

Hochschule Fresenius Fachbereich Wirtschaft & Medien Studiengang: Psychologie (M. Sc.)

Studienort: Köln

Erste Überprüfung und Validierung eines Fragebogens zur Erfassung von zugrundeliegenden Motiven der Selbstdarstellung auf Sozialen Netzwerken

Genehmigte MASTERTHESIS

zur Erlangung des akademischen Grades

des Master of Science (M. Sc.)

Luca Mario Derjung (B. Sc.)

Matrikelnummer: 400264504

Prüferin: Frau Dr. Dipl.-Psych. Verena Stumm
 Prüferin: Frau Prof. Dr. Petra Lindauer

Abgabedatum: 21.11.2022

Eigentlich für mich, aber am Ende dann doch nur für

Euch!

Danke für alles! Ich liebe euch.

Danksagung

An dieser Stelle ist es mir ein besonderes Bedürfnis, mich bei meiner Betreuerin Frau Dr. Dipl.-Psych. Verena Stumm zu bedanken, die mich über den gesamten Verlauf meines Studiums der Psychologie begleitet hat.

Neben der fachlichen Unterstützung habe ich insbesondere die direkte und sehr offene Kommunikation, die sich im Zuge der langen Zusammenarbeit etabliert hat, stets sehr geschätzt.

> Köln, im November 2022 Luca Mario Derjung

Zusammenfassung

In der vorliegenden Masterthesis wird ein Fragebogen zur Ermittlung von Motiven der Selbstdarstellung auf sozialen Netzwerken erstmalig untersucht und überprüft. Die Prüfung des Fragebogens erfolgt in zwei Schritten: Zunächst soll eine erste Überprüfung der Itemkennwerte und relevanten Gütekriterien erfolgen. Anschließend sollen erste Befunde zur Bestimmung der Kriteriumsvalidität des Fragebogens zusammengetragen werden. Entsprechend besteht die Arbeit aus zwei separaten Studien. Beide Studien basieren auf einem nicht-experimentellen quantitativen Untersuchungsansatz und zur Erhebung der Daten wird jeweils eine Querschnittbefragung durchgeführt. Im Rahmen der ersten Studie (N = 308) erfolgt die erste Erhebung und Itemanalyse sowie die Prüfungen der Reliabilität des Fragebogens. Anschließend wurde der Fragebogen gekürzt und umstrukturiert. Die finale Version besteht aus sieben faktoranalytisch ermittelten Skalen. Im Rahmen der zweiten Studie (N = 112) werden die Item- und Skalenwerte sowie Reliabilitätskoeffizienten erneut untersucht, um die Ergebnisse der ersten Studie bestenfalls zu replizieren. Im zweiten Teil erfolgte die Validierung des Fragebogens anhand verschiedener Außenkriterien. Dazu werden die folgenden Validierungsskalen in den Fragebogen integriert: die Skalen Neurotizismus, Extraversion und Leistungsmotiv des B5-T (Satow, 2020), die Skala Angst vor negativer Bewertung (SANB-5) (Kemper, Lutz, & Neuser, 2011), die Rosenberg Self-Esteem-Scale (RSES) (Collani & Herzberg, 2003), die Skala Anschluss & Kontaktbereitschaft des TEMEO-R (Deutscher & Mees, 2012) und die Skala Orientierung an Anderen des SMSC (Collani & Stürmer, 2009). Diese weisen entsprechende inhaltliche und theoretische Bezüge zu den Skalen des Fragebogens auf. Anschließend werden die Korrelationen zwischen den Skalen und den Außenkriterien untersucht. Zusammenfassend lässt sich im Hinblick auf die Validierung des Fragebogens berichten, dass sich für die Skalen Streben nach sozialer Anerkennung, soziale Vergleiche sowie Selbstkonzept und Selbstwert mittlere Werte für die Validitätskoeffizienten feststellen lassen. Für die Skala Konformitätsdruck und Gruppendynamik konnten lediglich niedrige Werte erreicht werden. Außerdem konnten im Hinblick auf die Skala Anschlussmotiv keine Befunde zur Validierung ermittelt werden.

Abstract

In this master's thesis, a questionnaire for determining motives of self-presentation on social networks is examined and reviewed for the first time. The examination of the questionnaire is carried out in two steps: First, an initial review of the item characteristics and relevant quality criteria will be conducted. Subsequently, initial findings will be compiled to determine the criterion validity of the questionnaire. Accordingly, the work consists of two separate studies. Both studies are based on a non-experimental quantitative research approach and a cross-sectional survey is conducted to collect the data. In the first study (N = 308), the initial survey and item analysis as well as tests of the reliability of the questionnaire are conducted. Subsequently, the questionnaire was shortened and restructured. The final version consists of seven factor-analysed scales. In the second study (N = 112) the item and scale values as well as reliability coefficients were examined again in order to replicate the results of the first study at best. In the second step, the questionnaire was validated using various external criteria. For this purpose, the following validation scales were integrated into the questionnaire: The Neuroticism, Extraversion, and Achievement Motive scales of the B5-T (Satow, 2020), the Fear of Negative Appraisal scale (SANB-5) (Kemper, Lutz, & Neuser, 2011), the Rosenberg Self-Esteem Scale (RSES) (Collani & Herzberg, 2003), the Connection & Sociability scale of the TEMEO-R (Deutscher & Mees, 2012), and the Orientation to Others scale of the SMSC (Collani & Stürmer, 2009). These show content and theoretical references to the scales. Accordingly, the correlations between the scales and the external criteria are examined. In summary, with regard to the validation of the scales, it can be reported that for the scales Striving for Social Recognition, Social Comparisons, and Self-Concept and Self-Esteem, medium values for the validity coefficient can be determined. For the scales Conformity Pressure and Group Dynamics only low values could be achieved. Furthermore, no validation findings could be determined with regard to the scale Connection Motive.

Inhaltsverzeichnis

1 Einleitung	1
2 Theoretischer Hintergrund	2
2.1 Theoretische Basis und Operationalisierung des Fragebogens	2
2.1.1 Selbstdarstellung	2
2.1.2 Modelllernen	2
2.1.3 Motive	3
2.1.4 Selbstbezogene Kognitionen	5
2.1.5 Mitmenschen	8
2.1.6 Narzissmus	12
2.1.7 Eskapismus	13
2.1.8 Fear of missing out	14
2.2 Testtheoretische Grundlagen	15
2.2.1 Begriffsbestimmung und Gegenstand einer Testtheorie	15
2.2.2 Die Klassische Testtheorie	16
2.2.3 Probabilistische Testtheorien (PTT)	22
2.2.4 Gegenüberstellung der KKT und der IRT	25
2.2.5 Gütekriterien	25
2.3 Zielsetzung	30
3 Studie I	30
3.1 Testtheoretischer Hintergrund – Studie I	30
3.1.1 Itemanalyse	31
3.1.2 Reliabilität und interne Konsistenz	35
3.1.3 Faktorenanalyse	35
3.1.4 Zielsetzungen	41
3.2 Methode – Studie I	42
3.2.1 Forschungsdesign	42
3.2.2 Anforderungen an die Stichprobe und Rekrutierung	42
3.2.3 Messinstrument	43
3.2.4 Datenerhebung	43

	3.2.5 Statistische Analysen	. 44
	3.3 Ergebnisse – Studie I	. 45
	3.3.1 Zusammensetzung der Stichprobe	. 45
	3.3.2 Itemanalyse und -selektion	. 46
	3.3.3 Interne Konsistenz der theoretisch angenommenen Skalen	. 49
	3.3.4 Exploratorische Faktorenanalysen	.50
	3.3.5 Skalenzusammensetzung und Kennwerte des revidierten Fragebogens	55
	3.4 Diskussion und Interpretation der Ergebnisse – Studie I	. 58
	3.4.1 Fragebogenanalyse und -revision	. 59
	3.4.2 Interpretation und Definition der Faktoren	. 60
	3.4.3 Kritische Reflexion – Studie I	. 64
	3.5 Fazit – Studie I	. 66
4	Studie II	. 66
	4.1 Testtheoretischer Hintergrund – Studie II	66
	4.1.1 Bedeutung der Validität für die psychometrische Diagnostik	67
	4.1.2 Prüfung der Validität	. 67
	4.1.3 Zielsetzung	.68
	4.2 Methode – Studie II	. 73
	4.2.1 Forschungsdesign	.73
	4.2.2 Anforderungen an die Stichprobe und Rekrutierung	. 73
	4.2.3 Messinstrumente	.74
	4.2.4 Datenerhebung	.76
	4.2.5 Statistische Analysen	. 77
	4.3 Ergebnisse – Studie II	. 77
	4.3.1 Zusammensetzung der Stichprobe	. 77
	4.3.2 Itemkennwerte des FBMSN-R	. 79
	4.3.3 Reliabilitätsanalyse	. 80
	4.3.3 Auswertung der Hypothesen zur Bestimmung der Kriteriumsvalidität	. 81
	4.4 Diskussion und Interpretation der Ergebnisse – Studie II	. 88
	4.4.1 Zusammenfassung und Interpretation der empirischen Befunde	. 88

4.4.2 Kritische Reflexion	94
4.5 Fazit – Studie II	95
5 Diskussion	96
5.1 Zusammenfassung und Interpretation der Ergebnisse	96
5.2 Kritische Schlussbetrachtung	98
5.2.1 Datenerhebung und Stichproben	98
5.2.2 Methodische Vorgehensweise	99
5.2.3 Reflexion der Ergebnisse	100
5.3 Schlussfolgerungen und Ausblick	101
6 Gesamtfazit	102
Literaturverzeichnis	VII
Anhang	XV
Anhangsverzeichnis	XV
Eidesstattliche Erklärung	XVI

Tabellenverzeichnis

Tab. 1: Kennwerte der Itemanalyse der Skalen des FBMSN
Tab. 2: Übersicht der Reliabilitätskoeffizienten (α) der Skalen des FBMSN vor und
nach der Itemselektion50
Tab. 3: Anzahl der zu extrahierenden Faktoren nach verschiedene Kriterien 52
Tab. 4: Rotierte Mustermatrix des FBMSN (Hauptkomponentenanalyse mit 7
Faktoren, Promax-Rotation)53
Tab. 5: Rotierte Mustermatrix des FBMSN (Hauptachsenanalyse mit 7 Faktoren,
Promax-Rotation)
Tab. 6: Skalen und Items des FBMSN-R inklusive Trennschärfe- und Reliabilitäts-
Koeffizienten56
Tab. 7: Interskalen-Korrelationen des FBMSN-R58
Tab. 8: Faktoren inklusive Itemanzahl60
Tab. 9: Itemanalyse Studie II inklusive der Trennschärfekoeffizienten aus Studie I 79
Tab. 10: Übersicht und Vergleich der Reliabilitätskoeffizienten80
Tab. 11: Korrelationsmatrix für die Skala Streben nach Sozialer Anerkennung 82
Tab. 12: Korrelationsmatrix für die Skala Soziale Vergleiche82
Tab. 13: Korrelationsmatrix für die Skala Konformitätsdruck und Gruppendynamik und
die Skala Anschluss & Kontaktbereitschaft (TEMEO-R)83
Tab. 14: Korrelationsmatrix für Skala Konformitätsdruck und Gruppendynamik,
SMSC: Orientierung an Anderen und SANB-585
Tab. 15: Korrelationsmatrix für die Skala Selbstkonzept und Selbstwert 86
Tab. 16: Korrelationsmatrix für die Skala Anschlussmotiv
Tab. 17: Übersicht über die Validitätskoeffizienten 88

Abbildungsverzeichnis

Abb. 1: Das Fogg Behavior Model (Fogg, 2009)11
Abb. 2: Darstellung des Standardfehlers als Abweichung vom wahren Wert im
entsprechend Konfidenzintervall (Pospeschill, 2022)21
Abb. 3: Beispielhafter Verlauf der logistischen IC-Funktionen rasch-homogener Items
(Pospeschill, 2022)24
Abb. 4: a: Beispiel für eine eindeutige Entscheidung durch Scree-Test. b: Beispiel
ohne deutlichen "Knick" (Brandt, 2020)
Abb. 5: Verteilung nach Geschlecht - Studie I45
Abb. 6: Nutzung von Social Media Plattformen – Studie I
Abb. 7: Verteilung nach Geschlecht - Studie II78
Abb. 8: Nutzung von Social Media Plattformen – Studie II

Abkürzungsverzeichnis

B5-T Big-Five-Persönlichkeitstest

CFA Konfirmatorische Faktorenanalyse

EFA Exploratorische Faktorenanalyse

FBMSN Fragebogen zur Erfassung von zugrundeliegenden Motiven der

Selbstdarstellung auf Sozialen Netzwerken

FBMSN-R Fragebogen zur Erfassung von zugrundeliegenden Motiven der

Selbstdarstellung auf Sozialen Netzwerken - Revidierte Version

FOMO Fear of missing out

IRT Item-Response-Theorie

KTT Klassische Testtheorie

MLF Maximum-Likelihood-Faktorenanalyse (maximum likelihood factor

analysis)

PCA Hauptkomponentenanalyse (principal components analysis)

PFA Hauptachsenanalyse (principal axes factor analysis)

PTT Probabilistische Testtheorie

RSES Rosenberg Self-Esteem-Scale

SANB-5 Skala Angst vor Negativer Bewertung

SMSC Fragebogen zum Verhalten in sozialen Situationen

TEMEO-R Fragebogen zur Erfassung telischer und metatelischer

Orientierungen in revidierter Form

1 Einleitung

"There are only two industries that call their customers `users': illegal drugs and software" (Orlowski, 2020).

Dieses Zitat aus der Dokumentation "Das Dilemma mit den sozialen Medien" (im Original The Social Dilemma) zieht einen erschreckenden Vergleich und weist damit auf eine bedenkliche Entwicklung in der heutigen Gesellschaft hin. Tatsache ist, dass Social-Media-Plattformen uns süchtig machen und dabei auf dieselben neurophysiologischen Mechanismen zurückgreifen wie Drogen. So folgen alle Plattformen demselben Prinzip wie Drogen und Rauschmittel: Sie sprechen das Lust- und Motivationszentrum des Gehirns an, wodurch Dopamin ausgeschüttet wird. Egal, ob wir Likes, Thumbs Ups oder andere Buttons und Emojis als Reiz präsentiert bekommen, immer erfahren wir kurze Dopaminausstöße (Arte, 2019). Grundsätzlich verfolgen Social-Media-Plattformen genau ein übergeordnetes Ziel: Den Konsumenten möglichst lange an den Bildschirm zu fesseln. Dabei bedienen sich die Plattformen verschiedener psychologischer Konzepte. So zielt beispielsweise Instagram auf das Bedürfnis nach sozialer Anerkennung ab, welches durch Likes befriedigt werden kann (Arte, 2019).

Im Grunde ist jeder Social Media Nutzer süchtig nach Dopamin. Doch nun stellt sich die entscheidende Frage, welche Motive der Nutzung von sozialen Netzwerken zugrunde liegen und letztlich dazu verleiten, sich dieser Sucht hinzugeben. Burdorf und Lux (2022) entwickeln auf Basis dieser Fragestellung einen Fragebogen, um die zugrundeliegenden Motive der Selbstdarstellung auf sozialen Netzwerken zu erfassen.

Die vorliegende Masterthesis soll den entwickelten Test indessen überprüfen und optimieren, um bestenfalls eine fertige Version zu liefern, die den Ansprüchen der Testtheorie genügt und für Forschungszwecke freigegeben werden kann. Dazu werden zwei separate Studien konzipiert. Zunächst soll in der ersten Studie eine erste Item- und Reliabilitätsanalyse sowie eine entsprechende Revision des Fragebogens erfolgen. Anschließend liegt der Fokus der zweiten Studie auf der Validierung der revidierten Version des Fragebogens. Für jede Studie wurde eine eigene Stichprobe gezogen.

An dieser Stelle ist darauf hinzuweisen, dass in dieser Thesis auf die gleichzeitige Verwendung von personenbezogenen Sprachformen verzichtet wird. Demzufolge wird die männliche Sprachform für sämtliche Personenbezeichnungen und Pronomen verwendet. Dies soll keine Diskriminierung oder Benachteiligung anderer Geschlechter darstellen, sondern ausschließlich der besseren Lesbarkeit dienen. Entsprechend gelten sämtliche Personenbezeichnungen für alle Geschlechter.

2 Theoretischer Hintergrund

Das folgende Kapitel beschäftigt sich mit der theoretischen Basis der in dieser Forschungsarbeit untersuchten Materie. Um ein umfangreiches Verständnis für die zu erfassenden Konstrukte zu gewährleisten, behandelt der erste Teil des Kapitels die theoretischen Fundamente der Skalen des Fragebogens zur Erfassung von zugrundeliegenden Motiven der Selbstdarstellung auf sozialen Netzwerken. Überdies sollen die testtheoretischen Grundlagen und Annahmen, auf denen diese Forschungsarbeit sowie die Überprüfung des Fragebogens basiert, ausführlich erläutert und ausgearbeitet werden. Am Ende des Kapitels werden die Zielsetzungen dieser Thesis vorgestellt, welche auf den vorgestellten theoretischen Inhalten fußen.

Fragebogen zur Erfassung von zugrundeliegenden Motiven der Selbstdarstellung auf sozialen Netzwerken wird im Folgenden durch die Bezeichnung FBMSN abgekürzt.

2.1 Theoretische Basis und Operationalisierung des Fragebogens

In Rahmen der nachfolgenden Abschnitte werden die von Burdorf und Lux (2022) operationalisierten Konstrukte thematisiert. Im Zuge dessen wird zunächst der Begriff "Selbstdarstellung" definiert. Weiterhin werden die zugrundeliegenden Theorien der einzelnen Skalen kurz und prägnant geschildert, um ein grundlegendes Verständnis für den Gegenstand des entwickelten Fragebogens zu gewährleisten.

2.1.1 Selbstdarstellung

Das Phänomen der Selbstdarstellung beschreibt die strategische Art und Weise wie Menschen versuchen, den Eindruck, den sie auf Andere machen, zu steuern. Wir versuchen uns so zu präsentieren, wie wir glauben zu sein. Einerseits zielen Menschen mit ihrem Verhalten in sozialen Situationen auf eine externe Bestätigung des eigenen Selbstkonzeptes ab (Konsistenzaspekt der Selbstdarstellung). Andererseits versuchen Menschen sich selbst in Szene zu setzen, um eine positive Resonanz des Publikums bzw. der Mitmenschen hervorzurufen (selbstwertdienlicher Aspekt der Selbstdarstellung) (Neyer & Asendorpf, 2018; Morf & Koole, 2014).

2.1.2 Modelllernen

Burdorf und Lux (2022) erarbeiten in ihren Fallanalysen die Skala *Modelllernen*, welche im Wesentlichen auf Albert Banduras *sozial-kognitiver Lerntheorie* und dem *Lernen am Modell*, basiert. Die konkreten Items sind dem Anhang A zu entnehmen.

Das Lernen durch Beobachtung von Modellen stellt eine schnelle und effiziente Form des menschlichen Lernens dar. So wird das Verhalten zunächst von einem Modell vorgeführt und anschließend vom Beobachter nachgeahmt (Gerrig, 2015). Laut Bandura

(1977) ist diese Form des Lernens der wesentliche Ursprung erworbener Verhaltensmuster. Grundsätzlich setzt sich der Prozess des Modelllernens aus vier Komponenten zusammen. Diese sind Aufmerksamkeit, Speicherung sowie Produktion und Motivation (Rammsayer & Weber, 2016). Grundvoraussetzung für den Lernprozess ist die vollkommene Ausrichtung der Aufmerksamkeit seitens des Beobachters auf das Verhalten des Modells sowie dessen Folgen (Kiesel & Koch, 2012). Dieser Prozess der Aufmerksamkeitsausrichtung wird sowohl von Ereignis- (Zugänglichkeit, Komplexität, Bedeutung und Instrumentalität) als auch von Beobachtermerkmalen (kognitive Fähigkeiten, Aktivierungsniveau und Wahrnehmungspräferenzen) beeinflusst. Der Speicherungsprozess ist im Wesentlichen von den kognitiven Fähigkeiten und Informationsverarbeitungsprozessen des Beobachters abhängig. Weiterhin gilt die motorische Reproduktionsfähigkeit als wichtiger Faktor. Schließlich sollte der Beobachter über die motorischen Fähigkeiten verfügen, um das Verhalten des Modells ausüben zu können. Hier wirken sich sowohl Feedback als auch intensives Üben positiv auf die Modifikation des Verhaltens aus (Rammsayer & Weber, 2016; Kiesel & Koch, 2012). Als unabdingbar für die tatsächliche Reproduktion des beobachteten Verhaltens beschreibt Bandura die Motivation. Diese kann durch äußere Anreize (Verstärker) sowie stellvertretende Verstärkung (z. B. resultierende Verhaltensfolgen für das Modell) gesteigert werden (Rammsayer & Weber, 2016).

2.1.3 Motive

Im wissenschaftlichen Kontext wird klar zwischen Motiven und Motivation unterschieden. Nachfolgend werden verschiedene Definitionen und Betrachtungsweisen der Motivation vorgestellt, um eine eindeutige Differenzierung zwischen Motivation und Motiven zu gewährleisten.

Grundsätzlich gilt Motivation als ein zeitlich begrenzter Prozess, welcher durch verschiedene interne und externe Faktoren beeinflusst wird. Entsprechend kann Motivation als gegenwärtige Aktivierung zur Erreichung eines bestimmten Zieles beschrieben werden (Neyer & Asendorpf, 2018). Weiterhin ist Motivation nicht direkt beobachtbar. Entsprechend ist die Motivation als hypothetisches Konstrukt bzw. Abstraktion zu betrachten (Benecke & Brauner, 2017). Allerdings kann Motivation auch eine angeborene Bedürfnishierarchie repräsentieren. Laut Maslow (1981) ist eine vollständige Befriedigung der Bedürfnisse ein seltener Zustand, der nur kurz aufrechterhalten werden kann.

Allgemein beschreibt Heckhausen (1980) Motive als individuelle Bewertungsdispositionen für bestimmte Handlungsfolgen. Entsprechend gelten Motive als zeitlich stabile Persönlichkeitsdispositionen, die in bestimmten Situationen das Verhalten und Erleben lenken (Hofer & Hagemeyer, 2018; Heckhausen & Heckhausen, 2018a).

Eine erste Spezifikation einzelner Motive geht auf Murray (1938) zurück. Dieser definiert zunächst Leistungs- (Streben nach Erfolg) sowie Macht- (Streben nach Einfluss) und Anschlussmotiv (Streben nach sozialen Interaktionen) als grundlegende Motive der Persönlichkeit. Im Zuge weiterer Forschungen wird diese klassische Motivtrias (McClelland, 1985; Murray, 1938) von McAdams (1980) durch das Intimitätsmotiv (Streben nach Vertrautheit und Nähe) ergänzt (Neyer & Asendorpf, 2018). Im Allgemeinen setzen sich die Motive aus zwei Komponenten zusammen. Diese sind definiert als Annäherungsund Vermeidungskomponente. Die Motivkomponenten sind maßgeblich für die Ausrichtung des Verhaltens in bestimmten Situationen (Heckhausen & Heckhausen, 2018b).

Im Rahmen der Fallanalysen von Burdorf und Lux (2022) ergeben sich lediglich Hinweise für das Anschluss- und Intimitätsmotiv. Daher wird auf eine ausführliche Behandlung der anderen Motive verzichtet und der Fokus auf die genannten Motive gerichtet.

2.1.3.1 Anschlussmotiv

Für diese Skala konzentrieren sich die Autorinnen auf das konkrete Interaktionsverhalten der Probanden und soziale Beziehungen auf Social-Media-Plattformen (Burdorf & Lux, 2022). Die Items für die Skala *Anschlussmotiv* sind im Anhang A hinterlegt.

