943
	Jura	.862	.618	1.917	1	.166	2.368
	IO	-.305	.207	2.220	1	.136	.737
	Textglaubwürdigkeit	-.452	.355	1.652	1	.199	.636
	Textglaubwürdigkeit ²	.106	.217	.237	1	.626	1.112
	Medienglaubwürdigkeit	.288	.302	.921	1	.337	1.334
	Medienglaubwürdigkeit ²	-.069	.203	.116	1	.734	0.934
	Autorenglaubwürdigkeit	.267	.311	.738	1	.390	1.306
	Autorenglaubwürdigkeit ²	.056	.209	.073	1	.788	1.058
	Konstante	.521	1.858	0.080	1	.788	1.684
3 ^c	Alter	-.082	.076	1.223	1	.269	.922
	HZB-Note	-.260	.447	.340	1	.560	.771
	Weiblich	-.723	.531	1.863	1	.172	.485
	Divers	-2.017	1.058	4.138	1	.042	.133
	Lehramt	2.208	.754	9.710	1	.002	9.101
	Jura	.792	.666	1.405	1	.236	2.208
	IO	-.550	.405	1.925	1	.165	.577
	Textglaubwürdigkeit	-.594	.417	2.099	1	.147	.552
	Textglaubwürdigkeit ²	.289	.313	.868	1	.352	1.335
	Medienglaubwürdigkeit	.563	.388	2.302	1	.129	1.756
	Medienglaubwürdigkeit ²	-.321	.297	1.264	1	.261	.725
	Autorenglaubwürdigkeit	.239	.344	.485	1	.486	1.270
	Autorenglaubwürdigkeit ²	-.030	.236	.016	1	.898	.970
	IO*Textglaubwürdigkeit	.845	.479	3.386	1	.066	2.329
	IO*Textglaubwürdigkeit ²	.977	.409	8.160	1	.004	2.656
	IO*Medienglaubwürdigkeit	-.199	.430	.215	1	.643	.819
	IO*Medienglaubwürdigkeit ²	-.646	.325	4.938	1	.026	.524
	IO*Autorenglaubwürdigkeit	-.541	.389	2.007	1	.157	.582
	IO*Autorenglaubwürdigkeit ²	.033	.305	.012	1	.914	1.033
	Konstante	1.508	2.125	0.523	1	.469	4.517

Tabelle A3.3: *Vollständige Ergebnisse der Berechnung der logistischen Regression von Modell II-C NfCC und Nachrecherche.*

Logistische Regression mit AV Nachrecherche und UV NfCC							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
1 ^a	Alter	.084	.117	.517	1	.472	1.088
	HZB-Note	-.632	.649	.999	1	.318	.531
	Weiblich	.264	.667	.160	1	.689	1.303
	Lehramt	1.373	.890	2.486	1	.115	3.948
	Jura	.826	.964	.729	1	.393	2.285
	Konstante	-2.846	3.009	.000	1	.	.058
2 ^b	Alter	.116	.125	.886	1	.346	1.123
	HZB-Note	-.924	.745	1.691	1	.193	.397
	Weiblich	.289	.806	.130	1	.718	1.335
	Lehramt	1.462	1.102	1.886	1	.170	4.314
	Jura	1.004	1.139	.782	1	.376	2.729
	NfCC	-.491	.394	1.690	1	.194	.612
	Textglaubwürdigkeit	.733	.735	1.024	1	.312	2.081
	Textglaubwürdigkeit ²	.248	.387	.412	1	.521	1.282
	Medienglaubwürdigkeit	-.658	.536	1.576	1	.209	.518
	Medienglaubwürdigkeit ²	-.074	.429	.030	1	.862	.928
	Autorenglaubwürdigkeit	-.267	.561	.229	1	.632	.765
	Autorenglaubwürdigkeit ²	.046	.410	.012	1	.911	1.047
	Konstante	-3.551	3.344	.000	0	.	.029
3 ^c	Alter	.104	.137	.575	1	.448	1.109
	HZB-Note	-.443	.962	.214	1	.643	.642
	Weiblich	.942	1.090	.789	1	.374	2.564
	Lehramt	1.245	1.516	.673	1	.412	3.473
	Jura	.186	1.346	.019	1	.890	1.204
	NfCC	-.262	.853	.095	1	.757	.769
	Textglaubwürdigkeit	-.159	1.102	.021	1	.885	.853
	Textglaubwürdigkeit ²	1.642	.839	4.923	1	.026	5.164
	Medienglaubwürdigkeit	-.141	.866	.026	1	.872	.868
	Medienglaubwürdigkeit ²	-.669	.542	1.654	1	.198	.512
	Autorenglaubwürdigkeit	-.117	.809	.021	1	.885	.890
	Autorenglaubwürdigkeit ²	-.665	.850	.709	1	.400	.514
	NfCC*Textglaubwürdigkeit	-.386	1.249	.097	1	.755	.680
	NfCC*Textglaubwürdigkeit ²	1.915	1.341	5.063	1	.024	6.786
	NfCC*Medienglaubwürdigkeit	.584	1.033	.329	1	.566	1.793
	NfCC*Medienglaubwürdigkeit ²	-.513	.933	.301	1	.583	.598
	NfCC*Autorenglaubwürdigkeit	-1.093	1.291	.741	1	.389	.335
	NfCC*Autorenglaubwürdigkeit ²	-2.381	1.473	6.369	1	.012	.092
	Konstante	-4.446	3.818	1.396	1	.237	.012

Tabelle A3.4: Vollständige Ergebnisse der schrittweisen Berechnung von Modell III-A NfC und Weiterverlinkung.

Logistische Regression mit AV Weiterverlinkung und UV NfC							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi²	df		
1 ^a	Alter	.091	.068	1.730	1	.188	1.095
	HZB-Note	-1.066	.508	5.009	1	.025	.344
	Weiblich	-.144	.468	.095	1	.758	.866
	Divers	-19.848	12905.5	3.646	1	.056	.000
			93				
	Lehramt	-.977	.668	2.264	1	.132	.376
	Jura	-.037	.633	.003	1	.953	.964
	Konstante	-1.308	1.788	0.528	1	.468	.270
2 ^b	Alter	.055	.073	.568	1	.451	1.057
	HZB-Note	-1.278	.609	5.070	1	.024	.278
	Weiblich	.309	.602	.267	1	.605	1.362
	Divers	-19.428	12508.4	2.475	1	.116	.000
			30				
	Lehramt	.252	.813	.096	1	.756	1.287
	Jura	.224	.702	.101	1	.750	1.251
	NfC	.020	.246	.006	1	.936	1.020
	Textglaubwürdigkeit	1.155	.575	4.961	1	.026	3.175
	Textglaubwürdigkeit ²	-.727	.486	2.557	1	.110	.484
	Medienglaubwürdigkeit	.972	.414	6.839	1	.009	2.643
	Medienglaubwürdigkeit ²	-.055	.371	.022	1	.881	.947
	Autorenglaubwürdigkeit	-.636	.383	2.864	1	.091	.530
	Autorenglaubwürdigkeit ²	-.079	.314	.064	1	.800	.924
	Konstante	-.538	2.014	0.071	1	.790	.584
3 ^c	Alter	.065	.076	.727	1	.394	1.067
	HZB-Note	-.909	.669	1.983	1	.159	.403
	Weiblich	.423	.623	.467	1	.494	1.526
	Divers	-19.272	12140.7	2.122	1	.145	.000
			54				
	Lehramt	-.057	.884	.004	1	.948	.944
	Jura	.019	.727	.001	1	.978	1.019
	NfC	.256	.474	.301	1	.583	1.292
	Textglaubwürdigkeit	1.134	.641	3.865	1	.049	3.107
	Textglaubwürdigkeit ²	-.807	.549	2.593	1	.107	.446
	Medienglaubwürdigkeit	1.176	.454	8.338	1	.004	3.242
	Medienglaubwürdigkeit ²	-.059	.394	.023	1	.880	.943
	Autorenglaubwürdigkeit	-.678	.418	2.767	1	.096	.508
	Autorenglaubwürdigkeit ²	-.049	.335	.022	1	.883	.952
	NfC*Textglaubwürdigkeit	-.087	.685	.016	1	.898	.917
	NfC*Textglaubwürdigkeit ²	.051	.582	.008	1	.930	1.052
	NfC*Medienglaubwürdigkeit	-.331	.462	.501	1	.479	.718
	NfC*Medienglaubwürdigkeit ²	-.408	.328	1.493	1	.222	.665
	NfC*Autorenglaubwürdigkeit	.500	.490	1.107	1	.293	1.649
	NfC*Autorenglaubwürdigkeit ²	.168	.431	.151	1	.697	1.182
	Konstante	-1.316	2.203	.357	1	0.550	.268

Tabelle A3.5: Vollständige Ergebnisse der schrittweisen Berechnung von Modell III-B IO und Weiterverlinkung.