Wie bereits im vorangegangenen Abschnitt angedeutet wurde, zeigen anschlussmotivierte Menschen eine ausgeprägte Präferenz für soziale Interaktionen und Kontakt mit unbekannten und bekannten Personen. Bei diesem Motiv werden die Komponenten Hoffnung auf Anschluss als Annäherungskomponente sowie Angst vor Zurückweisung als Vermeidungskomponente unterschieden (Brandstätter, Schüler, Puca, & Lozo, 2018). Konkret zeigen annäherungsmotivierte Personen ein besonderes Streben nach Gesellschaft und positiven sozialen Interaktionen. Dagegen neigen vermeidungsmotivierte Menschen aus der Angst vor Zurückweisung dazu, Konflikte zu vermeiden (Neyer & Asendorpf, 2018). Weiterhin zeigt Gable (2006), dass Menschen mit einer stärkeren Vermeidungstendenz intensiver auf Konflikte reagieren. Entsprechend sind diese Personen langfristig unzufriedener in ihren Beziehungen. Außerdem ist den Ausführungen von Asendorpf (1989) zu entnehmen, dass die beiden Komponenten dieses Motivs deutliche Bezüge zu den Charaktereigenschaften Geselligkeit und Schüchternheit aufweisen.

2.1.3.2 Intimitätsmotiv

Im Zuge der Fallanalyse leiten Burdorf und Lux (2022) primär solche Items für die Skala *Intimitätsmotiv* ab, die sich neben der Tendenz, sich freizügig auf Social Media zu präsentieren, auf den Aufbau intimer Liebesbeziehungen fokussieren. Die entsprechenden Items sind dem Anhang A zu entnehmen.

Das Intimitätsmotiv zeichnet sich durch ein Streben nach respektive eine Präferenz für Nähe und Vertrautheit mit anderen Menschen aus. Dabei bildet der Wunsch nach Erfahrungen des innigen und vertrauten Austausches von Gedanken und Gefühlen den Kern dieses Motivs. Entsprechend gehen diese Formen der sozialen Interaktion und Beziehung in ihrer Intensität weit über die, im Anschlussmotiv postulierten, gewünschten Erfahrungen hinaus (Hofer & Hagemeyer, 2018). Die Annäherungskomponente zeigt sich in der Hoffnung auf Anschluss oder in einem warmen und engen Kontakt mit anderen Menschen. Dagegen äußert sich die Vermeidungskomponente in einer Furcht vor Distanz und Einsamkeit (Neyer & Asendorpf, 2018). Die Forschung zeigt, dass Personen mit einem ausgeprägten Intimitätsmotiv Sexualität nutzen, um ihr Streben nach Intimität und Nähe zu befriedigen. Demgegenüber nutzen vermeidungsmotivierte Menschen Sexualität zur Vorbeugung von Einsamkeit und Ablehnung (Impett, Strachman, Finkel, & Gable, 2008).

2.1.4 Selbstbezogene Kognitionen

In ihren Ausarbeitungen gehen Burdorf und Lux (2022) auf verschiedene Facetten der selbstbezogenen Kognitionen ein. Im Folgenden soll auf die für den generierten Fragebogen relevanten Aspekte eingegangen werden. Die Skalen der Autorinnen zielen hier vorwiegend auf selbstwertrelevante Aspekte ab, die durch das individuelle Verhalten und Interaktionen auf Social-Media-Plattformen beeinflusst werden.

2.1.4.1 Selbstwahrnehmung

Im Rahmen ihrer Fallanalyse formulieren Burdorf und Lux (2022) für die Skala *Selbstwahrnehmung* vier Items, diese sind dem Anhang A zu entnehmen.

Im Fokus der Selbstwahrnehmung steht die Wahrnehmung des eigenen Körpers sowie physiologischer Prozesse und des eigenen Verhaltens. Durch verschiedenste Sinnesmodalitäten (z. B. visuelle, akustische und kinetische Wahrnehmungen) sammeln wir permanent Informationen über unser Aussehen und unser Verhalten. Diese Informationen sind jedoch in unterschiedlichem Ausmaß verzerrt. Schließlich sind unsere sensorischen Wahrnehmungsquellen nicht frei von Wahrnehmungstäuschungen. Das gilt vorwiegend für viele physiologische Prozesse, da wir über keinen direkten sensorischen Zugang verfügen. Das individuelle Selbstkonzept ist abhängig von der Genauigkeit der Selbstwahrnehmung, da diese eine wesentliche Informationsquelle darstellt. Die Genauigkeit der Selbstwahrnehmung ist bei Menschen häufig unterschiedlich stark ausgeprägt (Neyer & Asendorpf, 2018). Weiterhin liefert die Selbstwahrnehmung kein objektives Bild der Wirklichkeit, sondern ist durch unsere Erwartungen gesteuert. Entsprechend sammeln Menschen sensorische Information, um sie mit den Erwartungen über diese

Informationen zu vergleichen. Dies hat zur Folge, dass Menschen dazu tendieren, sich selbst so zu sehen, wie sie glauben zu sein (Neyer & Asendorpf, 2018).

2.1.4.2 Soziales Spiegeln

Aus den Fallanalysen leiten die Autorinnen die im Anhang A aufgelisteten Items für die Skala *Soziales Spiegeln* ab (Burdorf & Lux, 2022).

Unter dem sozialen Spiegeln wird eine weitere Informationsquelle des Selbstbildes verstanden, welche aus Interaktionen mit anderen Menschen gewonnen wird. Wir schließen aus der Reaktion des Gegenübers auf unser Selbstbild. Folglich tendieren wir dazu, uns selbst so zu sehen, wie wir glauben, dass Mitmenschen uns sehen. Jedoch ist festzuhalten, dass das Konzept des sozialen Spiegelns lediglich eine Projektion des eigenen Selbstbildes auf andere Menschen darstellt. Aus dieser subjektiven Perspektive resultiert die Annahme, dass andere Menschen uns so sehen, wie wir uns selbst wahrnehmen (Neyer & Asendorpf, 2018).

2.1.4.3 Sozialer Vergleich

Im Rahmen ihrer Fallanalysen generieren Burdorf und Lux (2022) acht Items, welche im Anhang A konkret aufgelistet sind. Die Items für die vorliegende Skala zielen hauptsächlich auf die Ergebnisse sozialer Vergleiche auf sozialen Netzwerken sowie deren Auswirkung auf das Selbstwertgefühl ab.

Als weitere wesentliche Quelle des menschlichen Selbstverständnisses dient der Vergleich der eigenen Merkmale und Fähigkeiten mit denen unserer Mitmenschen. Als Begründer dieses Phänomens gilt Leon Festinger, welcher in seiner Theorie des sozialen Vergleichs drei zugrundeliegende Motive beschreibt (Festinger, 1954). Dabei liegt der Fokus auf den Fähigkeiten ähnlicher Anderer, um einen nützlichen und interpretierbaren Bezugsrahmen zu finden, wenn objektive Standards fehlen (Morf & Koole, 2014). Festinger (1954) postuliert zunächst das Motiv akkurater Selbsteinschätzung, welches darauf abzielt, ein solides sowie realistisches Selbstbild zu konstruieren (Crusius, 2022). Zur Veranschaulichung dieses Motivs wird das Big Fish Little Pond Modell herangezogen, um den Einfluss des sozialen Vergleichs auf das Selbstwertgefühl zu verdeutlichen. In diesem Modell fungiert der kleine Fisch in einem großen Teich als Metapher für einen Vergleichsprozess mit einer Gruppe, welche in der relevanten Fähigkeit besser ist. Daraus resultiert ein niedrigeres Selbstwertgefühl. Ein höheres Selbstwertgefühl entsteht dagegen bei einem Vergleich mit einer Gruppe, welche in der relevanten Fähigkeit schlechter ist als das Individuum. Entsprechend dem Modell wird von einem großen Fisch in einem kleinen Teich gesprochen (Neyer & Asendorpf, 2018; Morf & Koole, 2014). Der direkte Einfluss des sozialen Vergleichs auf das Selbstwertgefühl ist in

verschiedenen Studien untersucht und bestätigt worden (z. B. Marsh & Hau, 2004; Asendorpf & van Aken, 1993).

Als zweites Motiv des sozialen Vergleichs wird das Motiv der Selbstverbesserung beschrieben. Den Kern dieses Motivs bildet die Wahl der Bezugs- oder Vergleichsgruppe eines Individuums, da diese Auswahl sowohl Einfluss auf die Motivation als auch auf die Zielsetzung des Individuums hat (Crusius, 2022). In der Praxis kann dieser Vergleichsprozess mit der gewählten Referenzgruppe zu einer gesteigerten Motivation und somit zu einer Möglichkeit, die eigenen Fähigkeiten zu verbessern führen (Morf & Koole, 2014). Im Kontext des Motives Streben nach einem positiven Selbstwert postulieren Guenther und Alicke (2010) den Better-than-average-effect. Dieser beschreibt das Phänomen, dass Menschen fortlaufend nach einem positiven Selbstbild streben und versuchen dieses aufrechtzuerhalten (Guenther & Alicke, 2010).

Zusammenfassend ist hervorzuheben, dass das Selbstwertgefühl bei einem sozialen Vergleichsprozess davon abhängig ist, wie das Individuum annimmt, sich von der Bezugsgruppe zu unterscheiden (Neyer & Asendorpf, 2018). Grundsätzlich kann jeder Mensch oder jede Gruppe als Vergleichsobjekt dienen. In Bezug auf Banduras Theorie des *Lernens am Modell* (siehe Kapitel 2.1.2 Modelllernen) ist die Identifikation des Individuums mit dem Vergleichsmodell bei der Auswahl entscheidend. Durch diese Identifikation und dem resultierenden Vergleich mit dem Modell kann eine erhebliche Steigerung der Leistungsmotivation erzielt werden (Crusius, 2022; Kiesel & Koch, 2012).

In der Literatur werden zwei Formen sozialer Vergleiche unterschieden. Zum einen gilt der soziale Aufwärtsvergleich als Vergleich mit ähnlichen Menschen, die im Hinblick auf ein relevantes Merkmal besser sind als das Individuum selbst. Der soziale Abwärtsvergleich dagegen beschreibt den Vergleich mit Personen, die sich in einer offenbar schlechteren Situation befinden (Peter, Fahr, & Früh, 2012). In einer Metaanalyse der Theorie des sozialen Vergleichs zeigt sich die allgemeine Tendenz, dass Menschen sich mit Personen vergleichen, die ihnen in irgendeiner Form überlegen sind, selbst wenn eine Bedrohung des Selbstwertgefühls vorliegt. Diese Vergleiche ziehen eine geringere Einschätzung der eigenen Fähigkeiten sowie eine schlechtere Stimmung nach sich (Gerber, Wheeler, & Suls, 2018).

2.1.4.4 Selbstkonzept

Die Items der Skala *Selbstkonzept* konzentrieren sich in erster Linie auf Reaktionen in Folge von Verletzungen des Selbstkonzeptes im Kontext der Nutzung von Social-Media-Plattformen (Burdorf & Lux, 2022). Die Items sind im genauen Wortlaut im Anhang A zu finden.

In der Sozialpsychologie wird das Selbstkonzept ganz allgemein als kognitive Repräsentation der eigenen Selbstkenntnis, die sich aus der Summe aller Informationen (im

Hinblick auf körperliche Erscheinung, Eigenschaften, Ziele, Motive, Bedürfnisse, Fähigkeiten und Beziehungen zu anderen), die das Individuum über sich selbst hat, zusammensetzt, definiert. Das Selbstkonzept verleiht den individuellen Erfahrungen sowie den zugehörigen Beziehungen zu anderen Menschen Bedeutung und Kohärenz (Morf & Koole, 2014).

Die persönlichkeitspsychologische Forschung geht auf William James zurück, dessen Ansätze bis heute gültig sind. James beschreibt das / und das Me als zwei Komponenten des Selbst. Das I wird in diesem Zusammenhang als Urheber des eigenen Handelns sowie Wissens beschrieben und konstruiert die Selbstdefinition. Wohingegen das Me den Gegenstand der Selbstdefinition darstellt (Neyer & Asendorpf, 2018; Rammsayer & Weber, 2016). Nach James bildet das Selbstkonzept einen dispositionalen Teil des Me (Never & Asendorpf, 2018). Laut Shavelson, Hubner und Stanton (1976) bestimmt das Selbstkonzept die Handlungsinterpretationen eines Individuums und bildet eine Moderatorvariable zu dessen Erklärung. Weiterhin wird zwischen kognitiven und evaluativen Merkmalen differenziert (Stiller & Alfermann, 2005). Des Weiteren postulieren Shavelson und Kollegen (1976) ein Modell, welches das allgemeine Selbstkonzept in die Kategorien akademisches und nicht akademisches Selbstkonzept unterteilt. Entsprechend werden durch das akademische Selbstkonzept die Bewertungen des eigenen Verhaltens sowie Leistungen in Schulfächern repräsentiert. Das nicht akademische Selbstkonzept bildet dagegen die Facetten soziales (Freunde Andere) sowie emotionales (Gefühle) und physisches Erscheinungsbild) Selbstkonzept ab (Shavelson, Hubner, & Stanton, 1976).

2.1.5 Mitmenschen

Burdorf und Lux (2022) fassen im Rahmen ihrer Fallanalysen die Konstrukte, welche im Wesentlichen abhängig von unseren Mitmenschen sind, zusammen. Im nachfolgenden Kapitel werden diese Konstrukte vorgestellt.

2.1.5.1 Gruppendynamik

Für die Operationalisierung dieser Skala werden relevante Theorien sowie Konzepte der Phänomenologie von Gruppen geschildert und ausgearbeitet. Weiterhin sollen wesentliche Begrifflichkeiten definiert und erläutert werden. Im Anhang A ist eine Tabelle hinterlegt, welcher die generierten Items für die Skala *Gruppendynamik* zu entnehmen sind (Burdorf & Lux, 2022). Die formulierten Items legen den Fokus vordergründig auf Phänomene des Konformitätsdrucks und versuchen das Verhalten auf sozialen Netzwerken in diesem Zusammenhang zu erfassen.

Grundsätzlich wird eine Gruppe als eine Anzahl von zwei oder mehr Individuen verstanden, die sich selbst als Mitglieder einer Gruppe verstehen, definiert. Weiterhin zeichnet

sich eine Gruppe durch besondere Gesetzmäßigkeiten, die von allen Mitgliedern akzeptiert werden, aus. Außerdem bilden direkte Interaktionen und Kontakt der Gruppenmitglieder den Kern für das Fortbestehen der Gruppe. Des Weiteren zeichnet sich die Gruppe durch eine Binnenstruktur aus, welche Rollen sowie Positionen und Aufgaben der einzelnen Mitglieder definiert. Überdies hat die Gruppe ein bestimmtes Ziel, auf das sich die Gruppenmitglieder bewusst oder unbewusst geeinigt haben. Aus diesem gemeinsamen Ziel wird bei den Mitgliedern ein *Wir-Gefühl* hervorgerufen (Nijstad & van Knippenberg, 2014). Der Grad der Zusammengehörigkeit innerhalb der Gruppe wird durch die Gruppenkohäsion bzw. Kohärenz bestimmt. Konkret bildet dies die Anziehungskraft der Gruppe, welche die Mitglieder an die Gruppe bindet (Nijstad & van Knippenberg, 2014).

Mit dem Fortbestand einer Gruppe entwickeln sich häufig bestimmte Verhaltensregeln sowie -vorschriften. Damit einher geht ein gewisser Druck (Konformitätsdruck), diese Regeln zu befolgen und sich konform zu verhalten. Widersetzt sich ein Gruppenmitglied diesen Regeln, kann dies Strafen sowie Überzeugungsversuche und gegebenenfalls den Ausschluss aus der Gruppe, nach sich ziehen (Hewstone & Martin, 2014).

Durch die Gruppenmitgliedschaft wird bei den Mitgliedern häufig ein Gefühl der Überlegenheit gegenüber anderen Individuen, außerhalb der Gruppe hervorgerufen (Ardelt-Gattinger, Lechner, & Schlögl, 1998). Janis (1991) beschreibt das Phänomen des Gruppendenkens (groupthink). Dieses tritt primär dann auf, wenn die Mitglieder einer Gruppe die eigenen Überzeugungen für allgemeingültig halten und sich gegenüber externen Meinungen sowie Kritik verschließen. Das Syndrom des Gruppendenkens hat häufig Beeinträchtigungen des Denkens sowie Störungen des Bewertens alternativer Handlungsverläufe der einzelnen Gruppenmitglieder, zur Folge (Janis, 1991).

2.1.5.2 Mehrheitseinfluss

Auf Basis ihrer Fallanalysen definieren Burdorf und Lux (2022) die Skala *Mehrheitsein-fluss* über acht Items (siehe Anhang A). Die theoretische Grundlage der Skala bilden hauptsächlich die Erkenntnisse zum Mehrheitseinfluss von Solomon E. Asch (1955). An dieser Stelle ist darauf hinzuweisen, dass auf eine detaillierte Beschreibung der Experimente von Salomon Asch verzichtet und lediglich auf die wesentlichen Ergebnisse sowie Schlussfolgerungen eingegangen wird.

Die Untersuchungen von Asch (1955) basieren auf den Ergebnissen von Mustafa Sherif. In seinen Experimenten zur Entwicklung individueller Normen aus dem Jahr 1935, ließ Sherif die Versuchspersonen schätzen, wie weit sich Lichtpunkte voneinander entfernt haben. Im Anschluss sollten die Probanden in Kleingruppen aus drei Personen über die Distanz diskutieren. Auf Grundlage seiner Experimente kommt Sherif zu dem Schluss, dass die Mehrheit maßgeblich für die Gruppennorm ist. Konkret zeigt sich diese

Beobachtung darin, dass die Entscheidung der Gruppe dazu tendiert, dass sich eine dritte Person der mehrheitlichen Schätzung anpasst (Germar, 2016).

Die Ergebnisse von Aschs Experimenten zur Wirkung von Mehrheitsaussagen auf die Meinung einer Einzelperson stützen das Phänomen des Mehrheitseinflusses. Aus den Ergebnissen geht hervor, dass sich 36,8 % der 123 Probanden der fehlerhaften Mehrheitsmeinung anpassten. Entsprechend können die Versuchspersonen in zwei Gruppen geteilt werden: Jene, die sich der Mehrheit fügten und diejenigen, die standhaft auf der eigenen Meinung beharrten (Asch, 1955). Auf der Verhaltensebene konnten diese Befunde in weiteren Studien repliziert werden (Germar, 2016).

Aus den infolge der Experimente durchgeführten Einzelgesprächen mit den Teilnehmern geht hervor, dass das Vertrauen in die eigenen Fähigkeiten und die eigene Person maßgeblich für die Beständigkeit des eigenen Urteils ist (Asch, 1955).

2.1.5.3 Minderwertigkeit

Die Items der Skala *Minderwertigkeit* zielen auf eine Einschätzung der Ausprägung des Minderwertigkeitsgefühls ab (Burdorf & Lux, 2022). Die konkreten Items für die Skala befinden sich in Anhang A. Zur Operationalisierung der Skala *Minderwertigkeit* wird von Burdorf und Lux (2022), das Modell der Individualpsychologie nach Alfred Adler herangezogen. Im Rahmen des nachfolgenden Abschnitts wird auf eine detaillierte Beschreibung der Theorie verzichtet. Stattdessen liegt der Fokus auf den Aspekten und Inhalten der Individualpsychologie, welche für die vorliegende Skala die höchste Relevanz aufweisen.

Alfred Adler gilt als Begründer der Individualpsychologie. Diese betrachtet das Individuum in seiner Ganzheitlichkeit, d. h. der Mensch wird als unteilbare Einheit bestehend aus Körper und Psyche gesehen. Daraus resultiert die Annahme, dass Störungen in einem Bereich (z. B. körperliche Beeinträchtigungen) sich auf andere Lebensbereiche (z. B. Minderwertigkeitsgefühle) auswirken. Grundsätzlich postuliert die Individualpsychologie ein positives Menschenbild. Nach Adler wird der Mensch als soziales Wesen geboren, das nach Geltung und Selbstverwirklichung strebt. Außerdem geht Adler davon aus, dass das Individuum grundsätzlich sowohl zum Guten als auch zum Bösen fähig ist. Wobei die guten Tendenzen durch einen liebevollen Umgang gefördert werden können. Entsprechend bildet die Kindeserziehung eine wesentliche Präventionsmöglichkeit, um Entwicklungsstörungen vorzubeugen. In diesem Zusammenhang geht Adler auf verschiedene Erziehungsfragen ein, um die Ausbildung einer stabilen Psyche und Persönlichkeit zu fördern (Hannich, 2018).

Den Kern der Persönlichkeitstheorie nach Adler bilden das *Minderwertigkeitsgefühl und* seine Kompensation. Zunächst beschränkt sich Adler auf die Organminderwertigkeit¹, welche ein Gefühl der Minderwertigkeit nach sich ziehen kann, welches durch das betroffene Individuum kompensiert werden will. Dabei sind Vergleichsprozesse, mit der konkreten Umwelt des Individuums oder mit individuellen Zielen im Hinblick auf die Zielerreichung, maßgeblich für die Stärke des Minderwertigkeitsgefühls. Später erweitert Adler diesen Ansatz und berücksichtigt in der Folge alle Minderwertigkeitsgefühle, die aus sozialer oder psychischer Unfähigkeit resultieren (Hannich, 2018).

Eine weitere Komponente Individualpsychologie nach Alfred Adler bildet das *Konzept der Finalität*. Die Individualpsychologie geht davon aus, dass menschliches Verhalten stets zielgerichtet ist und der Erreichung eines gewünschten Soll-Zustandes dient. Für ein umfassendes Verständnis des menschlichen Handelns stellt Adler neben der Wirkursächlichkeit, vor allem die Frage nach der Zweckursächlichkeit in den Vordergrund. Das Ziel selbst fungiert in einer kompensatorischen Rolle, da seine konkreten Inhalte sowie antizipierten Folgen, beim Individuum positive Emotionen auslösen (Hannich, 2018).

2.1.5.4 Soziale Anerkennung

Im Zuge der Fallanalysen leiten die Autorinnen des Fragebogens acht Items für die Skala Soziale Anerkennung ab (Burdorf & Lux, 2022). Die konkreten Items sind dem Anhang A zu entnehmen. Die Operationalisierung der Skala basiert insbesondere auf dem Verhaltensmodell von Brian J. Fogg (Fogg Behavior Model). Entsprechend wird das Modell im folgenden Abschnitt umfassend beschrieben und ausgearbeitet. Im Zuge dessen wird ebenfalls auf eine passende grafische Darstellung zurückgegriffen, um einen einfachen sowie verständlichen Zugang zu dem Modell zu gewährleisten (siehe Abb. 1).

Das Modell dient dem Verständnis und Modifikation des menschlichen Verhaltens. Nach Fogg (2009) wird das menschliche Verhalten durch die Komponenten Motivation sowie Fähigkeit und Auslöser (triggers) bestimmt (Fogg, 2009). Diese Komponenten lassen sich grafischen abbilden (siehe Abb. 1). Laut Fogg (2009) bildet das Verhältnis der

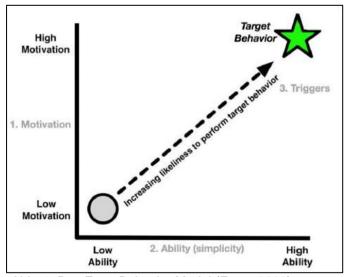


Abb. 1: Das Fogg Behavior Model (Fogg, 2009).