Logistische Regression mit AV Weiterverlinkung und UV IO							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test			
				Chi²	df	Sig.	Exp(B)
1 ^a	Alter	.091	.068	1.730	1	.188	1.095
	HZB-Note	-1.066	.508	5.009	1	.025	.344
	Weiblich	-.144	.468	.095	1	.758	.866
	Divers	-19.848	12905.593	3.646	1	.056	.000
	Lehramt	-.977	.668	2.264	1	.132	.376
	Jura	-.037	.633	.003	1	.953	.964
	Konstante	-1.308	1.788	0.528	1	.468	.270
2 ^b	Alter	.039	.075	.262	1	.609	1.039
	HZB-Note	-1.312	.615	5.293	1	.021	.269
	Weiblich	.304	.600	.259	1	.611	1.355
	Divers	-19.696	12311.299	2.946	1	.086	.000
	Lehramt	.533	.845	.400	1	.527	1.704
	Jura	.322	.708	.206	1	.650	1.379
	IO	-.359	.262	1.942	1	.163	.699
	Textglaubwürdigkeit	1.211	.568	5.568	1	.018	3.355
	Textglaubwürdigkeit ²	-.725	.485	2.566	1	.109	.484
	Medienglaubwürdigkeit	1.125	.441	8.367	1	.004	3.080
	Medienglaubwürdigkeit ²	-.122	.389	.105	1	.746	.885
	Autorenglaubwürdigkeit	-.767	.393	3.988	1	.046	.464
	Autorenglaubwürdigkeit ²	-.127	.320	.163	1	.687	.880
	Konstante	-.078	2.076	0.001	1	.970	.925
3 ^c	Alter	.071	.099	.497	1	.481	1.073
	HZB-Note	-2.048	.830	7.827	1	.005	.129
	Weiblich	-.180	.735	.060	1	.806	.835
	Divers	-20.543	11679.472	4.237	1	.040	.000
	Lehramt	1.748	1.112	2.649	1	.104	5.742
	Jura	-.157	.970	.026	1	.871	.854
	IO	-1.564	.727	6.003	1	.014	.209
	Textglaubwürdigkeit	2.110	.770	9.497	1	.002	8.250
	Textglaubwürdigkeit ²	-.838	.575	2.291	1	.130	.433
	Medienglaubwürdigkeit	1.689	.610	11.437	1	.001	5.413
	Medienglaubwürdigkeit ²	-.266	.557	.248	1	.619	.766
	Autorenglaubwürdigkeit	-1.765	.610	10.469	1	.001	.171
	Autorenglaubwürdigkeit ²	-.314	.411	.605	1	.437	.731
	IO*Textglaubwürdigkeit	2.645	1.050	8.572	1	.003	14.089
	IO*Textglaubwürdigkeit ²	2.267	.823	6.866	1	.009	9.647
	IO*Medienglaubwürdigkeit	-1.514	.696	5.261	1	.022	.220
	IO*Medienglaubwürdigkeit ²	.362	.600	.370	1	.543	1.436
	IO*Autorenglaubwürdigkeit	-.868	.814	1.200	1	.273	.420
	IO*Autorenglaubwürdigkeit ²	-1.771	.687	8.902	1	.003	.170
	Konstante	.165	2.706	0.004	1	.951	1.180

Tabelle A3.6: Vollständige Ergebnisse der schrittweisen Modellberechnung III-C NfCC und Weiterverlinkung anklicken.

Logistische Regression mit AV Weiterverlinkung und UV NfCC							
Mo- dellnr.	Variable	Regressi- onskoef- fizientB	Stan- dard- fehler	Likelihood-Quotien- ten-Test		Sig.	Exp(B)
				Chi ²	df		
1 ^a	Alter	.145	.122	1.430	1	.232	1.156
	HZB-Note	-.463	.793	.355	1	.551	.629
	Weiblich	-.726	.691	1.090	1	.297	.484
	Lehramt	-1.059	1.004	1.185	1	.276	.347
	Jura	-.740	1.116	.458	1	.499	.477
	Konstante	-3.432	2.976	1.357	1	.244	.032
2 ^b	Alter	.182	.184	1.100	1	.294	1.200
	HZB-Note	-.826	1.023	.689	1	.406	.438
	Weiblich	-2.010	1.107	3.691	1	.055	.134
	Lehramt	-.106	1.383	.006	1	.939	.899
	Jura	-.177	1.435	.015	1	.901	.837
	NfCC	.601	.477	1.728	1	.189	1.824
	Textglaubwürdigkeit	2.091	1.180	3.933	1	.047	8.089
	Textglaubwürdigkeit ²	-.788	.831	.953	1	.329	.455
	Medienglaubwürdigkeit	.522	.718	.556	1	.456	1.686
	Medienglaubwürdigkeit ²	-.959	.832	1.462	1	.227	.383
	Autorenglaubwürdigkeit	-1.756	.756	6.991	1	.008	.173
	Autorenglaubwürdigkeit ²	-.309	.480	.438	1	.508	.734
	Konstante	-2.522	4.369	0.341	1	.559	.080
3 ^c	Alter	.301	.255	1.660	1	.198	1.352
	HZB-Note	-1.192	1.384	.873	1	.350	.304
	Weiblich	-2.362	1.495	2.924	1	.087	.094
	Lehramt	1.183	1.871	.410	1	.522	3.264
	Jura	-1.341	1.958	.511	1	.475	.261
	NfCC	-.519	1.044	.247	1	.619	.595
	Textglaubwürdigkeit	3.020	1.690	4.263	1	.039	20.496
	Textglaubwürdigkeit ²	-1.518	1.210	1.876	1	.171	.219
	Medienglaubwürdigkeit	1.074	1.206	.940	1	.332	2.926
	Medienglaubwürdigkeit ²	-1.444	1.081	2.156	1	.142	.236
	Autorenglaubwürdigkeit	-2.237	1.085	6.070	1	.014	.107
	Autorenglaubwürdigkeit ²	.243	.703	.121	1	.728	1.275
	NfCC*Textglaubwürdigkeit	1.854	1.704	1.198	1	.274	6.383
	NfCC*Textglaubwürdigkeit ²	.273	1.818	.023	1	.881	1.313
	NfCC*Medienglaubwürdigkeit	-.784	1.732	.234	1	.629	.457
	NfCC*Medienglaubwürdigkeit ²	2.376	1.621	3.304	1	.069	10.767
	NfCC*Autorenglaubwürdigkeit	-1.082	1.065	1.056	1	.304	.339
	NfCC*Auto- renglaubwürdigkeit ²	-.707	.939	.638	1	.424	.493
	Konstante	-5.298	6.091	.000	0	.	.005