¹ Organminderwertigkeit: Angeborene oder später entwickelte körperliche Funktionsstörung und/oder Missbildung, aus der eine Minderung der physischen Belastbarkeit resultiert (Hannich, 2018).

individuellen Ausprägungen einer Person auf den Achsen Motivation (motivation) und Fähigkeit (ability), die Grundvoraussetzung für die Ausübung des gewünschten Zielverhaltens (target behavior). Entsprechend ist die Wahrscheinlichkeit oder Bereitschaft für die Ausführung des Verhaltens umso größer, je höher die Ausprägungen auf den beiden Achsen ausfällt. Ob das Zielverhalten tatsächlich gezeigt wird, ist abhängig von einem konkreten Auslöser (trigger). Dieser ist in der Grafik kurz vor dem Zielverhalten lokalisiert (siehe Abb. 1) und ist für die Ausübung des Zielverhaltens unabdingbar (Fogg, 2009). Im zweiten Teil dieses Abschnitts wird auf eine detaillierte Ausführung der drei Komponenten verzichtet. Stattdessen richtet sich der Fokus auf den Faktor der Motivation, da dieser Faktor den Gegenstand der konturierten Skala bildet.

Für den Faktor Motivation differenziert Fogg (2009) drei Kernmotivatoren. Diese lassen sich jeweils in einer dimensionalen Struktur aus zwei entgegengesetzten Polen abbilden (Fogg, 2009).

Der erste Kernmotivator wird durch die Komponenten Vergnügen und Schmerz repräsentiert. Diese Komponenten bilden zwei fest verankerte unmittelbare Reaktionen des Individuums auf bestimmte Situationen ab, welche in direktem Zusammenhang mit Selbsterhaltung respektive Vermehrung stehen. Konkret äußert sich dies in der Fortsetzung bzw. Unterbrechung des Verhaltens beim Empfinden von Vergnügen bzw. Schmerz (Fogg, 2009).

Den zweiten Motivator bilden Hoffnung und Angst (bzw. Verlust), welcher stets an einen bestimmten Zustand (erwünscht vs. gefürchtet) gekoppelt ist. Im Alltag wirkt dieser Motivator meist stärker als der Erste, da seine Wirkung sich über einen längeren Zeitraum erstreckt (Fogg, 2009).

Der Bezugsrahmen des letzten Motivators umfasst nach Fogg (2009) das gesamte Sozialverhalten des Individuums. Menschen streben nach sozialer Anerkennung und Bestätigung, um einen positiven Selbstwert aufrechtzuerhalten und ihre Identität zu festigen. Entsprechend zielen sämtliche Handlungen des Individuums auf die Anerkennung und den Zuspruch anderer Menschen hab. Den Gegenpol bildetet die soziale Ablehnung, welche primär durch die Angst vor Zurückweisung gekennzeichnet ist. Diese Angst ist ebenfalls auf evolutionäre Entwicklungen zurückzuführen. Schließlich ist der Mensch ein soziales Wesen, welches in Gruppen lebt und diese zum Überleben benötigt. Entsprechend werden sozial anerkannte Verhaltensweisen aufrechterhalten, um das Bedürfnis nach sozialer Anerkennung zu befriedigen, sowie die Akzeptanz der Gruppe zu gewährleisten (Fogg, 2009).

2.1.6 Narzissmus

Auf Basis der Fallanalysen konturieren Burdorf und Lux (2022) die Skala *Narzissmus*. Im folgenden Abschnitt wird die theoretische Basis der Operationalisierung des

Konstruktes vorgestellt. Im Anhang A sind die Items für die Skala *Narzissmus* nach Burdorf und Lux (2022) übersichtlich dargestellt.

Der Begriff Narzissmus wird in der Persönlichkeits- und Sozialpsychologie losgelöst von seiner klinisch-pathologischen Bedeutung verstanden. So wird Narzissmus in diesem Zusammenhang als eine Persönlichkeitsdimension definiert, die sich durch eine Tendenz zur Selbstüberschätzung, ein schwach ausgeprägtes Empathievermögen sowie eine Überempfindlichkeit gegenüber Kritik und eine starke Selbstbezogenheit auszeichnet (Neyer & Asendorpf, 2018; Lammers, Vater, & Roepke, 2013). Weiterhin machen Narzissten auf fremde Personen zunächst einen positiven Eindruck, welcher auf ihrem Charme sowie Kleidung und der scheinbaren Selbstsicherheit beruht (Dufner, Rauthmann, Czarna, & Denissen, 2013). Langfristig kann eine hohe Ausprägung dieser Persönlichkeitsdimension eher negativ auf das Umfeld wirken, da der Narzisst zu starken Stimmungsschwankungen neigt, wenn sein starkes Bedürfnis nach Wertschätzung nicht ausreichend gestillt wird (Neyer & Asendorpf, 2018; Lammers, Vater, & Roepke, 2013).

Die klinische Psychologie differenziert zwischen der gesunden und der pathologischen Ausprägung von Narzissmus. Die gesunde Ausprägung beschreibt, wie bereits thematisiert, eine bestimmte Persönlichkeitsdimension. Wohingegen die pathologische Ausprägung eine Persönlichkeitsstörung darstellt. Weiterhin werden im klinischen Kontext die drei Narzissmussubtypen, *grandios-maligne* sowie *vulnerabel-fragile* narzisstische Persönlichkeitsstörung und die exhibitionistische narzisstische Persönlichkeitsstörung mit hohem Funktionsniveau, unterschieden (Lammers, Vater, & Roepke, 2013). Aus epidemiologischer Sicht ergibt sich aus Untersuchungen von Bevölkerungsstichproben für die narzisstische Persönlichkeitsstörung eine Prävalenzrate² von unter 1 % (Fiedler, 2011). Aus den Daten von Trull et al. (2010) geht hervor, dass Männer (Prävalenz von 1,2 %) häufiger betroffen sind als Frauen (Prävalenz von 0,7 %) (Lammers, Vater, & Roepke, 2013).

2.1.7 Eskapismus

Burdorf und Lux (2022) generieren auf Grundlage ihrer Fallanalysen neun Items, welche die Skala *Eskapismus* repräsentieren (siehe Anhang A).

Grundsätzlich ist das Konstrukt des Eskapismus kaum zu operationalisieren, da eine einheitliche Definition dieses Phänomens bisher fehlt (Dimbath, 2015). Entsprechend werden im folgenden Abschnitt verschiedene Definitionen aus der wissenschaftlichen Literatur vorgestellt.

² Prävalenzrate: Prozentsatz aller Krankheitsfälle in einer bestimmten Population zu einem bestimmten Zeitpunkt bzw. innerhalb einer Zeitperiode (Wittchen & Jacobi, 2011).

Das Phänomen des Eskapismus wird als eine Form der Realitätsflucht dargestellt, welche nach Bonfadelli und Friemel (2017) dazu dient, den Alltag sowie die damit einhergehenden Sorgen und Probleme zu vergessen oder zu verdrängen. Weiterhin haben Menschen das Gefühl, sich entspannen und runterfahren zu können, weil in der alternativen Realität oder Welt zumeist Spaß und Vergnügen im Vordergrund stehen (Walden, 2020). Das Phänomen der Flucht in eine alternative Realität, welche vor allem Spaß bietet, bezeichnet Dimbath (2015) als Vergnügungs- und Zerstreuungssucht. Des Weiteren können gesellschaftliche Anforderungen sowie Normen, denen sich das Individuum verpflichtet fühlt und denen es gerecht werden will, Druck erzeugen. Übersteigt dieser Druck die eigenen Kapazitäten, so erscheint der Rückzug in eine alternative Realität verlockend (Dimbath, 2015). Ob eskapistisches Verhalten als dysfunktional angesehen wird oder nicht, ist abhängig vom zeitlichen Ausmaß des Verhaltens. Entsprechend kann das Fluchtverhalten dann als problematisch eingestuft werden, wenn es das Individuum fortlaufend an einer erfolgreichen Bewerkstelligung des Alltags in der tatsächlichen Realität hindert (Kuhlmann & Gehrau, 2011).

Zusammenfassend kann der Eskapismus als eine bewusste Handlung verstanden werden, die darauf abzielt, dem Alltag zu entfliehen sowie dessen scheinbare Bedrohungen zu verdrängen.

2.1.8 Fear of missing out

Das Konstrukt Fear of missing out (FOMO) oder Angst etwas zu verpassen wird trotz des offensichtlichen Fehlens einer direkten Verbindung zur Selbstdarstellung in die Operationalisierung einbezogen. Burdorf und Lux (2022) leiten aus der Analyse der Interviews ab, dass ein Zusammenhang zwischen der Erstellung sowie Nutzung von "Fake-Profilen" und dem Konstrukt FOMO besteht. Entsprechend generieren sie zwei Items für diese Skala (siehe Anhang A).

Allgemein beschreibt FOMO die Angst, ein spannendes Ereignis zu verpassen sowie das Gefühl, dass andere Personen ein ereignisreicheres und interessanteres Leben haben, als das Individuum selbst (Przybylski, Murayama, DeHaan, & Gladwell, 2013). In Bezug auf soziale Netzwerke resultiert häufig das Bestreben, den Kontakt zu anderen Menschen durchgängig aufrechtzuerhalten, um zu wissen, was sie erleben (Milyavskaya, Saffran, Hope, & Koestner, 2018). Häufig bezieht sich FOMO auf bestimmte Gruppen, denen das Individuum angehört (z. B. Freunde). In diesem Zusammenhang können die unangenehmen Gefühle noch stärker ausfallen, da Menschen grundsätzlich danach streben, die Bindung zu einer Gruppe aufrechtzuerhalten (Abel, Buff, & Burr, 2016).

Das Konzept FOMO gewinnt vor allem in der heutigen Zeit immer mehr an Bedeutung. Gründe dafür sind hauptsächlich der nahezu permanente Internetzugriff und

verschiedenste Social Media Plattformen (z. B. *Facebook*, *Instagram*, *TikTok* oder *Twitter*), welche das Leben anderer Personen in virtueller Form immer verfügbar machen. Ebenso führt die Etablierung und Nutzung von Smartphones dazu, dass Menschen nicht nur zu jeder Tageszeit, sondern auch an jedem Ort, Zugriff auf entsprechende Informationen über das Leben anderer haben (Abel, Buff, & Burr, 2016).

2.2 Testtheoretische Grundlagen

Im folgende Kapitel wird die testtheoretischen Basis der Fragebogenkonstruktion thematisiert. Entsprechend sollen wesentliche Grundannahmen der Klassischen Testtheorie herausgearbeitet und umfassend beschrieben werden, um die Basis für eine nachvollziehbare Prüfung des vorliegenden Messinstruments zu schaffen. Im Zuge dessen erfolgt ebenfalls eine kritische Auseinandersetzung mit der KTT. Ferner erfolgt die Darstellung von probabilistischen Ansätzen bzw. Modellen der Item-Response-Theorie als eine Alternative oder auch Ergänzung zur KTT.

2.2.1 Begriffsbestimmung und Gegenstand einer Testtheorie

In der psychologischen Forschung dienen verschiedene Testverfahren, wie Leistungstests, Fragebögen sowie standardisierte Interviews und Beobachtungen der Erfassung von bestimmten *Konstrukten*. Ein Konstrukt wiederum fasst bestimmte menschliche Verhaltens- und Erlebensweisen sowie typische Reaktionsdispositionen zu einem Merkmalskonzept zusammen. Ein Konstrukt (auch *latente Variable*) ist nicht direkt messbar und muss daher über verschiedene *Indikatoren* (*Testitems*) erfasst werden. Dabei dienen die Antworten der numerischen Erfassung des Verhaltens einer Testperson. Für die Konstruktion von Messinstrumenten, die ein bestimmtes Konstrukt abbilden sollen, ist von *Eindimensionalität* auszugehen. Entsprechend soll nur eine latente Variable für unterschiedliche Itemantworten verschiedener Personen verantwortlich sein. Folglich repräsentieren die Items diese latente Variable und sind somit als *reflektive Indikatoren* zu bezeichnen (Bühner, 2011).

Allerdings können Konstrukte auch aus *formativen Indikatoren* bestehen. Diese sind jedoch nicht Gegenstand der Klassischen Testtheorie. Entsprechend wird auf eine ausführliche Darstellung verzichtet und nur auf die Kernunterschiede eingegangen. Wie bereits erwähnt, basiert die Klassische Testtheorie (und viele probabilistische Modelle) auf eindimensionalen Messungen. In der Praxis geht eine Änderung des Konstrukts mit einer Veränderung der Ausprägung auf allen Items (Indikatoren) einher, da die latente Variable (Konstrukt) die Zusammenhänge der Indikatoren beschreibt. Wohingegen formative Indikatoren das Konstrukt erklären. Konkret bedeutet das, dass eine Änderung des Konstrukts auf nur einen bestimmten Indikator zurückgehen kann (Bühner, 2011).

Weiterhin sind bestimmte Gütekriterien definiert, die eine theoretische Orientierung bieten und testtheoretisch konstruiert sind. An diesen Kriterien müssen sich sämtlich psychometrische Testverfahren messen lassen, die den Anspruch haben, Merkmalsausprägungen zu messen oder Zuordnungen zu bestimmten Personenklassen vorzunehmen (Bühner, 2011).

2.2.2 Die Klassische Testtheorie

Die Klassische Testtheorie (KTT) bildet seit über 70 Jahren das theoretische Fundament für die Konstruktion psychodiagnostischer Testverfahren und deren Interpretation. Weiterhin sind viele bewährte Ansätze zur Prüfung verschiedener Gütekriterien von Messinstrumenten auf Basis der KTT konstruiert worden. Überdies besteht ein wesentlicher Vorteil der KTT in ihrer ökonomischen und praktikablen Anwendbarkeit (Pospeschill, 2022; Moosbrugger & Kelava, 2020).

Die KTT stellt im Wesentlichen eine Messfehlertheorie dar. Entsprechend ist das erklärte Ziel, die Trennung vom Messfehler und *wahrem Wert (True-Score)* (Moosbrugger, Schermelleh-Engel, Gäde, & Kelava, 2020).

2.2.2.1 Axiome der KTT

Die Grundannahmen der KTT sind als Axiome formuliert. Im folgenden Abschnitt werden die drei Axiome sowie weitere Zusatzannahmen beschrieben.

Zunächst wird das *Existenzaxiom* als erste Annahme der KTT postuliert. So geht aus diesem Axiom hervor, dass der *wahre Wert* τ_{vi} als *Erwartungswert E* der *Messung* x_{vi} einer *Versuchsperson* v in einem *Item i* existiert:

$$\tau_{vi} = E(x_{vi})$$

Konkret meint das, dass aus einer bestimmten Menge von Messwiederholungen auf die tatsächliche Merkmalsausprägung (wahrer Wert) im Hinblick auf ein Item geschlossen werden kann (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020).

Das zweite Axiom ist das *Verknüpfungsaxiom*. Dieses formuliert die Annahme, dass jede Messung x_{vi} den wahren Wert τ_{vi} sowie die Messfehlervariable (*Zufallsfehler*) ε_{vi} beinhaltet:

$$x_{vi} = \tau_{vi} + \varepsilon_{vi}$$

Entsprechend beinhaltet jede Messung die Summe aus der Messung des wahren Wertes (tatsächliches Merkmal) und der Messung eines Messfehlers (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020).

Aus der Verbindung von Existenz- und Verknüpfungsaxiom folgt, dass der Zufallsfehler einen Erwartungswert von null ausweist:

$$E(\varepsilon_{ij}) = 0$$

Entsprechend tendiert der Zufallsfehler bei unendlich vielen Wiederholungen der Messung gegen null (Pospeschill, 2022; Bühner, 2011).

Das Unabhängigkeitsaxiom postulliert, dass die wahren Werte τ und die Zufallsfehler ε im Hinblick auf beliebige Personen sowie beliebige Items unkorreliert sind:

$$Corr(\tau_{vi}, \varepsilon_{vi}) = 0$$

Entsprechend besteht keine spezifische Abhängigkeit zwischen der tatsächlichen Merkmalsausprägung und den Messfehlern (Pospeschill, 2022).

Unter den *Zusatzannahmen* fasst Pospeschill (2022) Unabhängigkeitsforderungen der KTT hinsichtlich der paarweisen Unabhängigkeit der Messfehler in Bezug auf zwei Personen sowie ein Itempaar zusammen. Entsprechend geht die KTT von einer Unkorreliertheit der Messfehler zwischen verschiedenen Testpersonen *v* und *w* für das gleiche Item *i* aus:

$$Corr(\varepsilon_{v_i}, \varepsilon_{w_i}) = 0$$

Überdies wird die Forderung der Unkorreliertheit der Messfehler in verschiedenen Tests oder Items formuliert. Entsprechend wird davon ausgegangen, dass bei zwei Items i und j die Messfehler ϵ_{vi} und ϵ_{vj} unabhängig sind (Pospeschill, 2022):

$$Corr(\varepsilon_{v_i}, \varepsilon_{v_i}) = 0$$

2.2.2.2 Bestimmung des wahren Wertes

Zur Bestimmung der wahren Testwerte erfolgt die Messung der latenten Merkmale anhand einer bestimmten Anzahl von eindimensionalen Items. Dieses Vorgehen entspricht einer Mehrfachmessung desselben Merkmals. Anschließend werden die Itemvariablen zu einer Testvariablen aufsummiert, um den Zufallsfehler zu neutralisieren. Entsprechend stellt der $Testwert x_v$ (Rohwert) den Summenwert der einzelnen Itemwerte x_{vi} dar (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020):

$$x_{v} = \sum_{i=1}^{m} x_{vi}$$

Anschließend wird der Erwartungswert des Testwerts x_v untersucht, um nachzuweisen, dass dieser dem wahren Wert τ_v entspricht. Entsprechend gilt, nach einsetzen in die Gleichung:

$$E(x_v) = E(\sum_{i=1}^m x_{vi}) = \sum_{i=1}^m E(x_{vi}) = \sum_{i=1}^m \tau_{vi} = \tau_v$$

Der Erwartungswert von x_v stimmt mit dem wahren Wert τ_v überein. Folglich kann der Summenwert x_v als Punktschätzung des wahren Wertes τ_v einer Person v genutzt werden (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020):

$$x_{v} = \hat{\tau}_{v}$$

An dieser Stelle weisen Moosbrugger und Kelava (2020) darauf hin, dass eine Punktschätzung nur bei Vorliegen einer hohen Reliabilität interpretiert werden kann.

2.2.2.3 Bestimmung von wahrer Varianz und Fehlervarianz

Die Varianzschätzung von wahrer Varianz $Var(\tau)$ und Fehlervarianz $Var(\epsilon)$ erfolgt auf Basis der Testwerte x der gesamten Stichprobe n. Entsprechend folgt die modellbasierte Varianzerlegung (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020):

$$Var(x) = Var(\tau + \varepsilon) = Var(\tau) + Var(\varepsilon) + 2 \cdot Cov(\tau, \varepsilon)$$

Gemäß dem Unabhängigkeitsaxiom sind Messfehler und wahrer Wert unkorreliert. Daraus folgt: $Cov(\tau, \varepsilon) = 0$. Damit gilt:

$$Var(x) = Var(\tau) + Var(\varepsilon)$$

Dem zur Folge ergibt sich die Testwertvarianz Var(x) aus wahrer Varianz $Var(\tau)$ und Fehlervarianz $Var(\epsilon)$ (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020).

2.2.2.4 Bestimmung der Reliabilität

Die Reliabilität gilt als Maß der Messgenauigkeit eines Tests. Diese lässt sich durch den Anteil der wahren Varianz an der Gesamtvarianz (beobachtete Varianz) bezeichnen und stellt einen Determinationskoeffizienten³ dar (Pospeschill, 2022; Moosbrugger & Kelava, 2020):

$$Rel = \frac{Var(\tau)}{Var(x)}$$

Dieser Reliabilitätskoeffizient Rel bestimmt die Messgenauigkeit des Verfahrens und kann Werte zwischen Null (keine Reliabilität) und Eins (höchste Reliabilität) annehmen.

³ Determinationskoeffizient: Kennwert der Regressionsanalyse, welcher eine Beurteilung der Vorhersagekraft einer abhängigen Variable durch die betrachtete unabhängige Variable (Eid, Gollwitzer, & Schmitt, 2017).

Entsprechend ist ein möglichst hoher Wert gewünscht (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020).

Wie bereits beschrieben, ist die wahre Varianz nicht messbar und muss daher geschätzt werden. Anhand der beobachteten Werte paralleler Tests (p und q) kann der Anteil der wahren Varianz an der beobachteten Varianz als *Test-Test-Korrelation* r_{tt} geschätzt werden (Pospeschill, 2022; Gäde, Schermelleh-Engel, & Werner, 2020):

$$Rel = \frac{Var(\tau)}{Var(x)} = \frac{Cov(x_p, x_q)}{SD(x_p) \cdot SD(x_q)} = Corr(x_p, x_q) = r_{tt}$$

Weitere differenzierte Methoden erlauben eine Bestimmung der Reliabilität aus der Testform selbst, ohne einen Paralleltest verwenden zu müssen (Pospeschill, 2022). Diese sind im Kapitelabschnitt 2.2.5.2 Reliabilität theoretisch beschrieben. Auf eine weitere Ausführung der testtheoretischen Inhalte bezüglich der Reliabilität wird an dieser Stelle verzichtet, da diese in Kapitel 3.1.2 Reliabilität und innere Konsistenz erneut aufgegriffen und ausführlich geschildert werden.

2.2.2.5 Reliabilität und Testlänge

Die Annahmen der KTT führen zu dem Schluss, dass mit zunehmender Zahl an Items zu einem bestimmten Merkmal die Genauigkeit der Messung steigt. Entsprechend kann von einer Reliabilitätssteigerung durch eine Verlängerung eines Tests durch parallele Testteile⁴ (weitere Items) gesprochen werden (Pospeschill, 2022).

Wenn ein Testteil p durch einen parallelen Teil q verlängert wird, ergibt sich die gemeinsame Testvarianz wie folgt (Pospeschill, 2022):

$$Var(x_p, x_q) = Var(\tau) + Var(\varepsilon) + Var(\tau) + Var(\varepsilon) + 2 \cdot Var(\tau) = 4 \cdot Var(\tau) + 2 \cdot Var(\varepsilon)$$

Aus dieser Gleichung geht hervor, dass bei einer Verdopplung (*k*=2) der Testlänge *I* durch einen parallelen Testteil die vierfache wahre Varianz resultiert (Pospeschill, 2022). Eine allgemeinere Darstellung liefert die *Spearman-Brown-Formel* (Fisseni, 2004):

$$r_{tt}(k \cdot l) = \frac{k \cdot r_{tt}}{1 + (k - 1) \cdot r_{tt}}$$

Weiterhin kann die Formel genutzt werden, um eine Reliabilitätsminderung bei einer Testverkürzung zu bestimmen (Pospeschill, 2022).

19

⁴ Parallele Testteile: Items mit gleichen wahren Werten und gleicher wahrer Varianz

2.2.2.6 Herleitung des Standardmessfehlers

Auf Basis des *Reliabilitätskoeffizienten Rel* lässt sich die Varianzzerlegung bei bekannter Testwertevarianz nach Pospeschill (2022) wie folgt vereinfachen:

$$Var(x) = Rel \cdot Var(x) + (1 - Rel) \cdot Var(x) = Var(\tau) + Var(\varepsilon)$$

Die Gleichung wird anschließend nach Var(ε) aufgelöst:

$$Var(\varepsilon) = Var(x) - Rel \cdot Var(x) = Var(x) \cdot (1 - Rel)$$

Daraus kann geschlossen werden, dass $Var(\varepsilon)$ der ungeklärte Anteil der Fehlervarianz der Testwertvarianz Var(x). Durch Ziehen der Wurzel lässt sich die Standardabweichung SD der Fehlervarianz, der sog. $Standardmessfehler\ SD(\varepsilon)$ bestimmen. Konkret verkleinert sich der Standardmessfehler mit wachsender Reliabilität, wodurch die Messgenauigkeit weiter steigt (Pospeschill, 2022):

$$SD(\varepsilon) = SD(x) \cdot \sqrt{1 - Rel}$$

Weiterhin kann für einen Test, welcher auf die Vorhersage eines bestimmten Kriteriums abzielt, der *Standardschätzfehler SD*(ε_{tc}) bestimmt werden. Dieser ermöglicht die Eingrenzung der Streuung um den wahren Kriteriumswert (Pospeschill, 2022):

$$SD(\varepsilon_{tc}) = SD(c) \cdot \sqrt{1 - Corr_{tc}^2}$$

Wobei SD(c) die Standardabweichung des Kriteriums sowie $Corr_{tc}$ die Korrelation zwischen Test und Kriterium darstellen (Pospeschill, 2022).