Tabelle A3.7: *Vollständige Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-A NfC, IO und systematische Begründung.*

Logistische Regression mit AV Systematische Begründung und UVs IO und NfC							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
1 ^a	Alter	.109	.098	1.321	1	.250	1.115
	HZB-Note	-.292	.421	.485	1	.486	.747
	Weiblich	.444	.455	.943	1	.331	1.559
	Divers	-.567	.847	.453	1	.501	.567
	Lehramt	-1.259	.573	5.076	1	.024	.284
	Jura	.725	.675	1.220	1	.269	2.065
	Konstante	-1.575	2.450	0.424	1	.515	.207
2 ^b	Alter	.105	.097	1.217	1	.270	1.111
	HZB-Note	-.295	.424	.487	1	.485	.744
	Weiblich	.465	.465	.988	1	.320	1.592
	Divers	-.505	.864	.343	1	.558	.604
	Lehramt	-1.274	.576	5.157	1	.023	.280
	Jura	.671	.682	1.017	1	.313	1.956
	NfC	.029	.205	.020	1	.887	1.030
	IO	.205	.212	0.946	1	.331	1.227
	Konstante	-1.446	2.449	.357	1	.550	.235

Tabelle A3.8: *Vollständige Ergebnisse der schrittweisen Berechnung der multiplen logistischen Regression von Modell V-B NfCC und systematische Begründung.*

Logistische Regression mit AV Systematische Begründung und UV NfCC							
Modellnr.	Variable	RegressionskoeffizientB	Standardfehler	Likelihood-Quotienten-Test		Signifikanz	Exp(B)
				Chi ²	df		
1 ^a	Alter	-.094	.152	.379	1	.538	.910
	HZB-Note	.754	.714	1.178	1	.278	2.126
	Weiblich	.701	.635	1.210	1	.271	2.016
	Lehramt	-1.629	.866	3.793	1	.051	.196
	Jura	.184	1.059	.030	1	.861	1.202
	Konstante	2.025	3.755	0.289	1	.591	7.578
2 ^b	Alter	-.093	.153	.368	1	.544	.911
	HZB-Note	.775	.724	1.216	1	.270	2.170
	Weiblich	.723	.645	1.251	1	.263	2.062
	Lehramt	-1.659	.881	3.828	1	.050	.190
	Jura	.146	1.076	.019	1	.892	1.157
	NfCC	-.065	.316	.042	1	.838	.937
	Konstante	1.961	3.772	0.269	1	.604	7.105

Lebenslauf

- 06/2024
Promotion (magna cum laude) zum Doktor der wirtschaftlichen Staatswissenschaften (Dr. rer. pol.) im Fach Wirtschaftspädagogik
Johannes Gutenberg-Universität Mainz
- Seit 09/2023
Wissenschaftliche Mitarbeiterin
Johannes Gutenberg-Universität Mainz
Institut für Publizistik
Prof. Dr. Christian Schemer
- 04/2021 – 09/2023
Wissenschaftliche Mitarbeiterin
Johannes Gutenberg-Universität Mainz
Lehrstuhl für Wirtschaftspädagogik
Prof. Dr. Olga Zlatkin-Troitschanskaia
- 10/2019 – 07/2021
Studium im Fach Management
Johannes Gutenberg-Universität Mainz
Abschluss: Master of Science (Note: 1.2)
- 04/2016 – 07/2019
Studium im Fach Wirtschaftspädagogik und Englisch
Universität des Saarlandes
Abschluss: Bachelor of Science (Note: 1.8)