Nach Fisseni (2004) lässt sich aus der Berechnung des Standardmessfehlers die *kritische Differenz diff_{crit}* ableiten. Diese Gleichung ermöglicht die Signifikanzprüfung der Differenz von zwei Testwerten (Fisseni, 2004):

$$diff_{crit}(x_1 - x_2) = z_a \cdot SD(x) \cdot \sqrt{2 - (r_{tt_1} + r_{tt_2})}$$

Die Reliabilitäten der Tests mit den Testwerten x_1 und x_2 sind bezeichnet durch r_{tt1} sowie r_{tt2} . Die Differenz zweier Testwerte gilt als signifikant, wenn diese größer ist als $diff_{crit}$ (Pospeschill, 2022; Fisseni, 2004).

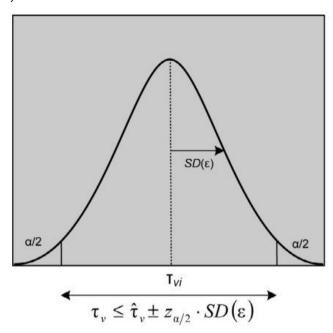
2.2.2.7 Konfidenzintervall des wahren Testwertes

Der Standardmessfehler $SD(\varepsilon)$ dient vorrangig der Schätzung eines Konfidenzintervalls des wahren Wertes τ_v (Pospeschill, 2022):

$$\tau_{v} \leq \hat{\tau}_{v} \pm z_{\frac{\alpha}{2}} \cdot SD(\varepsilon)$$

Entsprechend bestimmt das *Konfidenzintervall* den Vertrauensbereich, in dem 95 % (α = .05) bzw. 99 % (α = .01) aller wahren Wertes τ_{v} liegen, welche den Stichprobenschätzwert erzeugt haben. Zur Veranschaulichung ist dieser Sachverhalt in Abbildung 2 grafisch dargestellt (Pospeschill, 2022).

Grundvoraussetzung für die Bestimmung des Konfidenzintervalls ist eine Normalverteilung des Standardmessfehlers. Weiterhin wird eine Stichprobe von $n \ge 60$ verlangt. Alternativ kann bei einer Stichprobe von $n \le 60$ eine Betrachtung der t-Verteilung erfolgen. Des Weiteren muss eine Reliabilität von min. Rel ≥ .80 vorliegen. Denn eine reichende Reliabilität beein-Punktschätzungen. Woraus wiederum resultiert, dass das



flusst die Genauigkeit der Abb. 2: Darstellung des Standardfehlers als Abweichung vom wahren Wert im entsprechend Konfidenzintervall (Pospeschill, 2022).

Konfidenzintervall entsprechend breiter ausfällt (Pospeschill, 2022; Moosbrugger, Gäde, Schermelleh-Engel, & Rauch, 2020).

2.2.2.8 Kritische Betrachtung der KTT

Grundsätzlich liefert die KTT einen ökonomischen sowie praktikablen und vor allem bewährten theoretischen Ansatz zur Konstruktion von Testverfahren und der Interpretation der gemessenen Werte. Gleichwohl bestehen klare Schwächen und Einschränkungen der KTT, die ihrem Geltungsbereich ein klares Limit setzen (Pospeschill, 2022; Moosbrugger & Kelava, 2020).

Zunächst stellen wahrer Wert und Fehlerwert lediglich theoretische und damit nicht überprüfbare Größen dar. Entsprechend sind die Aussagen, die auf Grundlage des Vernüpfungsaxioms getroffen werden, lediglich Schätzungen (Pospeschill, 2022).

Einer der größten Kritikpunkte ist, dass die Axiome der KTT, wie die Unkorreliertheit von Messfehler und wahrem Wert, nicht empirisch prüfbar sind. Weiterhin stellt in diesem Zusammenhang auch die ausschließliche Anerkennung eindimensionaler Konstrukte einen Problemfaktor dar. Denn auch eine empirische Überprüfung der dieser Forderung findet in der KTT nicht statt (Bühner, 2011).

Außerdem geht die Klassische Testtheorie davon aus, dass die wahren Werte über mehrere Messungen konstant bleiben. Dieses Theorem kann maximal für kurze Zeiträume zwischen den Messungen gehalten werden. Schließlich muss zwangsläufig davon ausgegangen werden, dass bestimmte Merkmale sich über die Zeit verändern (Pospeschill, 2022; Moosbrugger & Kelava, 2020).

Ebenso geht die KTT von intervallskalierten Daten aus. Jedoch bietet die KTT keine Ansätze zur Prüfung des Skalenniveaus und ob dieses Kriterium über alle Tests gehalten werden kann, ist fraglich (Pospeschill 2022, Bühner, 2011).

Weiterhin sind die Parameter der KTT Population- und Stichprobenabhängig. Dies hat zur Folge, dass in Abhängigkeit von Heterogenität oder Homogenität der Stichprobe hinsichtlich eines Merkmals die Reliabilitätskoeffizienten unterschiedlich ausfallen (z. B. in Extremgruppen). Entsprechend ist die Generalisierbarkeit von Resultaten ebenfalls fraglich. Zudem gestaltet sich die Stichprobenabhängigkeit als besonders problematisch, für die Vergleichbarkeit über verschiedene Stichproben hinweg, da eine Stichprobe in weitere Substichproben zerfallen kann. Diese können wiederum hinsichtlich der Testwerte unterschiedliche Reliabilitäten und Validitäten aufweisen. Ferner lassen sich die Ergebnisse verschiedener Tests, die dasselbe Konstrukt messen wollen und auf Basis der KTT konstruiert wurden, nicht direkt vergleichen (Pospeschill 2022, Bühner, 2011).

Wie bereits beschrieben ist, der Rückschluss von Gruppenstatistiken auf den Einzelfall auf Basis der KTT grundsätzlich kritisch zu sehen, wenn die Reliabilität oder Validität unter Eins liegt. Denn eine Wahrscheinlichkeitsaussage bezieht sich immer auf Elemente mit einem bestimmten Umfang, kann aber nicht Gradmesser für die Vorhersagegenauigkeit eines einzelnen Falls sein (Pospeschill, 2022).

2.2.3 Probabilistische Testtheorien (PTT)

Probabilistische Testtheorien (PTT) oder auch Item-Response-Theorien (IRT) stellen eine wichtige und sinnvolle Ergänzung zur Klassischen Testtheorie dar. An dieser Stelle wird auf eine vollumfängliche Beschreibung und detaillierte Herleitung der PTT verzichtet, da die Überprüfung des Messinstruments in erster Linie auf Basis der KTT geschieht. Dennoch müssen vergleichbare Ansätze und Alternativen beleuchtet werden, um ein umfassendes Verständnis testtheoretischer Grundlagen zu gewährleisten. Entsprechend liefert dieses Kapitel einen Einblick in eine Alternative zur KTT. Wobei sich überwiegend auf das *Rasch-Modell* bezogen wird.

2.2.3.1 Allgemeine Annahmen probabilistischer Testtheorien

Die PTT bzw. IRT beschäftigt sich mit der folgenden Ausgangsfrage: Welche Rückschlüsse können bezüglich bestimmter Merkmale (Einstellungs- oder Fähigkeitsmerkmale) auf Basis der vorliegenden Antworten einer Testperson auf verschiedene Items

gezogen werden? Um diese Fragestellung zu beantworten, wird in probabilistischen Messmodellen zwischen *manifesten* und *latenten Variablen* unterschieden. Dadurch ist eine genaue Prüfung, ob alle Items das gleiche Merkmal messen, möglich. Manifeste Variablen ergeben sich durch beobachtbares Testverhalten einer Person in Form von spezifischen Itemantworten. Latente Variablen dagegen repräsentieren die Ausprägung nicht beobachtbarer Fähigkeits- und Persönlichkeitsmerkmale und sind von den manifesten Variablen abhängig. Weiterhin kann in spezifischen Modellen die Beziehung von diesen Variablen beschrieben werden. Überdies wird in Ansätzen der IRT eine *stochastische Beziehung* zwischen dem Antwortverhalten einer Testperson und dem definierten Modellparameter angenommen. Diese kann in einer *itemcharakteristischen Wahrscheinlichkeitsfunktion* (*IC-Funktion*) beschrieben werden (Pospeschill, 2022; Kelava & Moosbrugger, 2020b).

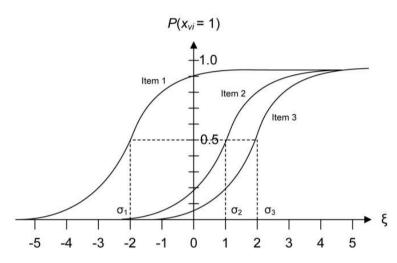
2.2.3.2 Das Rasch-Modell

Grundsätzlich zielt das *Rasch-Modell* darauf ab, wie das Antwortverhalten einer Testperson zustande kommt, dazu werden die Antwortmuster der Person untersucht. Diese Muster müssen wiederum einem bestimmten Modell folgen. Entsprechend wird ein Zusammenhang zwischen den Itemantworten (*manifeste Variablen*) und einem Personenmerkmal (*latente Variable*) ausgegangen. Konkret bedeutet das, dass mit höher Merkmalsausprägung der Person die Wahrscheinlichkeit einer Itemlösung steigt. Die Lösungswahrscheinlichkeit ist abhängig vom *Personenparameter* (z. B. Fähigkeit oder Eigenschaftsausprägung einer Person) und *Itemparameter* (z. B. Schwierigkeit eines Items). Diese Beziehung wird als *probabilistischer Zusammenhang* beschrieben, der sich über eine Wahrscheinlichkeitsfunktion (*IC-Funktion*) darstellen lässt. Entsprechend hat auch eine Person mit einer geringeren Merkmalsausprägung eine theoretische Chance ein Item zu lösen, das seine Fähigkeit übersteigt (Pospeschill 2022, Bühner, 2011).

Im Rahmen der IRT können unterschiedliche Modelltests durchgeführt werden. Wenn das Modell den Modelltest besteht und entsprechend beibehalten wird, bildet die Summe aus allen Itemantworten alle Informationen über die Merkmalsausprägung einer Person auf einer latenten Variablen ab. Dies hat zur Folge, dass der Summenwert einer Testperson eine *erschöpfende Statistik* der Personenfähigkeit darstellt. Konkret bedeutet erschöpfend, dass der Summenwert alle Informationen über die Merkmalsausprägung (z. B. bestimmte Fähigkeit) einer Person liefert. Daraus resultiert, dass eine Interpretation des Antwortmusters hinfällig ist, da diese keine weiteren Informationen über die Personenfähigkeit liefert (Bühner, 2011).

Bezogen auf die Eignung eines Items als Indikator für eine latente Variable kann festgehalten werden, dass die Lösung des Items ausschließlich auf die Fähigkeitsausprägung auf der latenten Variable zurückzuführen ist (und nicht auf eine andere Variable). Wenn die Items sich ausschließlich in ihrer Schwierigkeit unterscheiden, gilt *Rasch-Homogenität* für diese Items. Daher lassen sich *rasch-homogene Items* als einzelne Funktion auf einer gemeinsamen eindimensionalen Skala *(joint scale)* abbilden (Pospeschill 2022, Bühner, 2011).

In Abbildung 3 sind diese Funktionen beispielhaft dargestellt. Entsprechend sind die IC-Funktionen der einzelnen Items nur durch den Schwierigkeitsparameter σ_i verschoben (Pospeschill, 2022).



Zusammenfassend lassen sich folgende

Abb. 3: Beispielhafter Verlauf der logistischen IC-Funktionen rasch-homogener Items (Pospeschill, 2022).

Annahmen für das Rasch-Modell festhalten: Ein Test umfasst eine endliche Menge von Items. Weiterhin ist der Test homogen, d. h. dass alle Items dasselbe Merkmal messen. Außerdem sind die IC-Funktionen monoton steigend (d.h. je leichter das Item, desto höher ist die Lösungswahrscheinlichkeit). Überdies besteht *lokale stochastische Unabhängigkeit*. Im Wesentlichen besteht lokale stochastische Unabhängigkeit, wenn die Lösungswahrscheinlichkeit für ein Item ausschließlich von der Schwierigkeit und der Personenfähigkeit abhängig ist (Pospeschill, 2022; Kelava & Moosbrugger, 2020b).

Eine praktische Anwendung eines Modells der probabilistischen Testtheorie stellt das adaptive Testen dar. Hier handelt es sich um eine Form der Testung, bei der die Vorgabe der Items durch ein Regelsystem gesteuert wird und sich am Antwortverhalten des Probanden orientiert. Die praktische Umsetzung des Regelsystems (z. B. adaptiver Algorithmus) erfolgt zumeist computergestützt. Die Auswahl erfolgt aus einem bestimmten Itempool, wobei möglichst solche Items ausgewählt werden sollen, die die größte diagnostische Information liefern. Es erfolgt also eine sukzessive Präzisierung des geschätzten Personenparameters durch die Minimierung der redundanten Informationen. Dies geschieht durch die Vorgabe von Items mit entsprechend angepasster Itemschwierigkeit. Dieses Verfahren bietet vor allem eine besondere Steigerung der Messeffizienz und

ist äußerst ökonomisch für Testleiter sowie Probanden (Pospeschill, 2022; Kelava & Moosbrugger, 2020b).

2.2.4 Gegenüberstellung der KKT und der IRT

Im folgenden Kapitel erfolgt eine kurze Gegenüberstellung wesentlicher Unterschiede zwischen den testtheoretischen Ansätzen, um sowohl Stärken als auch Schwächen des jeweiligen Ansatzes herauszustellen.

Zunächst besteht ein wesentlicher Unterschied darin, dass die KTT bestimmte Forderungen (Axiome) formuliert, welche als allgemeingültig für die KTT gelten. Diese können jedoch nicht empirisch geprüft werden. Bei Ansätzen der PTT werden die definierten Annahmen empirisch geprüft (z. B. Itemhomogenität oder lokale stochastische Unabhängigkeit). Ein großer Nachteil der PTT im Vergleich zur KTT ist, dass eine Testkonstruktion auf Basis eines probabilistischen Ansatzes deutlich aufwendiger und zeitintensiver ist. Das liegt hauptsächlich daran, dass der verfügbare Itempool an modellkonformen Items lediglich in bestimmten sehr ausführlich erforschten Bereichen mit entsprechend ausgearbeiteten Konstrukten vorliegen. Innerhalb dieser Bereiche hat sich die PTT allerdings hervorragend bewährt. Eine besondere Stärke der PTT ist jedoch, dass die Parameter weniger Stichprobenabhängig sind als bei der KTT. Allerdings sind klassische Gütekriterien nicht einfach auf die PTT bzw. IRT übertragbar. Das liegt vordergründig daran, dass ein Test mit einem bestimmten Itemsatz in Abhängigkeit von der Testperson unterschiedlich reliabel ausfällt (Pospeschill 2022, Bühner, 2011).

Abschließend lässt sich dennoch festhalten, dass beide Theorien ihre eigenständige Daseinsberechtigung aufweisen, da sich die Ansätze in einigen Punkten vorteilhaft ergänzen (Moosbrugger & Kelava, 2020b).

2.2.5 Gütekriterien

Zur Einschätzung der Qualität eines Testverfahrens werden verschiedene anerkannte Gütekriterien betrachtet. So bilden die *Objektivität* sowie *Reliabilität* und *Validität* die traditionellen Hauptgütekriterien. Unter den Nebengütekriterien sind die Kriterien *Skalierung, Normierung, Ökonomie, Nützlichkeit, Zumutbarkeit* sowie *Fairness* und *Unverfälschbarkeit* eines Testinstrumentes zusammengefasst (Moosbrugger & Kelava, 2020a).

2.2.5.1 Objektivität

Das Kriterium der *Objektivität* ist definiert als Grad, zu dem die Testergebnisse einer Person unabhängig vom Untersuchenden ist. Im Zusammenhang mit psychometrischen Testverfahren werden die Formen *Durchführungs-*, *Auswertungs-* und *Interpretationsobjektivität* unterschieden.

Zur Gewährleistung der *Durchführungsobjektivität* müssen für alle Teilnehmer möglichst die gleichen Untersuchungsbedingungen geschaffen und eingehalten werden. Hier ist eine genaue Definition und klare Instruktion entscheidend. Entsprechend sollten die Bedingungen standardisiert sein, um jegliche Einflüsse (mit Ausnahme des zu messenden Merkmals) auf das Testverhalten konstant zu halten, damit die Ergebnisse nicht verzerrt werden (Moosbrugger & Kelava, 2020a).

Die Auswertungsobjektivität ist gegeben, wenn genaue Auswertungsanweisungen vorliegen und somit jede Person, die die Daten auswertet, den gleichen Punktwert für einer Versuchsperson ermittelt. Dabei ist primär das gewählte Antwortformat maßgeblich dafür, inwieweit dieses Kriterium als erfüllt betrachtet werden kann. Hier sind bei Tests Auswertungsschablonen sowie verhaltensverankerte Ratingskalen bei Verhaltensbeobachtungen oder Interviews bewährte Hilfsmittel, um die Auswertungsobjektivität zu sichern (Moosbrugger & Kelava, 2020a; Bühner, 2011).

Interpretationsobjektivität liegt vor, wenn die Interpretion der gleichen Testergebnisse unterschiedlicher Probanden im Hinblick auf das gemessene Merkmal unabhängig vom Auswerter ist. Entsprechend soll jeder Testanwender zu der gleichen Beurteilung der Ergebnisse kommen. Hier ist zu unterstreichen, dass sich die Interpretationsobjektivität nur auf das erfasste Merkmal und nicht auf die Schlussfolgerungen, die augrund der Merkmalsausprägung gezogen werden, bezieht (Moosbrugger & Kelava, 2020a; Bühner, 2011).

2.2.5.2 Reliabilität

Die Reliabilität bzw. Zuverlässigkeit bezeichnet den Grad der Genauigkeit eines Wertes, der im Rahmen einer Testung gemessen wird. Damit gilt dieses Kriterium als erfüllt, wenn der Test das zu messende Merkmal eindeutig erfasst. Um die Reliabilität einzuschätzen, wird der Reliabilitätskoeffizient (*Rel*) bestimmt. Dazu können verschiedene Verfahren herangezogen werden, welche auf den Annahmen der KTT beruhen. Im Folgenden sollen vier Verfahren der Reliabilitätsschätzung kurz vorgestellt werden (Moosbrugger & Kelava, 2020a; Bühner, 2011).

Die *Retest-Reliabilität* (Merkmalstabilität) wird durch die Korrelation der Testwerte derselben Stichprobe aus zwei verschiedenen Messzeitpunkten bestimmt. Dies geschieht unter der Annahme, dass zwischen den Testungen keine Veränderungen des Merkmals vorliegt. Allerdings ist zu berücksichtigen, dass die Korrelation von der Länge des Zeitintervalls abhängig ist. So können sowohl tatsächliche Ereignisse als auch Übungs- oder Erinnerungseffekte die Messwerte beeinflussen (Moosbrugger & Kelava, 2020; Bühner, 2011).

Eine weitere Methode ist die Bestimmung einer Paralleltestkorrelation, *Paralleltest-Reliabilität* genannt. Um diese zu Berechnen, werden zwei parallele Testformen

konstruiert, welche aus Itemzwillingen⁵ bestehen. Diese Testformen werden dann anhand einer Stichprobe zu zwei Messzeitpunkten erhoben. Entsprechend geht aus der Korrelation die Invarianz der Testwerte im Hinblick auf die Variation der Itemauswahl und der Messzeitpunkte hervor. Folglich stellt diese Methode einen Kennwert der Bedingungsstabilität dar, da sie die Kontrolle von Verzerrungen durch Übungs- oder Erinnerungseffekte sowie Merkmalsveränderungen ermöglicht (Moosbrugger & Kelava, 2020a; Bühner, 2011).

In der Praxis ist die Erstellung eines Paralleltests, aus verschiedenen Gründen oft schwierig umzusetzen, daher stellt die Bestimmung der *Split-Half-Reliabilität* eine nützliche Alternative dar. Hier wird der Test in zwei parallele Hälften unterteilt und miteinander korreliert. An dieser Stelle ist darauf hinzuweisen, dass die Reliabilität eines Tests mit wachsender Itemanzahl steigt. Folglich fällt Korrelation der Testhälften häufig niedriger als die tatsächliche Reliabilität des Gesamttests aus. Daher erfolgt eine Korrektur über die *Spearman-Brown-*Formel, um die Gesamtreliabilität zu berechnen (Moosbrugger & Kelava, 2020a; Bühner, 2011).

Eine der verbreitetsten Methoden ist die Bestimmung der internen oder auch inneren Konsistenz durch den Koeffizienten *Cronbachs Alpha* (α). Diese stellt eine Verallgemeinerung der Split-Half-Reliabilität dar. So wird jedes Item als eigener Testteil betrachtet und miteinander korreliert. Je stärker diese Teile positiv korrelieren, umso höher fällt die Reliabilität des Testwertes aus. Daher gibt jene die Genauigkeit des Wertes einer Messung zu einem bestimmten Zeitpunkt an (Moosbrugger & Kelava, 2020a; Bühner, 2011).

2.2.5.3 Validität

Als wichtigstes Hauptgütekriterium gilt die *Validität* oder auch Gültigkeit. Dieses gibt an, ob das Messinstrument das Merkmal misst, welches es messen soll oder zu messen beansprucht. Objektivität und Reliabilität sind notwendige, aber keine hinreichenden Bedingungen für eine hohe Validität (Moosbrugger & Kelava, 2020a).

In erster Linie werden drei unterschiedliche Aspekte der Validität betrachtet: *inhaltliche Validität* sowie *Kriteriums*- und *Konstruktvalidität*. Diese sind im Folgenden ausführlich beschrieben (Bühner, 2011).

Laut Moosbrugger und Kelava liegt *Inhaltsvalidität* vor, wenn "die Testitems im Zuge der Operationalisierung so konstruiert und ausgewählt werden, dass sie das interessierende Merkmal repräsentativ abbilden" (Moosbrugger & Kelava, 2020, S. 32).

An dieser Stelle zeigt sich, dass der Inhaltsvalidität eine besondere Bedeutung zuteil wird und sie sich in ihrer Feststellung klar von den anderen Formen abhebt. Ihre Prüfung ist empirisch nicht möglich, daher wird sie zumeist indirekt über die anderen beiden

_

⁵ Itemzwillinge: Inhaltlich und formal möglichst ähnliche Items

Aspekte erfasst. So bestimmen *Kriteriums*- und *Konstruktvalidität* nicht die Validität, wie sie allgemein definiert ist. Vielmehr geben sie Aufschluss über die Validität der Schlussfolgerungen, welche auf Basis der Kennwerte getroffen werden. Entsprechend stellen diese, genau genommen, validitätsbezogene Belege dar (Bühner, 2011).

Die Kriteriumsvalidität wird definiert als der Zusammenhang der Testergebnisse mit bestimmten Außenkriterien, mit denen das Instrument aufgrund des Messanspruches korrelieren sollte. Folglich liegt Kriteriumsvalidität vor, wenn vom gewonnenen Testwert auf ein Kriterium bzw. Verhalten außerhalb der Testsituation geschlossen werden kann (Moosbrugger & Kelava, 2020a; Bühner, 2011). Bühner (2011) unterscheidet verschiedene Formen der Kriteriumsvalidität, welche nachfolgend kurz beschrieben werden.

Zum einen bietet sich die Möglichkeit, Zusammenhänge mit zeitlich später erhobenen Kriterien zu untersuchen. Diese Form des Korrelationsschlusses wird als *Vorhersage-validität* oder auch *prognostische* oder *prädiktive Validität* bezeichnet. Werden die Kriterien nahezu zeitgleich mit den Testwerten erhoben, so wird die *Übereinstimmungsvalidität* oder auch *kongruente Validität* bestimmt. Die *retrospektive Validität* bezieht sich entsprechend auf die Korrelationen mit zeitlich vorher ermittelten Kriterien. Unabhängig von der zeitlichen Feststellung der Außenkriterien, bezeichnet die *inkrementelle Validität* den Beitrag oder Mehrwert an Informationen, den ein Test zur Vorhersage eines Kriteriums über ein anderes Instrument hinaus bietet (Bühner, 2011).

Um die Konstruktvalidität eines Messinstruments zu bestimmen, erfolgt eine Analyse der theoretischen Fundierung des Merkmals, welches der Test abbilden soll. Dabei liegt der Fokus auf der Dimensionalität und Struktur des Merkmals (Moosbrugger & Kelava, 2020a). Entsprechend kann ein Test als valide im Sinne des Konstrukts bezeichnet werden, wenn "die Zusammenhangstruktur zwischen den Testitems und den interessierenden (Persönlichkeits-) Merkmalen ("Konstrukte", "latente Variablen", "Traits", "latente Klassen", z. B. Fähigkeiten, Dispositionen, Charakterzüge oder Einstellungen) wissenschaftlich fundiert ist" (Moosbrugger & Kelava, 2020, S. 34). In der Praxis werden im Vorfeld bestimmte Erwartungen über die Zusammenhänge des vorliegenden Tests mit anderen Instrumenten formuliert. Unter der Konstruktvalidität sind die Aspekte konvergente sowie diskriminante bzw. divergente und faktorielle Validität zusammengefasst (Bühner, 2011).

Zur Feststellung der konvergenten Validität werden Zusammenhänge (Korrelationen) mit Verfahren eines ähnlichen Bereichs untersucht, um nachzuweisen, dass wirklich das interessierende Merkmal erfasst. Entsprechend werden hier hohe Zusammenhänge erwartet. Wohingegen die diskriminante bzw. divergente Validität sich auf Verfahren anderer Gültigkeitsbereiche bezieht. An dieser Stelle wird von niedrigen Zusammenhängen ausgegangen, da eine Abgrenzung zu ähnlichen Konstrukten erfolgen soll. Die faktorielle Validität basiert auf der Analyse der faktoriellen Struktur des Tests. So kann die

Beurteilung der Konstruktvalidität nach diesem Ansatz über struktursuchende und strukturprüfenden methodische Ansätze erfolgen. Struktursuchende Verfahren dienen in diesem Zusammenhang der Ableitung von Hypothesen über die Dimensionalität der Merkmalsstruktur. Während strukturprüfende Verfahren der statistischen Prüfung der angenommenen Dimensionalität dienen (Moosbrugger & Kelava, 2020a; Bühner, 2011).

2.2.5.4 Nebengütekriterien

In den folgenden Abschnitten sollen die einzelnen Nebengütekriterien *Skalierung, Normierung, Ökonomie, Nützlichkeit, Zumutbarkeit* sowie *Fairness* und *Unverfälschbarkeit* kurz beschrieben werden.

Das Gütekriterium der *Skalierung* bezieht sich auf die Gültigkeit der Verrechnungsregel, auf dessen Basis ein Testwert gewonnen wird. Folglich ist das Kriterium erfüllt, wenn "die laut Verrechnungsregel resultierenden Testwerte (numerisches Relativ) die tatsächlichen Merkmalsrelationen (empirisches Relativ) adäquat abbilden" (Moosbrugger & Kelava, 2020a, S. 20). Konkret bedeutet das, dass eine empirisch größere Ausprägung des Merkmals mit einem höheren Testwert einhergehen muss (Moosbrugger & Kelava, 2020a).

Ziel der *Normierung (Eichung)* eines Testverfahrens ist die Erstellung eines Bezugssystems, dass einen Vergleichsrahmen für die Testergebnisse einzelner Person bietet. Dazu muss eine repräsentative Stichprobe der Zielgruppe als Eichstichprobe gezogen werden. Dies erlaubt eine eindeutige Interpretation und Einordnung der Testwerte eines Probanden im Vergleich zur Zielgruppe (Moosbrugger & Kelava, 2020a).

Laut Moosbrugger und Kelava (2020a) gilt ein Test als ökonomisch, wenn der finanzielle und zeitliche Aufwand für die Durchführung der Testung in einem angemessenen Verhältnis zum diagnostischen Erkenntnisgewinn stehen.

Weiterhin erfüllt ein Testverfahren das Güterkriterium der Nützlichkeit, wenn der Test ein bestimmtes Merkmal erfasst, für das ein praktisches Bedürfnis besteht. Weiterhin müssen die auf der Basis des Tests abgeleiteten Maßnahmen einen adäquaten Mehrwert bzw. Nutzen bieten (Moosbrugger & Kelava, 2020a; Bühner, 2011).

Die Zumutbarkeit bezieht sich auf die zeitliche sowie psychische und physische Beanspruchung der Probanden. Entsprechend muss der erwartete Informationsgewinn der Testergebnisse in einem angemessenen Verhältnis zur Beanspruchung der Versuchsteilnehmer (Moosbrugger & Kelava, 2020a).

Die resultierenden Messwerte und Schlussfolgerungen bezogen auf eine Testperson dürfen keine Diskriminierungen aufgrund ihrer ethnischen sowie soziokulturellen und geschlechtsspezifischen Gruppenzugehörigkeit aufweisen. Ist dies gegeben, wird das

Kriterium der *Fairness* für ein Verfahren als erfüllt betrachtet (Moosbrugger & Kelava, 2020a; Bühner, 2011).

Um die *Unverfälschbarkeit* eines Tests zu gewährleisten, muss dieser so konstruiert sein, dass die Probanden die Ausprägungen ihrer Testwerte nicht willentlich oder unwillentlich in eine bestimmte Richtung verzerren oder verfälschen können (Moosbrugger & Kelava, 2020a; Bühner, 2011).

2.3 Zielsetzung

Das übergeordnete Ziel der vorliegenden Masterthesis ist eine erste Überprüfung sowie Validierung des FBMSN nach Burdorf und Lux (2022). Die Prüfung erfolgt auf Basis der Klassischen Testtheorie.

Diese Forschungsarbeit setzt sich aus zwei unabhängige Studien zusammen. Die erste Studie erfasst den FBMSN in seiner Rohform und dient der Analyse der generierten Items. Außerdem wird die Kategorisierung der Items nach Themenblöcken von Burdorf und Lux (2022) als Basis für theoretisch angenommene Struktur des Fragebogens beibehalten. Entsprechend bilden diese Themenblöcke, die vermuteten Subskalen des FBMSN und dienen als erster Orientierungspunkt für die Prüfung der faktoriellen Struktur. Auf Basis der Ergebnisse der ersten Studie soll eine Revision und gegebenenfalls eine Umstrukturierung des Fragebogens erfolgen.

Die revidierte Version wird dann im Rahmen der zweiten Studie anhand einer weiteren Stichprobe erhoben. Zudem wird der Fragebogen durch passende Validierungsskalen ergänzt, um die Validität auf Basis der erhobenen Außenkriterien zu bestimmen.

Im Optimalfall liefert diese Arbeit eine Version des Fragebogens, die für weitere Forschungszwecke freigegeben und gegebenenfalls optimiert werden kann.

3 Studie I

Dieser Teil der Thesis behandelt die erste Erhebung, welche im Rahmen der Überprüfung des Instrumentes durchgeführt wurde. Zunächst sollen die testtheoretische Basis und relevante statistische Verfahren erläutert werden, um die Nachvollziehbarkeit der Vorgehensweise zu sichern. Weiterhin wird auf die methodische Herangehensweise eingegangen. Überdies werden die Ergebnisse präsentiert und auf Basis einer kritischen Diskussion der Resultate, Schlussfolgerungen für die Revision und Optimierung des Fragebogens gezogen.

3.1 Testtheoretischer Hintergrund – Studie I

Im Rahmen des nachfolgenden Abschnitts sollen die relevanten testtheoretischen Hintergründe sowie statistischen Methoden weiter ausgearbeitet und ausführlich behandelt werden, um eine stringente Prüfung und Überarbeitung des FBMSN zu gewährleisten.

3.1.1 Itemanalyse

Nach der Planungs- und Entwicklungsphase eines Testverfahrens besteht der nächste Schritt in einer deskriptivstatistischen Evaluation der generierten Items. Daher wird im Rahmen dieser ersten Studie der Fragebogen anhand einer, für die Zielgruppe möglichst repräsentativen Stichprobe erhoben. Die Itemanalyse ermöglicht eine genauere Qualitätsbeurteilung des neuen Messinstruments, um festzustellen, welche Items für die Verwendung geeignet bzw. ungeeignet sind. Dazu folgt die Analyse den folgenden Schritten und fußt auf den entsprechenden Kennwerten (Döring & Bortz, 2016; Kelava & Moosbrugger, 2020a):

- Schwierigkeitsanalyse und Bestimmung des Schwierigkeitsindex Pi
- Bestimmung der Itemvarianzen Var(x)
- Bestimmung und Analyse der Trennschärfe r_{it}

Diese sollen im Rahmen des folgenden Kapitels ausführlich behandelt und vorgestellt werden. Auf Basis der Itemanalyse erfolgt dann eine Zusammenstellung der Items zu einem Test bzw. Subtest, der das interessierende Konstrukt bestmöglich abbildet.

3.1.1.1 Schwierigkeitsanalyse und Bestimmung des Schwierigkeitsindex Pi

Die Schwierigkeit eines Items *i* wird durch den Schwierigkeitsindex P_i beschrieben. Dieser ist laut Kelava und Moosbrugger (2020a) definiert als "der Quotient aus der bei diesem Item tatsächlich erreichten Punktsumme aller n Testpersonen und der maximal erreichbaren Punktsumme aller n Testpersonen (d. h. die Summe der höchstmöglichen Zustimmungsstufen, multipliziert mit dem Faktor 100)" (Moosbrugger & Kelava, 2020, S.146). Auf Basis dieser Definition ergibt sich für die Berechnung des *Schwierigkeitsindex P_i* die folgende Formel (Pospeschill, 2022; Kelava & Moosbrugger, 2020a):

$$P_{i} = \frac{\sum_{v=1}^{n} x_{vi}}{n \cdot (k-1)} \cdot 100$$

Wobei x_{vi} den Itemwerten und k der Anzahl der Antwortstufen der entsprechenden Ratingskala entsprechen (Pospeschill, 2022).

Ein Item ist somit umso schwieriger, je niedriger seine Lösungs- bzw. Zustimmungsrate (für symptomatisches Verhalten) in der vorliegenden Stichprobe ausfällt. Entsprechend wird ein besonders leichtes Item von fast allen Probanden korrekt bzw. zustimmend beantwortet (Döring & Bortz, 2016). In der Regel werden Items so konstruiert, dass eine Zustimmung als Indikator für eine höhere Merkmalsausprägung gilt. Allerdings werden häufig auch *invertierte Items* verwendet, um z. B. den Einfluss von Antworttendenzen abzuschwächen. Diese sind wiederum so konstruiert, dass eine Zustimmung als symptomatisch für eine geringe Merkmalsausprägung zu verstehen ist. Entsprechend müssen

die Antworten für diese Items vor der Itemanalyse "umgepolt" werden. Folglich wird beispielsweise eine vierstufige Kodierung 0, 1, 2, 3 in 3, 2, 1, 0 reinvertiert (Kelava & Moosbrugger, 2020a).

Um einen besonders hohen Informationsgehalt der Items zu gewährleisten, empfiehlt es sich eine möglichst breite Schwierigkeitsstreuung für den Test zu realisieren. Entsprechend sollte auf extrem leichte (bzw. schwierige) Items verzichtet werden, da diese keine Differenzierung zwischen den einzelnen Testpersonen erlauben. Grundsätzlich bedeuten mittlere Werte für P_i die größtmögliche Streuung der Itemantworten. Entsprechend erlauben Items mittlerer Schwierigkeit eine hohe Differenzierung zwischen Testpersonen mit hoher bzw. niedriger Merkmalsausprägung. Allgemein betrachtet werden zumeist Items mit einem Schwierigkeitsindex zwischen $P_i = 20$ und $P_i = 80$ präferiert. Items mit höheren ($P_i \ge 80$) bzw. niedrigeren ($P_i \le 20$) Werten erlauben eine genauere Differenzierung für das gesamte Merkmalsspektrum hinweg (Prospechill, 2022; Kelava & Moosbrugger, 2020a; Döring & Bortz, 2016).

3.1.1.2 Bestimmung der Itemvarianzen *Var(x)*

Die *Itemvarianz Var(x)* bietet einen Kennwert für die Beurteilung der *Differenzierungsfähigkeit* eines Items *i* in einer Stichprobe. Daraus resultiert ein direkter Zusammenhang zwischen Itemschwierigkeit und -varianz. So ist die mögliche Ausprägung der Differenzierungsfähigkeit (bzw. Itemvarianz) durch die Schwierigkeit des Items begrenzt. Demzufolge können Items mittlerer Schwierigkeit viele Differenzierungen (hohe Itemvarianz) liefern, wohingegen Items hoher oder niedriger Schwierigkeit weniger Differenzierungen (geringe Itemvarianz) ermöglichen. Ungeeignet für eine Differenzierung sind Items, die eine extreme Schwierigkeit (extrem leicht oder schwer) aufweisen (Pospeschill, 2022; Kelava & Moosbrugger, 2020a).

Zur Berechnung der *Differenzierungsfähigkeit (oder Itemvarianz) Var(x_i)* gelten die folgenden Formeln (Pospeschill, 2022):

$$Var(x_i) = \frac{\sum_{v=1}^{n} (x_{vi} - \overline{x}_i)^2}{n}$$
 oder $Var(x_i) = \frac{\sum_{v=1}^{n} (x_{vi} - p_i \cdot (k-1))^2}{n}$

Wobei $p_i = P_i/100$ und k der Anzahl der Antwortstufen des Items i entspricht (Pospeschill, 2022).

Für die Zusammenstellung eines möglichst repräsentativen Itemsatzes für den untersuchten Merkmalsbereich sind zunächst die Ergebnisse der deskriptivstatistischen Analyse der Items im Hinblick auf Itemschwierigkeit und Differenzierungsfähigkeit (bzw. Itemvarianz) zu betrachten. Wie bereits in den vorangegangenen Abschnitten beschreiben wurde, sind liefern Items mittlerer Schwierigkeit (P_i = 50) die meisten Informationen und erlauben eine deutliche Differenzierung zwischen Probanden höherer und niedriger Merkmalsausprägung. Weiterhin sollte ein Test eine Differenzierung über das gesamte

Merkmalsspektrum hinweg ermöglichen. Entsprechend empfiehlt es sich, Items aller Schwierigkeitsgrade in den Itempool aufzunehmen. Dabei gilt für die Zusammensetzung die Faustformel: 50 % mittelschwere, 25 % schwere und 25 % leichte Items (Kelava & Moosbrugger, 2020a; Döring & Bortz, 2016).

3.1.1.3 Bestimmung der Trennschärfe r_{it}

Die *Trennschärfe bzw. der Trennschärfekoeffizient oder Trennschärfeindex* r_{it} eines Items i kann als Item-Test-Korrelation beschrieben werden. Entsprechend gibt diese Aufschluss über das Ausmaß des korrelativen Zusammenhangs zwischen einem einzelnen Itemwert x_{vi} einer Testperson v und dem Gesamttestwert x_v . Folglich wird zur Bestimmung der Trennschärfe die Korrelation zwischen einem Itemwert und einem Testwert, der sich aus den anderen Items der Skala ergibt, berechnet. Die Trennschärfe drückt somit aus, ob und wie gut die Merkmalsdifferenzierung durch das Items i mit der Differenzierung, durch den Gesamttest, übereinstimmt (Pospeschill, 2022; Kelava & Moosbrugger, 2020a). Die Trennschärfekoeffizienten bewegen sich in einem Wertebereich von $-1 \le r_{it} \le +1$. Außerdem sind möglichst hohe Trennschärfen bei eindimensionalen Tests zu erwarten (Döring & Bortz, 2016).

Zur Berechnung des *Trennschärfekoeffizienten* r_{it} wird für alle Testpersonen n die Korrelation eines Items i mit dem Testwert x_v (Pospeschill, 2022):

$$r_{it} = r(x_{vi}, x_{v})$$

Der Testwert x_v stellt dabei zumeist einen vorläufigen (d. h. testtheoretisch ungesicherten) Wert dar. Der vorläufige Testwert x_v einer Testperson v lässt sich am einfachsten aus dem Summenwert aller Itemwerte bestimmen (Pospeschill, 2022; Kelava & Moosbrugger, 2020a):

$$x_{v} = \sum_{i=1}^{m} x_{vi}$$

Wenn der Test aus wenigen Items besteht, sollte der Testwert als Summenwert ohne das betreffende Item i gebildet werden, um die Trennschärfe nicht zu überschätzen. Die auf Basis des korrigierten Testwerts bestimmte Trennschärfe wird als *part-whole-korrigierte Trennschärfe* $r_{it(i)}$ bezeichnet. Entsprechend ermittelt sich dann (Pospeschill, 2022; Kelava & Moosbrugger, 2020a):

$$r_{it(i)} = r(x_{vi}, x_{v(i)})$$

Auf Basis der Part-Whole-Korrektur (Pospeschill, 2022):

$$x_{v} = \left(\sum_{i=1}^{m} x_{vi}\right) - x_{vi}$$

Um eine möglichst genaue Erfassung des interessierenden Merkmals durch den Fragebogen oder Test zu gewährleisten, müssen die Items eine ausreichend hohe Trennschärfe aufweisen. Diese kann durch eine hohe Itemvarianz begünstigt werden. Grundsätzlich sind Items mit einer negativen oder einer Trennschärfe nahe null zu eliminieren (Pospeschill, 2022; Kelava & Moosbrugger, 2020a). Weiterhin sind bei eindimensionalen Testverfahren hohe Trennschärfen erstrebenswert, da eine hohe Trennschärfe dahin gehend interpretiert werden kann, dass die jeweiligen Items dasselbe eindimensionale Konstrukt messen. Items mit niedrigen Trennschärfen sind bei einem engen Konstrukt (bzw. homogenen Test) als schlechte Indikatoren zu bewerten. Entsprechend sind diese aus einem eindimensional angelegten Test zu eliminieren, um die Testhomogenität zu gewährleisten (Döring & Bortz, 2016).

Für die Interpretation des Trennschärfekoeffizienten und entsprechenden Schlussfolgerungen postulieren Kelava und Moosbrugger (2020a) bestimmte *Fallunterscheidungen*, da der Item-Trennschärfekoeffizient in Bezug auf die Höhe sowie die Art des Vorzeichens verschiedene Ausprägungen ergeben kann. Im Folgenden sind die drei *Fallunterscheidungen* kurz beschrieben (Kelava & Moosbrugger, 2020a):

- 1. Trennschärfe hochpositiv (bzw. nahe dem Wert +1): Ein hoher positiver Wert kann dahin gehend interpretiert werden, dass die einzelnen Items homogen gegenüber dem Gesamttest messen. Entsprechend leisten Items mit einer hohen Trennschärfe einen positiven Beitrag zu Messgenauigkeit und Validität des Tests. Inhaltlich betrachtet, wird das Item i von einem Probanden mit einem hohen Testwert (hohe Merkmalsausprägung) gelöst, während es von einer Testperson mit geringem Testwert (niedrige Merkmalsausprägung) nicht gelöst wird. Hier gelten Werte zwischen .4 und .7 als "gute" Trennschärfen (Pospeschill, 2022; Kelava & Moosbrugger, 2020a).
- 2. *Trennschärfe nahe null:* Ein niedriger Wert weist darauf hin, dass kein Zusammenhang zwischen der Differenzierung des Gesamttests und der Differenzierung durch das Item *i* besteht. Entsprechend ist das Item ungeeignet, um zwischen Probanden mit hoher und niedriger Merkmalsausprägung zu differenzieren (Pospeschill, 2022; Kelava & Moosbrugger, 2020a).
- 3. Trennschärfe hochnegativ (bzw. nahe dem Wert -1): Negative Werte deuten oft auf invertierte oder fehlerhaft formulierte Items hin. Entsprechend kann dieser Wert durch eine Änderung der Kodierungsrichtung korrigiert werden, sofern sich dies inhaltlich-theoretisch rechtfertigen lässt. Formal betrachtet zeigt eine negative Trennschärfe an, dass Testpersonen mit niedriger Merkmalsausprägung symptomatisch antworten und Probanden mit hoher Merkmalsausprägung nicht (Pospeschill, 2022; Kelava & Moosbrugger, 2020a).

3.1.2 Reliabilität und interne Konsistenz

Wie bereits in den Kapitelabschnitten 2.2.2.4 Bestimmung der Reliabilität und 2.2.5.2 Reliabilität sowohl inhaltlich als auch testtheoretisch ausführlich beschrieben wurde, dient die Reliabilität der Einschätzung der Genauigkeit einer Messung. Daher ist es für eine erste Analyse der Itemkennwerte unabdingbar, die Messgenauigkeit des Tests zu bestimmen. Dazu erfolgt eine Einschätzung der Reliabilität auf Basis der Konsistenzmethode mithilfe des Cronbachs Alpha-Koeffizienten. Im Wesentlichen stellt die Bestimmung der inneren Konsistenz eine Erweiterung der Testhalbierungsmethode dar (siehe 2.2.5.2 Reliabilität). Inhaltlich betrachtet erfolgt eine Überprüfung des inneren Itemzusammenhangs unter Berücksichtigung der Testlänge. Diese erfolgt auf Basis von Itemvarianzen sowie Itemkovarianzen. Formal betrachtet entspricht der Alpha-Koeffizient der mittleren Split-Half-Reliabilität des Tests für alle möglichen Testhalbierungen. Dieser ist wie folgt definiert (Döring & Bortz, 2016; Bühner, 2011):

$$Rel_{interne \ Konsistenz} = \alpha = \frac{p}{p-1} \cdot \left(1 - \frac{\sum_{i=1}^{p} Var_{x_i}}{Var_x}\right)$$

Entsprechend fällt der Alpha-Koeffizient umso höher aus, je mehr Items ein Test umfasst $(p = \text{Itemanzahl}; x_i: \text{Itemwerte von 1 bis } p; x: \text{Testwert als Summe aller Itemwerte})$ und je größer die Item-Interkorrelationen sind (Döring & Bortz, 2016).

Voraussetzung für eine Konsistenzanalyse auf Basis des Alpha-Koeffizienten ist, dass Eindimensionalität vorliegt, d. h. dass alle Items eines Tests bzw. Subtests dasselbe Merkmal in ähnlicher Art und Weise messen. Außerdem müssen die Items zumindest essenzielle τ-Äquivalenz aufweisen. Diese liegt vor, wenn die Items denselben Anteil an wahrer Varianz aufweisen sind. In der Praxis sind diese Voraussetzungen häufig nicht erfüllt. Daher liefert der Koeffizient lediglich eine Mindestschätzung der Reliabilität, was wiederum im Zuge der Interpretation zu berücksichtigen ist (Gäde, Schermelleh-Engel, & Werner, 2020; Bühner, 2011). Nach Bühner (2011) sind Reliabilitätskoeffizienten unter .60 inakzeptabel. Werte über .70 und .80 gelten wiederum als gut bzw. sehr gut. Werte über .90 deuten meist auf redundante Items hin. Weiterhin sind auch die Reliabilitätskoeffizienten solcher Skalen genauer zu betrachten, welche durch eine besonders hohe Anzahl von Items repräsentiert sind. Schließlich steigt die Reliabilität eines Tests generell durch die Verwendung zusätzlicher Items (Pospeschill, 2022).

3.1.3 Faktorenanalyse

Die Faktorenanalyse fasst eine Gruppe multivariabler Analyseverfahren, welche sich im Hinblick auf ihr konkretes Ziel differenzieren lassen, zusammen. Entsprechend wird zwischen der Exploratorischen Faktorenanalyse (EFA) und der Konfirmatorischen Faktorenanalyse (CFA) unterschieden. Diese dienen zum einen als struktursuchende

Verfahren (EFA) der Generierung von Hypothesen, sowie zum anderen als strukturprüfende Verfahren (CFA) der Prüfung von Hypothesen, über die faktorielle Struktur (Pospeschill, 2022; Bühner, 2011). Im Allgemeinen ist das Ziel einer Faktorenanalyse die Aufklärung der Korrelationen der Items untereinander durch eine geringe Anzahl homogener Faktoren. Diese Faktoren repräsentieren zumeist verborgene latente Variablen, welche die beobachteten Variablen beeinflussen. Im Rahmen der exploratorischen Faktorenanalyse wird die Item-Korrelationsmatrix analysiert (Bühner, 2011). An dieser Stelle liegt der Fokus hauptsächlich auf der exploratorischen Faktorenanalyse, da es sich in *Studie I* um eine erste Prüfung der faktoriellen Struktur des Messinstruments handelt und keine konkreten Hypothesen über diese Struktur vorliegen.

Vor der Durchführung der exploratorische Faktorenanalyse sind zunächst drei bestimmte Entscheidungen im Hinblick auf die Festlegung der Methoden und Kriterien der Faktorenanalyse zu treffen. Die Auswahl richtet sich dabei nach den verfolgten Zielen (Pospeschill, 2022; Bühner, 2011): Festlegung der Extraktionsmethode sowie Bestimmung von Eigenwert und Kommunalitäten, Festlegung des Abbruchkriteriums und Festlegung der Rotationsmethode.

Im Folgenden werden diese Komponenten der Faktorenanalyse ausführlich behandelt. Zunächst soll allerdings auf Grundannahmen der Faktorenanalyse eingegangen. Anschließend wird auf verschiedene Methoden und Kriterien der EFA eingegangen.

3.1.3.1 Grundannahmen und Kennwerte der Faktorenanalyse

Die Faktorenanalyse basiert auf der Annahme, dass der Testwert eines Probanden auf einer bestimmten Variable durch gewichtete Ausprägungen der Testperson auf den Faktoren (Faktorwerten) und einen Fehler bestimmt werden kann. Entsprechend dem *Fundamentaltheorem* der Faktorenanalyse kann eine *beobachtete standardisierte Messung zvi* einer Testperson v bei einem Item i in einer Linearkombination dargestellt werden. Diese umfasst die Faktorladungen λ_{ik} gewichteter Faktorenwerte f_{kv} sowie einer Fehlerkomponente ε_{vi} (Pospeschill, 2022; Moosbrugger & Harting, 2003):

$$z_{vi} = \lambda_{i1}f_{1v} + \lambda_{i2}f_{2v} + \dots + \lambda_{ik}f_{kv} + \dots + \lambda_{iq}f_{qv} + \varepsilon_{vi} = \sum_{k=1}^{q} (\lambda_{ik}f_{kv}) + \varepsilon_{vi}$$

Weiterhin kann eine Linearkombination für die standardisierte Variable z_i eines Items i aus den Faktorladungen λ_{ik} gewichteter Faktoren F_k und einer Fehlerkomponente ε_{vi} bestimmt werden (Pospeschill, 2022; Bühner, 2011):

$$z_i = \lambda_{i1} F_1 + \lambda_{i2} F_2 + \dots + \lambda_{ik} F_k + \dots + \lambda_{iq} F_q + \varepsilon_i = \sum_{k=1}^{q} (\lambda_{ik} F_k) + \varepsilon_i$$

Die Gewichtungskoeffizienten λ_{ik} der Variablen z_i auf den Faktoren F_k bestimmen die Faktorladungen. Bei einer orthogonaler Rotation, d. h. unter der Annahme, dass die

Faktoren unkorreliert sind, können die Faktorenladungen Werte zwischen -1 und +1 annehmen. Zudem können die Faktorladungen als Korrelationskoeffizienten für den Zusammenhang zwischen den beobachteten Variablen sowie dem Faktor verstanden und interpretiert werden. Überdies ist die Varianz der standardisierten Variablen $Var(z_i)$ als Summe aus den quadrierten Faktorladungen λ_{ik}^2 sowie der Varianz der Fehlerkomponenten $Var(\varepsilon_i)$ beschrieben (Pospeschill, 2022: Brandt, 2020):

$$Var(z_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \dots + \lambda_{ik}^2 + \dots + \lambda_{iq}^2 + Var(\varepsilon_i) = 1$$

Wobei der Index q die Faktorenanzahl bestimmt. Durch die Standardisierung ist die $Var(\varepsilon i) = 1$ (Pospeschill, 2022: Brandt, 2020).

Bei einer obliquen Rotation, d. h. unter der Annahme, dass die Faktoren korrelieren, werden die Korrelationen von Faktor und Item in einer *Strukturmatrix* sowie partielle⁶ standardisierte Regressionsgewichte in einer *Mustermatrix* abgebildet (Pospeschill, 2022: Brandt, 2020).

Der *Eigenwert* bezeichnet die jeweils durch einen Faktor aufgeklärte Varianz, die sich über alle standardisierten Variablen ergibt. Dieser ist bestimmt durch die Summe der quadrierten Faktorladungen des entsprechenden Faktors λ_{ik}^2 aller Variablen z_i (Pospeschill, 2022):

$$Eig\left(F_{k}\right) = \sum_{i=1}^{m} \lambda_{ik}^{2}$$

Des Weiteren kann für jede Variable z_i genau ermittelt werden, wie groß der Anteil mit wie vielen extrahierten Faktoren an der Varianzaufklärung ist. Dazu erfolgt die Bestimmung der *Kommunalität* h_i^2 , durch die Summierung der quadrierten Ladungen über sämtliche Faktoren:

$$h_i^2 = \sum_{k=1}^q \lambda_{ik}^2$$

Der Faktorwert gibt Aufschluss über den Ausprägungsgrad einer Testperson auf einem bestimmten Faktor. Für die Interpretation ist festzuhalten, dass hohe positive Werte auf eine hohe Ausprägung auf dem entsprechenden Faktor hindeuten, während hohe negative Werte für eine niedrige Ausprägung sprechen (Pospeschill, 2022; Brandt, 2020).

3.1.3.2 Extraktionsmethoden

Im folgenden Abschnitt sollen die drei etablierten und am häufigsten verwendeten Methoden, Hauptkomponenten- sowie Hauptachsenanalyse und Maximum-Likelihood-Faktorenanalyse kurz beschrieben werden. An dieser Stelle ist darauf hinzuweisen, dass es

⁶ Partiell: Für jedes Item werden die Zusammenhänge zu anderen Faktoren auspartialisiert bzw. eliminiert (Pospeschill, 2022).

neben diesen drei Varianten noch weitere Methoden gibt, welche allerdings nicht weiter thematisiert werden.

Die Hauptkomponentenanalyse (PCA) zielt auf eine Dimensions- bzw. Datenreduktion ab. Weiterhin soll durch Extraktion von Faktoren möglichst viel Varianz eines Datensatzes für die beobachteten Variablen aufgeklärt werden. Dabei geht die PCA von der Annahme aus, dass sich 100% der Varianz durch die Faktoren erklären lässt. Entsprechend bleibt die Messfehlervarianz der beobachteten Variablen unberücksichtigt. Die PCA standardisiert die Varianz jeder Variablen auf einen Wert von Eins. Entsprechend bildet die Gesamtvarianz die Anzahl m der beobachteten Variablen (Testitems). Praktisch bündeln die extrahierten Hauptkomponenten mehr Varianz als eine einzelne Variable. Dennoch verbleibt ein Fehlerterm aus den Anteilen der nicht berücksichtigten Hauptkomponenten (Pospeschill, 2022; Bühner, 2011).

Ziel der *Hauptachsenanalyse (PFA)* ist die Identifikation von Faktoren, die eine Aufklärung der Beziehungsstruktur der beobachteten Variablen zulässt. Im Gegensatz zur PCA geht die PFA davon aus, dass die Gesamtvarianz der beobachteten Variablen sich aus wahrer Varianz und Fehlervarianz zusammensetzt. Die Faktoren werden somit anhand der wahren Varianzanteile bestimmt. Bei der Analyse wird die Korrelationsmatrix der quadrierten multiplen Korrelationen jeder einzelnen Variable mit dem restlichen Variablensatz betrachtet (Pospeschill, 2022; Bühner, 2011).

Die *Maximum-Likelihood-Faktorenanalyse (MLF)* dient der Schätzung von Populationsparametern aus Stichprobenstatistiken. Dabei wird die Korrelationsmatrix der Stichprobe als Grundlage für die Schätzung betrachtet. Anschließend werden die Faktoren mithilfe des χ^2 – Test (Goodness-of-Fit-Test) extrahiert. Dieser wird so oft wiederholt, bis er nicht mehr signifikant wird. Entsprechend werden schrittweise Faktoren extrahiert, die zur Aufklärung der Varianz beitragen. Die Ergebnisse einer MLF können explizite Rückschlüsse auf Zusammenhänge in der Population liefern. Allerdings nur dann, wenn das Modell durch Modelltests hinsichtlich seiner Anpassungsgüte erfolgreich geprüft wurde. Weiterhin erfordert die MLF eine sehr große Stichprobe sowie multivariat-normalverteilter Variablen (Pospeschill, 2022; Bühner, 2011).

3.1.3.3 Extraktionskriterien

Zur Feststellung der relevanten Faktorenanzahl können verschiedene Kriterien herangezogen werden, wenn kein explizites hypothetisches Modell vorliegt. Diese Kriterien beziehen sich auf den Eigenwertverlauf (Pospeschill, 2022). Im Folgenden werden das Kaiser-Guttman-Kriterium, der Scree-Test nach Cattel und der Minimum-Average-Partial-Test von Velicer (MAP-Test) kurz beschrieben. Eine weitere Möglichkeit zur Bestimmung der zu extrahierenden Faktorenanzahl ist die Parallelanalyse nach Horn. Diese wird nicht weiter behandelt, da sie im Rahmen dieser Studie ungeeignet für die

Bestimmung der Faktoren ist. Das *Kaiser-Guttman-Kriterium* fußt auf der Annahme, dass alle Faktoren relevant sind, die eine größere Menge an Varianz aufklären als eine einzelne Variable. Entsprechend muss der Faktor einen Eigenwert aufweisen, der größer als Eins ist. Bei einer größeren Anzahl an beobachteten Variablen kommt es häufig vor, dass eine zu stark differenzierende Aufgliederung in Faktoren erfolgt. Das hat zur Folge, dass die Anzahl an relevanten Faktoren überschätzt wird (Pospeschill, 2022; Bühner, 2011).

Weiterhin kann der *Scree-Plot* im Rahmen des *Scree-Tests* nach Cattell als zusätzliches Kriterium herangezogen werden. Dieser bildet den Eigenverlauf vor der Rotation grafisch ab. Zur Interpretation wird auf dem Graphen nach einem deutlichen Eigenwertabfall (d. h. einem deutlichen "Knick" in der Linie) gesucht. Zur Veranschaulichung werden in Abbildung 4 zwei Beispiele für einen Eigenwertverlauf dargestellt. Gemäß dem Scree-Test werden alle Faktoren als relevant betrachtet, die vor diesem Eigenwertabfall liegen (siehe Abb. 4a). In der Praxis hat sich dieses Kriterium bewährt. Trotzdem besteht ein klarer Nachteil darin, dass es häufig zu Situationen kommt, in denen der Eigenwertverlauf keinen eindeutigen "Knick" aufweist (siehe Abb. 4b). Dies hat zur Folge, dass die Wahl der Faktorenanzahl subjektiv variieren kann (Pospeschill, 2022; Brandt, 2020).

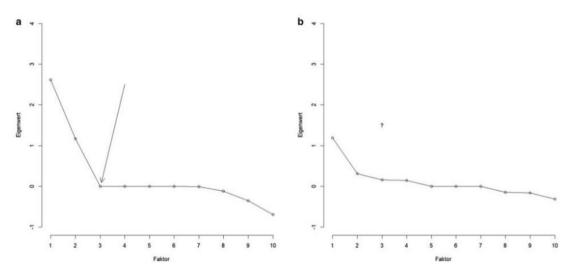


Abb. 4: a: Beispiel für eine eindeutige Entscheidung durch Scree-Test. b: Beispiel ohne deutlichen "Knick" (Brandt, 2020).

Der Minimum-Average-Partial-Test von Velicer (MAP-Test) beruht auf den Partialkorrelationen der jeweils extrahierten Komponenten (Hauptkomponentenanalyse) aus der Korrelationsmatrix der beobachteten Variablen. Zur Berechnung des MAP-Tests mit SPSS liefert O'Connor (2000) eine Syntax (Pospeschill, 2022; Bühner, 2011).

Wie bereits erwähnt wird die *Parallelanalyse nach Horn* nicht weiterbehandelt, da von einer hohen Korrelation der Faktoren ausgegangen werden muss und somit die Genauigkeit dieser Methode schwindet (Pospeschill, 2022).

3.1.3.4 Rotationsmethoden

Rotationsmethoden dienen der Erreichung eines interpretierbaren Faktorenmusters, das einer *Einfachstruktur* entspricht. Diese liegt vor, wenn alle beobachteten Variablen auf einem bestimmten Faktor eine hohe (Primär-)Ladung sowie auf den anderen Faktoren keine bzw. eine niedrige (Sekundär-)Ladung aufweisen. Grundsätzlich wird zwischen *orthogonalen* und *obliquen* Rotationsmethoden differenziert. Orthogonale Verfahren, gehen von unkorrelierten Faktoren und somit einer unabhängigen Interpretation der Faktoren aus. Diese Annahme wird hingegen bei obliquen Verfahren, zugunsten einer besseren Interpretierbarkeit aufgegeben (Pospeschill, 2022; Brandt, 2020). Im Rahmen dieses Kapitels liegt der Fokus auf der obliquen Rotation, da in dieser Studie nicht von unkorrelierten Faktoren ausgegangen werden kann.

Eine oblique Rotation der Faktoren verfolgt das Ziel, eine klare Einfachstruktur im Hinblick auf die semipartiellen standardisierten Regressionsgewichte der beobachteten Variablen auf den Faktoren zu erreichen. Diese werden in der Mustermatrix abgebildet. Weiterhin liefert ein obliques Rotationverfahren eine Strukturmatrix, welche die Korrelationen zwischen den Faktoren und den Items darstellt. Zur Interpretation wird in erster Linie die Mustermatrix betrachtet (Mulaik, 2010). Nichtsdestotrotz ist es bei korrelierten Faktoren wichtig darauf zu verweisen, dass, selbst wenn in der Mustermatrix keine "Nebenladungen" für die beobachteten Variablen vorliegen, sämtliche Items mit den anderen Faktoren substanziell korrelieren (Bühner, 2011). Laut Bühner (2011) sind die Direct-Oblimin- und die Promax-Rotation die beiden wichtigsten Verfahren obliquer Methoden der Rotation. Eine Oblimin-Rotation, verfolgt das Ziel, eine simultane Optimierung bezüglich eines orthogonalen und obliquen Rotationskriteriums zu erreichen. Durch die Festlegung einer bestimmten Winkelgröße kann der Grad der Faktoreninterkorrelation spezifiziert werden. Diese Methode wird am häufigsten verwendet, wenn ein konkretes theoretisches Modell über die Korrelation der Faktoren gegeben ist. Die Promax-Methode ist eine schiefe Rotation, bei der die Faktoren miteinander korrelieren dürfen. Diese Rotation kann schneller berechnet werden als eine Oblimin-Rotation (Brandt, 2020; Bühner; 2011).

3.1.3.5 Voraussetzungen für die Durchführung einer Faktorenanalyse

Bevor eine Faktorenanalyse letztlich erfolgreich durchgeführt werden kann, ist von verschiedene Voraussetzungen bezüglich der verwendeten Items und der Stichprobe abhängig. Diese werden im Folgenden thematisiert. Dabei wird ebenfalls auf die entsprechenden Kennwerte sowie Prüfungsmethoden eingegangen.

Zunächst weist Bühner (2011) darauf hin, dass eine Faktorenanalyse nur bei einer ausreichend hohen Item-Korrelationen durchgeführt werden kann. Ob substanzielle Korrelationen in der beobachteten Korrelationsmatrix vorliegen, wird anhand des Kaiser-

Meyer-Olkin-Koeffizienten (KMO) quantifiziert. Dabei sollte der Koeffizient einen Wert von mindestens .60 erreichen. Weiterhin wird der Bartlett-Test auf Sphärizität verwendet, um zu prüfen, ob die Korrelationen signifikant von null abweichen (Pospeschill, 2022; Bühner, 2011).

Des Weiteren bildet ein ausreichender Itemsatz die Basis eine erfolgreiche Durchführung der Faktorenanalyse. Dabei gelten vier Items pro Faktor als Untergrenze und es wird definitiv zur Verwendung von mehr Items geraten (Bühner, 2011). Überdies muss eine angemessen große Stichprobe vorliegen. Nach Bühner (2011) ist ein Stichprobeumfang von n = 100 bis 200 als "ausreichend" bzw. "befriedigend". Jedoch konnten Mac-Callum, Widaman, Zhang und Hong (1999) zeigen, dass eine Stichprobe von 60 Personen ausreichend sein kann. Allerdings nur unter der Bedingung, dass jedes Item eine Kommunalität von mindestens $h^2 = .60$ aufweisen muss. Für die Beurteilung der Qualität der Items sollte die Kommunalität grundsätzlich mitberücksichtigt werden. Allgemein lässt sich festhalten, dass eine Faktorlösung umso stabiler ist, je höher der Stichprobenumfang sowie die Kommunalitäten ausfallen (Pospeschill, 2022; Bühner, 2011). Schließlich ist die eine gleichmäßige Verteilung der Items anzustreben, da sich diese auf die Höhe der Korrelation auswirkt. Außerdem können Extrem- oder Ausreißerwerte eine Verzerrung der Korrelationen hervorrufen. Daher sollten die Streudiagramme der Items ebenfalls betrachtet werden, auch wenn ein Antwortformat mit mehr als zwei Stufen verwendet wird (Pospeschill, 2022; Bühner, 2011).

3.1.4 Zielsetzungen

Der Fragebogen von Burdorf und Lux (2022) soll verschiedene Motive der Selbstdarstellung auf sozialen Netzwerken abbilden. Das Ziel der ersten Studie ist es, eine erste Itemund Reliabilitätsanalyse durchzuführen. Weiterhin soll eine Revision der Erstversion des Fragebogens erfolgen, um eine Version des Fragebogens zu entwickeln, die den testtheoretischen Anforderungen genügt. Im Zuge dessen soll die faktorielle Struktur des Fragenbogens analysiert werden, um verschiedene Motive der Selbstdarstellung auf sozialen Netzwerken anhand einzelner Subskalen abzubilden. Im Zuge dessen sollen folgende Aspekte empirisch geprüft werden:

- Die Items sollen auf Basis geeigneter Kriterien analysiert werden. Im Zuge dessen sollen redundante sowie ungeeignete Items zur Kürzung des Fragebogens eliminiert werden.
- Die Reliabilität soll für die beobachteten Variablen bestimmt werden, um eine Einschätzung über die Messgenauigkeit der Skalen zu gewinnen.
- Die angenommene Struktur des Fragebogens soll analysiert werden. Entsprechend soll die faktorielle Struktur des Fragebogens empirisch geprüft werden.

Die methodische Vorgehensweise dieser Studie ist weitestgehend explorativer Natur. Daher werden an dieser Stelle keine konkreten Hypothesen formuliert.

3.2 Methode - Studie I

Das folgende Kapitel beschäftigt sich mit der methodischen Untersuchung und empirische Überprüfung auf Basis des beschriebenen testtheoretischen Fundaments. Im Zuge dessen werden zunächst das Studiendesign, relevante Informationen zur Rekrutierung der Stichprobe sowie die Datenerhebung dargestellt und das verwendete Messinstrument beschrieben. Abschließend sollen die geplanten statistischen Analysen konkret dargestellt werden.

3.2.1 Forschungsdesign

Das Forschungsdesign bildet den Rahmen der Untersuchung und hat direkten Einfluss auf den Ablauf der Studie. Die vorliegende Studie fußt auf einem nicht-experimentellen quantitativen Untersuchungsansatz. Für eine erste Überprüfung und Optimierung des FBMSN wird eine Querschnittbefragung durchgeführt (Richter, Paier, & Reiger, 2021; Eid, Gollwitzer, & Schmitt, 2017). Der Erhebungszeitraum der Befragung erstreckte sich vom 2. August 2022 und endet fünf Wochen später, am 10. September 2022. Der Fragebogen wird in seiner Rohform über eine Online-Umfrage anhand einer Analysestichprobe vollständig erhoben.

3.2.2 Anforderungen an die Stichprobe und Rekrutierung

Die Rohversion des Fragebogens wird anhand einer Gelegenheitsstichprobe einmalig in vollem Umfang erhoben. Eine Gelegenheitsstichprobe untersucht die Personen, die freiwillig an der Untersuchung teilnehmen und auf einfachem Wege erreicht werden können (Renner, Heydasch, & Ströhlein, 2012). Der wesentliche Unterschied zur Zufallsstichprobe besteht darin, dass nicht alle theoretisch möglichen Stichproben der Größe *n* die gleiche Wahrscheinlichkeit aufweisen, aus einer bestimmten Population gezogen zu werden (Eid, Gollwitzer, & Schmitt, 2017).

Laut Bühner (2011) gilt eine Stichprobe der Größe n=200 als "befriedigend" und eine Stichprobengröße von n=300 als "gut" geeignet für die Durchführung einer Faktorenanalyse. Daher wird eine Stichprobengröße von mindestens 200 Personen angepeilt. In erster Linie sollen nur Personen für die Stichprobe rekrutiert werden, die zum Zeitpunkt der Erhebung über mindestens einen Social Media Account verfügen und diesen regelmäßig nutzen. Des Weiteren sollte die Stichprobe eine ausgeglichene Geschlechterverteilung aufweisen und auf eine Altersspanne von 18 bis 40 Jahren begrenzt sein. Zur Rekrutierung der Probanden bietet sich eine Kontaktaufnahme über verschiedene Social Media Kanäle sowie E-Mail-Kontakte an, um eine große Anzahl potenzieller

Probanden schnell und kostenfrei anzusprechen. Ansonsten wird die Umfrage im digitalen Versuchspersonenstundensystem *Sona* der Hochschule Fresenius eingebunden, um weitere Teilnehmer zu akquirieren, diese können sich für ihre Teilnahme 0,25 Versuchspersonenstunden anrechnen lassen. Überdies wird der Fragebogen auf den Plattformen *SurveyCircle* und *Thesius* geteilt, um die Studie über möglichst viele verschiedene Kommunikationskanäle zu verbreiten. Außerdem werden als zusätzlicher Anreiz, drei Amazon-Gutscheinkarten unter den Teilnehmenden verlost.

3.2.3 Messinstrument

In der ersten Studie wurde die Rohversion des Fragebogens zur Erfassung von zugrundeliegenden Motiven der Selbstdarstellung auf sozialen Netzwerken (FBMSN) nach Burdorf und Lux (2022) verwendet.

Bei der Konstruktion des Fragebogens wurden ausschließlich geschlossene Fragen verwendet. Weiterhin stützen sich die Fragen- und Antwortformulierungen auf verschiedene Ich-Aussagen, die eine Meinung oder Einstellung abbilden. Die Probanden werden instruiert, den Grad der eigenen Zustimmung auf einer sechsstufigen Likert-Skala mit verbalen Ankern (*Stimme überhaupt nicht zu, Stimme nicht zu, Stimme eher nicht zu, Stimme eher zu, Stimme zu, Stimme völlig zu*) anzugeben. Eine Einstufung erfolgt von 1 (Stimme überhaupt nicht zu) bis 6 (Stimme völlig zu) (Burdorf & Lux, 2022).

Die Skalen werden als Themenblöcke nacheinander abgefragt, wobei auf eine konkrete Benennung der Themenblöcke verzichtet wird, um Antwortverzerrungen im Sinne der sozialen Erwünschtheit vorzubeugen (Burdorf & Lux, 2022).

Für die Skala *Modelllernen* formulieren Burdorf und Lux (2022) eine konkrete Instruktion: "Bitte denken Sie an Ihren Lieblingsinfluencer:in, Star oder die von Ihnen meist verfolgte Person auf Social Media (in den folgenden Fragen als 'die Person' betitelt). In den folgenden Fragen geht es um diese Person".

Der Fragebogen besteht aus 14 Skalen bzw. Themenblöcken mit insgesamt 99 Items. Die Anzahl der Items pro Skala variiert zum Teil recht stark. So sind einzelne Skalen lediglich durch zwei oder drei Items repräsentiert. Wohingegen andere Skalen neun oder mehr Items umfassen. Die Skalen sowie die zugehörigen Items sind dem Anhang A zu entnehmen.

3.2.4 Datenerhebung

Die Beantwortung der Fragestellung beruht auf Selbstauskünften der Teilnehmer. Dazu wurde ein Fragebogen konzipiert, welcher mithilfe der Online-Plattform *Unipark* erhoben wird. Der Zugang erfolgt über einen Link, der direkt zum konzipierten Fragebogen führt. Dieser beginnt mit einer Startseite, der wesentlicher Informationen zur Teilnahme an der Untersuchung zu entnehmen sind. Dabei werden die ethischen Richtlinien der

Deutschen Gesellschaft für Psychologie (DGP) beachtet. Es wird auf das Recht hingewiesen, die Umfrage jederzeit beenden zu können. Außerdem werden die Teilnehmer darüber aufgeklärt, dass ihre Daten anonym und vertraulich behandelt sowie nur zu wissenschaftlichen Zwecken verwendet werden. Zusätzlich wird auf den E-Mailkontakt des Versuchsleiters hingewiesen, falls sich für die Probanden Fragen zu diesem Thema ergeben sollten. Im Anschluss folgt die Einverständniserklärungsabgabe zum Datenschutz. Auf der nachfolgenden Seite wird die ungefähr benötigte Zeit zur vollständigen Beantwortung des Fragebogens von zehn Minuten angegeben. Des Weiteren wird auf die Möglichkeit zur Anrechnung von Versuchspersonenstunden sowie die Teilnahme am Gewinnspiel hingewiesen. Überdies werden die Versuchspersonen dazu aufgerufen, die Fragen möglichst spontan zu beantworten, da es keine richtigen oder falschen Aussagen gibt. Im nächsten Schritt wird eine Filterfrage installiert, welche das wesentliche Einschlusskriterium (Nutzung sozialer Netzwerke) abfragt. So kann eine Verfälschung der Daten durch ungeeignete Probanden einfach verhindert werden. Auf die Filterfrage folgt die Abfrage soziodemografischer Daten (Alter, Geschlecht, Beziehungsstatus, Bildungsabschluss sowie gegenwärtiges Beschäftigungsverhältnis und Religionszugehörigkeit) und Fragen zum individuellen Nutzungsverhalten von sozialen Netzwerken (genutzte Plattformen und Nutzungszeit in Stunden pro Tag). Anschließend beginnt die Erhebung der Rohversion des FBMSN. Am Ende der Befragung kann sich der Teilnehmer für das Gewinnspiel anmelden und sich zum digitalen Versuchspersonenstundensystem Sona der Hochschule Fresenius weiterleiten lassen. Der Fragebogen ist in Anhang A hinterlegt.

3.2.5 Statistische Analysen

Zunächst erfolgt eine ausführliche Itemanalyse. Im Zuge dessen wird der Trennschärfe-koeffizient sowie die Varianz und der Schwierigkeitsindex für jedes Item berechnet. Des Weiteren wird die Dimensionalität der Items und Skalen geprüft. Anschließend werden die Items anhand dieser Kennwerte analysiert und selektiert. Anschließend werden die Reliabilitätskoeffizienten (Cronbachs-Alpha-Koeffizienten) der theoretisch gebildeten Skalen berechnet. Zur Aufklärung der faktoriellen Struktur des Fragebogens werden explorative Faktorenanalysen berechnet und ausgewertet.

Bei der Aufbereitung und Berechnung aller relevanten Daten wurde die Software *IBM SPSS Statistics* verwendet. Die Software gehört zu den marktführenden Statistik-Analyse-Programmen und ermöglicht explorative Analysen sowie Vorhersageanalysen. Weiterhin ermöglicht das Programm auch, grafische Darstellungen umzusetzen. Das Programm *IBM SPSS Statistics* deckt nahezu alle Anforderungen der deskriptiven und inferenzstatistischen Datenauswertung und Analyse ab. Im Rahmen dieser Untersuchung wurde es für alle Datenauswertungen und -analysen verwendet (IBM, o.J.).

3.3 Ergebnisse - Studie I

Im folgenden Kapitel werden die Ergebnisse der ersten Studie präsentiert. Im Zuge dessen werden die Zusammensetzung der Stichprobe, die Kennwerte der Itemanalyse sowie die Ergebnisse Reliabilitätsanalyse und der explorativen Faktorenanalysen präsentiert. Alle Ergebnisse, der im Rahmen dieser Untersuchung durchgeführten statistischen Analysen sind im Anhang B hinterlegt.

3.3.1 Zusammensetzung der Stichprobe

An der Umfrage nahmen in der fünfwöchigen Untersuchungsphase insgesamt 373 freiwillige Probanden teil. Von diesen beendeten 317 den Fragebogen vollständig. Dies entspricht einer Abbruchquote von ca. 15 %. Nach Bereinigung des Datensatzes ergeben sich 308 verwendbare und vollständig ausgefüllte Fragebögen. Neun Versuchspersonen mussten aufgrund ihrer Überschreitung der festgelegten Altersspanne von 18 bis 40 Jahren aus den Berechnungen exkludiert werden. Alle unvollständigen Fragebögen wurden nicht in die weiteren Berechnungen und deskriptiven Statistiken einbezogen. Entsprechend umfasst die betrachtete Stichprobe 308 Versuchspersonen (*N*).

Bei Betrachtung der Häufigkeitsverteilung des Geschlechts zeigt sich, dass 59 % (n = 183) der Probanden weiblichen und 41 % (n = 125) männlichen Geschlechts sind (siehe Abb. 5).

Im Durchschnitt sind die Versuchspersonen 24,67 Jahre alt, bei einer Standardabweichung von 5,12 Jahren. Die Stichprobe umfasst eine Altersspanne von 18 bis 40 und beträgt somit 22 Jahre. Die meisten Probanden (59,7

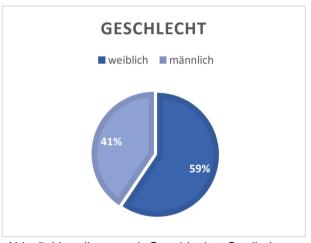


Abb. 5: Verteilung nach Geschlecht - Studie I.

%) sind zwischen 21 und 25 Jahre alt. Weiterhin befinden 61 % (n = 188) der Probanden zum Zeitpunkt der Erhebung in einer Beziehung. Insgesamt zeigt sich im Rahmen der Erhebung der soziodemografischen Daten, dass 61,7 % (n = 190) ein Abitur oder Fachabitur als höchsten Bildungsabschluss haben. Des Weiteren verfügen 25,6 % (n = 79) der Versuchspersonen über einen Bachelor- oder Masterabschluss. Ein großer Teil der Probanden sind Studenten (57,5 %; n = 157). Außerdem zeigen die Ergebnisse, dass 48,7 % der Teilnehmer sich keiner Religion zugehörig fühlen oder diese praktizieren.

Das folgende Säulendiagramm (siehe Abb. 6) dient der Veranschaulichung der Ergebnisse zum Nutzungsverhalten der Teilnehmer von Social-Media-Plattformen. Es vergleicht die einzelnen Plattformen im Hinblick auf die Nutzer unter Probanden dieser

Stichprobe. Auffällig ist, dass *Instagram* (87,7 %; n = 270) die mit Abstand (ca. 30 % mehr als *Snapchat*) am häufigsten genutzte Plattform in der vorliegenden Stichprobe ist. Weiterhin ist festzuhalten, dass die relativ junge Plattform *TikTok* in der vorliegenden Stichprobe fast genauso viele Nutzer vorweisen kann, wie die seit Langem etablierte Plattform *Facebook*. Im Hinblick auf die Nutzungsdauer geht aus den Daten hervor, dass der

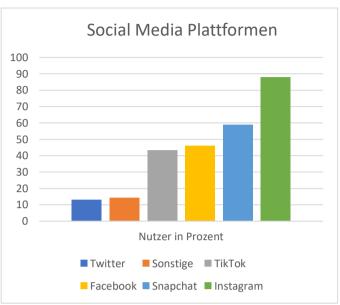


Abb. 6: Nutzung von Social Media Plattformen – Studie I.

Großteil der Untersuchungsteilnehmer (61,6 %) ein bis drei Stunden täglich auf verschiedenen Social-Media-Plattformen verbringt.

3.3.2 Itemanalyse und -selektion

Auf Grundlage der testtheoretischen Annahmen sowie den Erläuterungen zur Berechnung der Kennwerte sowie deren Interpretation, welche im Rahmen des Kapitels 3.3.1 Itemanalyse ausführlich beschrieben wurden, umfasst der folgende Abschnitt die Ergebnisse der deskriptivstatistischen Itemanalyse.

Auf Basis der ermittelten Kennwerte erfolgt die Selektion und somit der Ausschluss ungeeigneter Items. Für die Selektion liegt der Fokus primär auf den Trennschärfekoeffizienten (r_{it}) sowie dem Schwierigkeitsindex (P_i) und der Differenzierungsfähigkeit (Var(x)). Die ermittelten Kennwerte der Itemanalyse sind in der nachfolgenden Tabelle übersichtlich dargestellt.

Tab. 1: Kennwerte der Itemanalyse der Skalen des FBMSN Kennwerte der Itemanalyse

		\overline{X}	SD	Var(x)	r it	Pi
Modellernen	ML_01	3,52	1,517	2,302	0,51	50,4
	ML_02	2,15	1,267	1,606	0,68	23
	ML_03	1,71	1,157	1,338	0,54	14,2
	ML_04	1,89	1,318	1,738	0,54	17,8
	ML_05	2,39	1,388	1,926	0,65	27,8

	ML_06	2,10	1,361	1,854	0,60	22
	ML_07	2,67	1,449	2,098	0,56	33,4
	ML_08	2,54	1,447	2,093	0,42	30,8
	ML_09	3,41	1,279	1,637	0,11	48,2
	ML_10	2,41	1,332	1,774	0,56	28,2
	ML_11	3,34	1,494	2,233	0,45	46,8
		X	SD	Var(x)	r _{it}	Pi
Intimitätsmotiv	IM_01	1,78	1,165	1,356	0,52	15,6
	IM_02	1,72	1,133	1,285	0,51	14,4
	IM_03	2,81	1,546	2,390	0,48	36,2
	IM_04	3,39	1,996	3,984	0,63	47,8
	IM_05	2,72	1,867	3,484	0,62	34,4
	IM_06	2,12	1,409	1,985	0,64	22,4
	IM_07	1,79	1,152	1,327	0,62	15,8
		X	SD	Var(x)	r _{it}	Pi
Anschlussmotiv	AM_01	3,06	1,625	2,641	0,40	41,2
	AM_02	2,64	1,454	2,113	0,56	32,8
	AM_03	3,23	1,511	2,284	0,53	44,6
	AM_04	2,63	1,492	2,227	0,64	32,6
	AM_05	1,82	1,152	1,326	0,51	16,4
	AM_06	3,09	1,615	2,607	0,42	41,8
	AM_07	3,18	1,490	2,221	0,52	43,6
	AM_08	2,43	1,390	1,933	0,50	28,6
	AM_09	2,89	1,490	2,220	0,56	37,8
	AM_10	3,28	1,679	2,819	0,03	45,6
	AM_11	3,09	1,482	2,196	0,53	41,8
	AM_12	1,76	1,027	1,054	0,49	15,2
	AM_13	2,34	1,308	1,712	0,48	26,8
	AM_14	2,63	1,457	2,123	0,66	32,6
	AM_15	2,44	1,331	1,772	0,64	28,8
	AM_16	1,78	1,091	1,191	0,56	15,6
		\overline{X}	SD	Var(x)	r _{it}	Pi
Soziales Spie-	SS_01	3,81	1,226	1,504	0,48	56,2
geln	SS_02	3,25	1,222	1,493	0,31	45
	SS_03	3,50	1,270	1,612	0,45	50
		\overline{X}	SD	Var(x)	r _{it}	Pi
Selbstwahrneh-	SW_01_invers	4,76	1,345	1,810	0,02	75,2
mung	SW_02	3,69	1,568	2,457	0,48	53,8
	SW_03	3,78	1,502	2,255	0,53	55,6
	SW_04	4,62	1,383	1,911	0,07	72,4
		\overline{X}	SD	Var(x)	r _{it}	Pi

Sozialer Ver-	SV_01	2,83	1,501	2,252	0,57	36,6
gleich	SV_02	3,44	1,412	1,994	0,40	48,8
	SV_03	3,15	1,399	1,958	0,47	43
	SV_04	3,9	1,405	1,974	0,51	58
	SV_05	2,52	1,356	1,840	0,73	30,4
	SV_06	2,27	1,356	1,838	0,68	25,4
	SV_07	2,65	1,444	2,084	0,69	33
	SV_08	2,38	1,332	1,774	0,69	27,6
		\overline{X}	SD	Var(x)	r _{it}	Pi
Selbstkonzept	SK_01	3,10	1,562	2,440	0,68	42
	SK_02	3,24	1,604	2,572	0,57	44,8
	SK_03	3,95	1,482	2,196	0,46	59
	SK_04	2,81	1,544	2,383	0,72	36,2
	SK_05	2,84	1,480	2,191	0,34	36,8
		\overline{X}	SD	Var(x)	r _{it}	Pi
Gruppendyna-	GD_01	2,63	1,314	1,727	0,57	32,6
mik	GD_02	1,99	1,147	1,316	0,70	19,8
	GD_03	1,76	1,158	1,341	0,72	15,2
	GD_04	3,92	1,346	1,810	0,31	58,4
	GD_05	2,02	1,206	1,456	0,67	20,4
	GD_06	2,82	1,529	2,338	0,55	36,4
	GD_07	1,79	1,225	1,501	0,60	15,8
	GD_08	1,43	0,919	0,845	0,59	8,6
		\overline{X}	SD	Var(x)	r _{it}	Pi
Mehrheitsein-	ME_01	2,63	1,437	2,064	0,44	32,6
fluss	ME_02	3,60	1,459	2,129	0,21	52
	ME_03	3,12	1,514	2,292	0,36	42,4
	ME_04_invers	3,34	1,575	2,479	0,08	46,8
	ME_05	2,34	1,223	1,495	0,17	26,8
	ME_06	2,95	1,681	2,825	0,26	39
	ME_07	1,97	1,213	1,472	0,48	19,4
	ME_08	3,53	1,692	2,863	0,31	50,6
		\overline{X}	SD	Var(x)	r _{it}	Pi
Minderwertig-	MW_01	1,91	1,243	1,545	0,68	18,2
keit	MW_02	2,64	1,513	2,289	0,61	32,8
	MW_03	2,36	1,481	2,192	0,69	27,2
	MW_04	2,28	1,466	2,148	0,52	25,6
	MW_05	3,06	1,433	2,054	0,38	41,2
		\overline{X}	SD	Var(x)	r _{it}	Pi
Soziale Aner-	SA_01	3,48	1,489	2,218	0,76	49,6

	SA_03	3,75	1,336	1,784	0,54	55
	SA_04	4,45	1,219	1,486	0,61	69
	SA_05	3,42	1,504	2,263	0,68	48,4
	SA_06	2,86	1,606	2,579	0,69	37,2
	SA_07	1,52	1,041	1,084	0,18	10,4
	SA_08	3,56	1,606	2,580	0,61	51,2
		\overline{X}	SD	Var(x)	r _{it}	Pi
Narzissmus	N_01	2,48	1,378	1,899	0,40	29,6
	N_02	4,61	1,357	1,840	-0,03	72,2
	N_03	2,91	1,301	1,692	0,49	38,2
	N_04	2,36	1,330	1,770	0,51	27,2
	N_05	2,22	1,307	1,708	0,48	24,4
		\overline{X}	SD	Var(x)	r _{it}	Pi
Eskapismus	E_01	1,69	1,033	1,067	0,43	13,8
	E_02	2,69	1,374	1,888	0,45	33,8
	E_03	3,71	1,537	2,364	0,41	54,2
	E_04	3,64	1,571	2,467	0,39	52,8
	E_05	3,21	1,433	2,054	0,35	44,2
	E_06	2,38	1,438	2,068	0,36	27,6
	E_07	3,40	1,565	2,450	0,19	48
	E_08	2,42	1,378	1,899	0,50	28,4
	E_09	3,38	1,669	2,784	0,50	47,6
		\overline{X}	SD	Var(x)	r _{it}	Pi
Fear of missing	FOMO_01	2,32	1,794	3,217	0,88	26,4
out	FOMO_02	2,34	1,806	3,262	0,88	26,8

Anmerkungen: Mittelwert (\overline{x}), Standardabweichung (SD), Varianz (Var(x)), Trennschärfe (r_{it}), Schwierigkeitsindex (P_i)

Alle Items mit einem Trennschärfekoeffizienten r_{it} unter .40 wurden ausgeschlossen. Bezüglich des Schwierigkeitsindizes wurden solche Items bevorzugt, die sich in einem Wertebereich von P_i = 20 und P_i = 80 bewegten. Bei höheren bzw. niedrigeren Schwierigkeiten wurden die Items beibehalten, wenn sie eine hohe Trennschärfe aufwiesen.

3.3.3 Interne Konsistenz der theoretisch angenommenen Skalen

Im Folgenden werden die Reliabilitätskoeffizienten nach Cronbachs Alpha für die theoretisch angenommenen Skalen des FBMSN dargestellt. Die Bestimmung der Internen Konsistenz fußt auf den testtheoretischen Annahmen, welche im Kapitel 2.2 Testtheoretische Grundlagen und insbesondere im Abschnitt 3.1.2 Reliabilität und interne Konsistenz, dargelegt wurden. Nach Bühner (2011) sind Reliabilitätskoeffizienten unter .60 inakzeptabel. Werte über .70 und .80 gelten wiederum als gut bzw. sehr gut. Werte über .90 deuten meist auf redundante Items hin. Weiterhin weist Bühner (2011) darauf hin,

dass eine Reliabilitätsanalyse auf Basis der Konsistenzmethoden nur bei eindimensionalen Testverfahren sinnvoll ist. Entsprechend wurde im Rahmen der Itemanalyse auch die Eindimensionalität bzw. Testhomogenität berücksichtigt. Somit wurden ausschließlich Items mit einer hohen Trennschärfe beibehalten, um eine möglichst genaue Reliabilitätsschätzung zu gewährleisten (Döring & Bortz, 2016). Die Trennschärfen sind dem Anhang B zu entnehmen.

Tab. 2: Übersicht der Reliabilitätskoeffizienten (α) der Skalen des FBMSN vor und nach der Itemselektion

Reliabilitätskoeffizienten (α)

		α (vor Selek-	α (nach Se-	Itemanzahl vor Selek-
		tion)	lektion)	tion (nach Selektion)
Skala	Modelllernen	.83	.83	11 (8)
	Anschlussmotiv	.82	.87	16 (12)
	Intimitätsmotiv	.82	.82	7 (7)
	Soziales Spiegeln	.60	.64	3 (2)
	Selbstwahrnehmung	.46	.83	4 (2)
	Sozialer Vergleich	.85	.86	8 (7)
	Selbstkonzept	.77	.80	5 (4)
	Gruppendynamik	.84	.86	8 (7)
	Mehrheitseinfluss	.59	.67	8 (2)
	Minderwertigkeit	.79	.81	5 (4)
	Soziale Anerkennung	.86	.88	8 (7)
	Narzissmus	.60	.71	5 (3)
	Eskapismus	.72	.63	9 (4)
	Fear of Missing out	.94	.94	2 (2)
Items Ge-	,		,	99 (73)
samt				

An dieser Stelle wurden alle Skalen, die durch weniger als vier Items repräsentiert sind, ausgeschlossen und nicht in weitere Analysen einbezogen. Denn Bühner (2011) weist darauf hin, dass für die Durchführung einer Faktorenanalyse pro erwarteten Faktor mindestens vier Items verwendet werden sollten. Für die verbliebenden Skalen konnten gute bis sehr gute Werte für die interne Konsistenz beobachtet werden. Lediglich die Skala *Eskapismus* erreicht nur einen akzeptablen Wert für Cronbachs Alpha.

3.3.4 Exploratorische Faktorenanalysen

Vor der Durchführung der Faktorenanalysen wurden zunächst die notwendigen Voraussetzungen, die in Kapitel 3.1.3 Faktorenanalyse beschrieben sind, geprüft. Entsprechend wurde der Datensatz auf das Vorliegen substanzielle Korrelationen, die eine Durchführung erlauben, geprüft. Der KMO-Koeffizient (Maß der Stichprobeneignung

nach Kaiser-Meyer-Olkin) liegt mit .924 für den Fragebogen laut Bühner (2011) in einem "sehr guten" Bereich. Weiterhin beweist der Bartlett-Test auf Sphärizität den Korrelationen des Fragebogens eine signifikante Abweichung von null (p=,000). Die Variablen sind somit für Faktorenanalysen geeignet (siehe Anhang D).

3.3.4.1 Erste Prüfung der Skalenstruktur des FBMSN

Um die erhobenen Daten genau zu untersuchen wurden zunächst exploratorische Faktorenanalysen für die, auf Basis der Itemanalyse gekürzte Rohversion des Fragebogens durchgeführt. Dabei wurden die von Burdorf und Lux (2022) entwickelten Themenblöcke beibehalten und als Basis für die theoretisch angenommene faktorielle Struktur genutzt. Allerdings lieferten verschiedene Analysemethoden keine interpretierbaren Ergebnisse. Die Ergebnisse dieser ersten Prüfungsversuche der faktoriellen Struktur sind der Vollständigkeit halber im Anhang D aufgeführt.

Daher erfolgte eine umfangreiche Überarbeitung sowie detaillierte Selektion der Items und Skalen, welche in weitere Analysen einbezogen wurden. Im Zuge dessen wurden sowohl die Kennwerte der Itemanalyse sowie konkrete Inhalte der einzelnen Items als auch die theoretische Operationalisierung der Skalen nach Burdorf und Lux (2022) erneut betrachtet sowie kritisch analysiert.

3.3.4.2 Überarbeitung des Fragebogens

Auf Basis der Itemanalyse wurde eine erste Überarbeitung und Kürzung des Fragebogens vorgenommen. Im Zuge dessen wurden auf Grundlage der ermittelten Kennwerte ungeeignete bereits Items eliminiert.

Weiterhin wurden alle Items im Hinblick auf ihre Formulierung überprüft und über den gesamten Itempool hinweg miteinander verglichen, um inhaltlich redundante Items sowie Wiederholungen für verschiedene Skalen zu identifizieren.

Abschließend erfolgte eine kritische Überprüfung der theoretischen Operationalisierung der einzelnen Konstrukte von Burdorf und Lux (2022). Dabei wurde besonders auf den Bezug zum eigentlichen Ziel des Fragebogens geachtet. Schließlich soll der Fragebogen in erster Linie der Erfassung der zugrundeliegenden Motive der Selbstdarstellung auf sozialen Netzwerken dienen. Bei Betrachtung der einzelnen Skalen und Operationalisierungen zeigte sich, dass diese Zielsetzung zum Teil stark überstiegen worden ist. Auf diesen Umstand verweisen auch Burdorf und Lux (2022) in ihrer Ausarbeitung. In Verbindung mit den bereits dargelegten Ausschlusskriterien wurden daher die Skalen Modelllernen sowie Eskapismus und FOMO aufgrund mangelnder Operationalisierung und/oder fehlender Bezüge zu Motiven der Selbstdarstellung nicht in weitere Analysen einbezogen. Überdies wurde die Skala Minderwertigkeit ebenfalls von weiteren Analysen ausgeschlossen, da die generierten Items keinen direkten Bezug zu sozialen Medien

aufweisen. Außerdem konnten für die Skala ausschließlich Itemschwierigkeiten von P_i < 40 nachgewiesen werden, dies hat zur Folge, dass diese Skala nur wenig Informationen liefert.

3.3.4.3 Analyse der faktoriellen Struktur des überarbeiteten Fragebogens

Analog zur ersten Analyse der faktoriellen Struktur des FBMSN wurden zunächst die Voraussetzungen für die Durchführung einer Faktorenanalyse überprüft. Entsprechend wurde der Datensatz erneut auf das Vorliegen substanzielle Korrelationen, die eine Durchführung erlauben, geprüft. Der KMO-Koeffizient (Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin) liegt mit .923 für den Fragebogen laut Bühner (2011) in einem "sehr guten" Bereich. Weiterhin bescheinigt der Bartlett-Test auf Sphärizität den Korrelationen des Tests eine signifikante Abweichung von null (p=,000). Entsprechend sind die Variablen für die Durchführung von Faktorenanalysen geeignet.

Anschließend wurde die Anzahl der zu extrahierenden Faktoren festgelegt. Dazu wurden folgende Kriterien betrachtet (Pospeschill, 2022; Bühner, 2011):

- Kaiser-Guttman- Kriterium
- Scree-Test nach Cattel
- Minimum-Average-Partial-Test von Velicer (MAP-Test)

Für die Durchführung des MAP-Tests wurde die Syntax von O'Connor (2000) eingesetzt. In der folgenden Tabelle ist die Anzahl der zu extrahierenden Faktoren nach den verschiedenen Kriterien dargestellt.

Tab. 3: Anzahl der zu extrahierenden Faktoren nach verschiedene Kriterien

		Anzahl der Faktoren
Kriterium	Kaiser-Guttmann-Kriterium	8
	Scree-Test nach Cattel	5 - 7
	MAP-Test	7

Nach einer ersten exploratorischen Faktorenanalyse auf Grundlage des Kaiser-Gutmann- bzw. Eigenwertkriteriums lieferte bereits eine sinnvoll interpretierbare Faktorlösung. Allerdings verlieb ein Faktor, der nicht weiter interpretiert werden konnte. Daher fiel die Wahl, auf Grundlage des MAP-Tests, auf eine Lösung mit sieben Faktoren. Diese konnten alle sinnvoll interpretiert werden. Der Scree-Test wurde aufgrund seiner Subjektivität bei der Beurteilung des Eigenwertverlaufs als grobe Schätzung eines Bereichs, in dem die Faktorenanzahl sich bewegen könnte, genutzt.

Entsprechend wurde für den FBMSN eine Hauptkomponentenanalyse mit sieben Komponenten sowie anschließender Promax – Rotation durchgeführt. Die Faktorlösung mit den sieben Faktoren erklärt 62,651 % der Gesamtvarianz.

In der nachfolgenden Tabelle ist die rotierte Mustermatrix dargestellt. Faktorladungen unter .30 wurden nicht berücksichtigt und sind entsprechend nicht mitaufgeführt.

Tab. 4: Rotierte Mustermatrix des FBMSN (Hauptkomponentenanalyse mit 7 Faktoren, Promax-Rotation)

Mustermatrix inklusive Varianzaufklärung

Komponente	11	2	3	4	5	6	7
Varianzaufklärung	33,846 %	8,181 %	6,392 %	5,034 %	3,536 %	2,907 %	2,756 %
SA_04	,822		-,310				
SA_01	,817						
SA_03	,751						
SA_02	,742						
SA_06	,730						
SA_08	,694						
SA_05	,641						
AM_11	,566						
SV_07		,903					
SV_06		,861					
SV_05		,826					
SV_08		,769					
AM_02		,533					
SV_01	,350	,503					
GD_02			,871				
GD_03			,764				
GD_05			,760				
GD_08			,678				
GD_01			,585				
GD_07			,480				,373
GD_06	,316		,374				
AM_06				,731		,327	
AM_14			,338	,681			
AM_15				,675			
AM_08				,670			
AM_09				,653			
AM_03				,566			
AM_07	,445			,448			
AM_04				,402			
SK_03					,822		
SK_04					,690		
SK_01					,643		
AM_13					,635		
SK_02					,535		
SV_04		,332			,375		
IM_05						,868	
IM_04						,843	

IM_06	,680	
IM_02		,858
IM_01		,782
IM 07	.314	,427

Anschließend wurde für die überarbeitete Fragebogenversion eine Hauptachsenanalyse mit Promax-Rotation für die sieben Faktoren durchgeführt. Diese ist darauf ausgerichtet, latente Konstrukte in Form von Faktoren zu identifizieren, durch die sich die Beziehungsstruktur der beobachteten Variablen erklären lässt (Pospeschill, 2022). Im Zuge dessen sollte die faktorielle Struktur des Fragebogens bestätigt werden. Der Fokus lag dabei vor allem auf den Items SV_04 und AM_07, da die Faktorladungen für zwei Faktoren besonders nah beieinander liegen. Die nachfolgende Tabelle zeigt die rotierte Mustermatrix der Hauptachsenanalyse. Faktorladungen unter .30 wurden nicht berücksichtigt und sind somit nicht aufgeführt.

Tab. 5: Rotierte Mustermatrix des FBMSN (Hauptachsenanalyse mit 7 Faktoren, Promax-Rotation)

NM	uste	$\triangle rm$	19t	rıv

<u>-</u>	Faktor								
	1	2	3	4	5	6	7		
SA_01	,835								
SA_04	,792	-,302							
SA_02	,723								
SA_03	,684								
SA_06	,684								
SA_08	,656								
SA_05	,640								
AM_11	,551								
AM_07	,420				,402				
GD_02		,804							
GD_03		,794							
GD_08		,743							
GD_05		,719							
GD_07		,501							
GD_01		,408							
GD_06		,340							
SK_04			,888,						
SK_01			,713						
AM_13			,651						
SK_03			,550						
SK_02			,485						
SV_04			,349						
SV_05				,838					
SV_06				,826			5		

SV_07			,798			
SV_08			,664			
SV_01	,326		,428			
AM_02		,319	,379			
AM_14				,665		
AM_15				,661		
AM_06				,630		
AM_08				,586		
AM_09				,573		
AM_03				,483		
AM_04				,357		
IM_05					,843	
IM_04					,785	
IM_06					,591	
IM_02						,865
IM_01						,775
IM_07						,334

Insgesamt bestätigt die Hauptachsenanalyse die Ergebnisse der Hauptkomponentenanalyse. Bezüglich des Items SV_04 konnte Klarheit über die Faktorzugehörigkeit geschaffen werden. Das Item AM_07 wurde eliminiert, da auch die Hauptachsenanalyse keine eindeutige Lösung bietet. Bevor das Item endgültig ausgeschlossen wurde, ist dessen Einfluss auf die Reliabilität berücksichtigt worden.

3.3.5 Skalenzusammensetzung und Kennwerte des revidierten Fragebogens

Das nachfolgende Kapitel bildet die endgültige Version des Fragebogens inklusive relevanter Kennwerte ab. Zunächst ist an dieser Stelle darauf hinzuweisen, dass mit Selektion der Items sowie Umstrukturierung des Fragebogens eine Veränderung der Datenmatrix einhergeht. Daraus resultiert, dass die Testwerte und die relevanten Koeffizienten neu berechnet werden müssen (Kelava & Moosbrugger, 2020a). Weiterhin enthält dieses Kapitel die Skalen-Korrelationsmatrix, welche später als eine Grundlage für Interpretation der Faktoren bzw. Skalen genutzt werden soll.

Entsprechend zeigt die nachfolgende Tabelle die neu berechneten Trennschärfekoeffizienten der verbliebenen Items sowie Reliabilitätskoeffizienten nach Cronbachs Alpha für die faktoranalytisch ermittelten Skalen in ihrer finalen Zusammensetzung. Die Benennung der Faktoren ist hier zunächst vorläufig und somit allgemein gehalten, da eine detaillierte Interpretation noch aussteht (siehe dazu Kapitel 3.4.2 Interpretation der Faktoren). Für die revidierte Version des FBMSN wird im Folgenden die Bezeichnung FBMSN-R festgelegt.

Tab. 6: Skalen und Items des FBMSN-R inklusive Trennschärfe- und Reliabilitäts-Koeffizienten

FBMSN-R

		Items	r _{it}	α
to-	Soziale Aner- kennung	SA_01 Likes auf Social Media geben mir ein gutes Gefühl.	.76	.89
	3	SA_02 Likes auf Social Media bestärken mich in meiner Person.	.73	
		SA_03 lch bekomme viel positives Feedback auf meine Beiträge.	.53	
		SA_04 Ich freue mich über positives Feedback unter meinen Beiträgen.	.64	
		SA_05 Ich hoffe auf Feedback, wenn ich etwas hochlade.	.72	
		SA_06 Ich kontrolliere regelmäßig die Anzahl meiner Li-	.67	
		kes. SA_08 Ich mag es positives Feedback für mein Aussehen auf Social Media zu bekommen.	.62	
		AM_11 Ich erwarte Feedback, wenn ich etwas gepostet habe.	.64	
	Soziale Verale	i-SV_01 Ich vergleiche mein Social Media Profil mit den	.59	.8.
	che	Profilen meiner Freunde.	.59	.0
		SV_05 Wenn ich mir andere Profile anschaue, emp-	.78	
		finde ich Neid.		
		SV_06 Wenn ich mir andere Profile schaue, empfinde	.76	
		Eifersucht.		
		SV_07 Wenn ich mir andere Profile anschaue, bin ich unzufrieden mit mir.	.73	
		SV_08 Wenn ich mir andere Profile anschaue, bin ich unzufrieden mit meinem Leben.	.73	
		AM_02 Ich habe Angst, dass mich andere langweilig	.61	
		finden.		
	Konformitäts-	GD_01 Die Bilder auf meinem Profil ähneln denen meiner Freunde.	.54	.8
	druck	GD_02 Ich poste meistens, wenn meine Freunde auch etwas posten.	.71	
		GD_03 Ich fühle mich unter Druck gesetzt, wenn meine Freunde etwas gepostet haben und ich nicht.	.74	
		GD_05 Meine Freunde bringen mich dazu, mich auf Social Media zu präsentieren.	.69	
		GD_06 Ich poste Bilder mit meinen Freunden, um ihnen zu zeigen, wie wichtig sie mir sind.	.54	
		GD_07 Ich treffe mich mit meinen Freunden zu Fotoshootings, um Fotos für Social Media zu machen.	.61	
		GD_08 Ich muss Bilder auf Social Media hochladen, um	.63	
		in meinem Freundeskreis akzeptiert zu werden		

	AM_04 Mir ist es wichtig, dass ich mit meinen Beiträgen niemanden langweile.	.54	
	AM_06 Ich habe auf Social Media schon einmal Menschen getroffen, die mich verstehen.	.48	
	AM_08 Ich lerne Freunde über Social Media kennen.	.53	
	AM_09 Durch meine Beiträge kann ich mich ausdrücken.	.60	
	AM_14 Durch meine Beiträge erhoffe ich mir Kommuni-	.65	
	kation mit anderen.		
	AM_15 Ich nutze Social Media, um zwischenmenschli-	.65	
	che Beziehungen aufzubauen.		
Selbstkonzept	SK_01 Negative Kommentare verletzen mich sehr.	.73	.84
und Selbstwert	SK_02 Ich möchte nicht, dass meine Makel auf Social Media gesehen werden.	.60	
	SK_03 Wenn jemand meine Beiträge nicht mag, soll die Person es für sich behalten.	.48	
	SK_04 Durch negative Kommentare zweifle ich an mir.	.77	
	SV_04 Ich bewundere das Aussehen einiger Menschen	.53	
	auf Social Media.		
	AM_13 Mit negativem Feedback in der Realität wäre ich	.61	
	überfordert.		
Intimitätsmotiv Liebesbezie-	-IM_04 Ich habe mich schon mal über Social Media zu einem Date verabredet.	.71	.82
hungen	IM_05 Ich habe schon mal eine Liebesbeziehung über Social Media gestartet/angefangen.	.76	
	IM_06 Ich nutze Social Media, um Liebesbeziehungen	.59	
	aufzubauen.		
Intimitätsmotiv	-IM_01 Ich nutze Social Media, um anderen meinen Kör-	.75	.83
Freizügigkeit	per zu zeigen. IM_02 Ich zeige mich gerne freizügig auf Social Media.	.74	
	IM_07 Ich beabsichtige mit meinen Beiträgen die Men-	.59	
	schen zu verführen.		

Anmerkungen: Trennschärfe (r_{it}), Reliabilitätskoeffizient (α).

Für alle Subskalen des revidierten Fragebogens konnten sehr gute Werte für die interne Konsistenz nach Cronbachs Alpha beobachtet werden. Weiterhin fallen die Trennschärfen aller verbliebenden Items des FBMSN-R "gut" bis "sehr gut" aus.

3.3.5.1 Interskalenkorrelation

Wie bereits im Zuge der Faktorenanalyse erwähnt wurde, wird von einer Korrelation der Faktoren ausgegangen. Allerdings kann das Ausmaß der Korrelation genutzt werden, um die Skalenzusammensetzung inhalts-theoretisch zu interpretieren und zu rechtfertigen. Die Skalenwerte werden als Rohwerte durch Summierung der Itemwerte der jeweiligen Skala bestimmt. Entsprechend wird eine Produkt-Moment-Korrelation nach

Pearson für die Skalen berechnet. Somit zeigt die folgende Tabelle die Inter-Skalenkorrelationsmatrix.

Tab. 7: Interskalen-Korrelationen des FBMSN-R

Korrelationen

Korreia	ationen							
		SA	SV	GD	AM	SK	IM_LB	IM_FG
SA	r	1						
	Sig. (2-seitig)							
	N	308						
SV	r	,534**	1					
	Sig. (2-seitig)	<,001						
	N	308	308					
GD	r	,485**	,591**	1				
	Sig. (2-seitig)	<,001	<,001					
	N	308	308	308				
AM	r	,516**	,536**	,629**	1			
	Sig. (2-seitig)	<,001	<,001	<,001				
	N	308	308	308	308			
SK	r	,577**	,699**	,451**	,445**	1		
	Sig. (2-seitig)	<,001	<,001	<,001	<,001			
	N	308	308	308	308	308		
IM_LB	r	,222**	,273**	,317**	,400**	,199**	1	
	Sig. (2-seitig)	<,001	<,001	<,001	<,001	<,001		
	N	308	308	308	308	308	308	
IM_F	r	,402**	,493**	,614**	,551**	,368**	,423**	1
G	Sig. (2-seitig)	<,001	<,001	<,001	<,001	<,001	<,001	
	N	308	308	308	308	308	308	308

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Anmerkungen: r = Korrelationskoeffizient nach Pearson, SA = Soziale Anerkennung, SV = Sozialer Vergleich, GD = Gruppendynamik, AM = Anschlussmotiv, SK = Selbstkonzept, IM_LB = Intimitätsmotiv Liebesbeziehungen, Intimitätsmotiv Freizügigkeit.

Wie erwartet weisen alle Skalen unterschiedlich hohe signifikant positive Korrelationen auf. Dennoch korrelieren alle unterschiedlich stark. Im Rahmen der Diskussion sollen die Korrelationen ausführlicher interpretiert werden.

3.4 Diskussion und Interpretation der Ergebnisse - Studie I

Im Folgenden werden die Ergebnisse der ersten Studie inhaltlich interpretiert und im Hinblick auf die testtheoretische Basis sowie wissenschaftliche Literatur diskutiert. Ferner soll auf Einschränkungen und Limitierungen dieser Ersterhebung des Fragebogens eingegangen werden.

3.4.1 Fragebogenanalyse und -revision

Nach Abschluss der Itemanalyse konnte eine erste Kürzung auf Basis der Kennwerte vorgenommen werden. Im Zuge dessen wurden sämtliche Items eliminiert, die einen Trennschärfekoeffizienten r_{it} unter .40 erreichten. Bezüglich des Schwierigkeitsindizes wurden solche Items bevorzugt, die sich in einem Wertebereich von P_i = 20 und P_i = 80 bewegten. Bei höheren bzw. niedrigeren Schwierigkeiten wurden die Items beibehalten, wenn sie eine hohe Trennschärfe aufwiesen. Die gesetzten Vorgaben wurden relativ streng gewählt. Laut Pospeschill (2022) deuten Trennschärfekoeffizienten von .30 bis .39 darauf hin ,dass die Items ggf. durch eine Anpassung verbessert werden könnten. Weiterhin kennzeichnen Werte im Bereich von .20 und .29 Items, welche definitiv verändert werden sollten. Nach Pospeschill (2022) sollten Items erst ab einem Wert von unter .19 definitiv eliminiert werden. Allerdings wurde aufgrund des zeitlichen Rahmens dieser Masterthesis auf eine Anpassung oder Neuformulierung der Items verzichtet.

Weitere Selektionen geschahen auf Basis der Annahmen von Bühner (2011) zur ausreichenden Itemanzahl pro Faktor. Entsprechend sind sämtliche Skalen vollständig entfernt worden, die nach der Itemanalyse aus weniger als vier Items bestanden. Anschließend wurde eine erste exploratorische Faktorenanalyse für die Skalen Modelllernen, Anschlussmotiv, Intimitätsmotiv, Sozialer Vergleich, Selbstkonzept, Gruppendynamik, Minderwertigkeit sowie Soziale Anerkennung und Eskapismus anhand der verbliebenen 56 Items durchgeführt. Diese lieferte jedoch keine interpretierbaren Ergebnisse. Daher wurde eine erneute Überprüfung der überarbeiteten Version vorgenommen. Im Zuge dessen wurden alle Items hinsichtlich auf ihrer inhaltlichen Formulierung überprüft und über den gesamten Itemsatz hinweg miteinander verglichen, um redundante Items sowie Wiederholungen für verschiedene Skalen zu identifizieren. Abschließend wurden die generierten Items im Hinblick auf die Operationalisierung von Burdorf und Lux (2022) erneut kritisch betrachtet. Außerdem wurde in diesem Zusammenhang Bezug auf das eigentliche Ziel des Fragebogens genommen. Im Zuge dessen zeigte sich, dass die Zielsetzung zum Teil stark überstiegen worden ist. Auf diese Tatsache verweisen auch Burdorf und Lux (2022) in ihrer Ausarbeitung. In Verbindung mit den bereits dargelegten Ausschlusskriterien wird die Skala Eskapismus aufgrund mangelnder Operationalisierung nicht in weitere Analysen einbezogen. So wurde im Rahmen des Kapitels 2.1.7 Eskapismus bereits beschrieben, dass Eskapismus ein Konstrukt ist, welches kaum zu operationalisieren ist. Weiterhin finden sich in der Literatur viele verschiedene Definitionen. Des Weiteren wurde die Skala Minderwertigkeit ebenfalls von weiteren Analysen ausgeschlossen, da die generierten Items ohne direkten Bezug zu sozialen Medien formuliert wurden. Für die Skala Modelllernen lässt sich festhalten, dass die Items nach der Itemselektion, sich eher auf die Identifikation mit einem Influencer und weniger auf

Motive der Selbstdarstellung auf sozialen Netzwerken beziehen. Daher wurde auch diese Skala von weiteren Analysen ausgeschlossen.

Nach Durchführung der Faktorenanalysen konnte eine umstrukturierte Version des Fragebogens erstellt werden (FBMSN-R). Diese setzt sich aus 40 Items zusammen, welche auf sieben interpretierbare Faktoren verteilt sind. Für alle Subskalen konnten sehr gute Werte für die interne Konsistenz nach Cronbachs Alpha beobachtet werden. Des Weiteren liegen die Trennschärfen aller verbliebenden Items des FBMSN-R einem, "gut" bis "sehr gut" Bereich (siehe Kapitel 3.3.5 Skalenzusammensetzung und Kennwerte des revidierten Fragebogens).

Im Hinblick auf die Zielsetzung der ersten Studie lässt sich festhalten, dass eine revidierte Version des Fragebogens konzipiert werden konnte, die den testtheoretischen Anforderungen genügt und dessen Skalenstruktur empirisch geprüft ist. Entsprechend setzt sich der FBMSN-R aus reliablen Skalen mit geeigneten Items zusammen. Des Weiteren kann festgehalten werden, dass die revidierte Version mit deutlich weniger Items und Skalen auskommt als die Erstversion.

3.4.2 Interpretation und Definition der Faktoren

Im folgenden Absatz erfolgt die inhaltliche Interpretation der Faktoren auf Basis des theoretischen Hintergrundes und aktuellen Forschungen. Die Faktoren sind in der folgenden Tabelle kurz dargestellt, eine ausführliche Fassung inklusive der ausformulierten Items und relevanten Kennwerte findet sich in Kapitel 3.3.5 Skalenzusammensetzung und Kennwerte des revidierten Fragebogens.

Tab. 8: Faktoren inklusive Itemanzahl

		Anzahl der Items
Faktor	1: Streben nach sozialer Anerkennung	8
	2: Soziale Vergleiche	6
	3: Konformitätsdruck und Gruppendynamik	7
	4: Anschlussmotiv	7
	5: Selbstkonzept und Selbstwert	6
	6: Liebesbeziehungen	3
	7: Freizügige Selbstdarstellung	3

In den nachfolgenden Abschnitten wird jeder Faktor aufgrund der zugeordneten Items im Hinblick auf die theoretischen Hintergründe interpretiert und somit die gewählte Faktorlösung untermauert. Außerdem werden ebenfalls Ergebnisse aktueller Forschungen miteinbezogen.

3.4.2.1 Streben nach sozialer Anerkennung

Die Ergebnisse der Faktorenanalyse entsprechen für die Skala Streben nach Sozialer Anerkennung dem bereits ausführlich, im Rahmen des Kapitels 2.1.5.4 Soziale Anerkennung geschilderten Verhaltensmodell nach Fogg (2009). In diesem Zusammenhang ist hauptsächlich der dritte Kernmotivator von besonderer Bedeutung. Nach Fogg (2009) bezieht sich dieser auf das gesamte Sozialverhalten des Individuums. So streben Menschen stets nach sozialer Anerkennung und Bestätigung, um ein positives Selbstwertgefühl aufrechtzuerhalten (Fogg, 2009). Aus vielen Studien geht hervor, dass die Selbstdarstellung auf Social Media eng mit Wohlbefinden und Zufriedenheit der Nutzer verbunden sind (Valkenburg & Peter, 2009). Dem zur Folge können Interaktionen und Feedback auf einer Social-Media-Plattform Einfluss auf das Selbstwertgefühl haben. Die Ergebnisse der Studien von Gonzales und Hancock (2011) stützen diese Interpretation: So führt schon die Betrachtung des eigenen Profils eine Selbstwertsteigerung herbei. Nach Leary (1995), versuchen Menschen einen positiven Eindruck zu hinterlassen, der zu sozialen Belohnungen (z. B. Komplimente, Einladungen zu gemeinsamen Aktivitäten) führt. Entsprechend kann dieser Faktor als ein Motiv der Selbstdarstellung auf sozialen Medien für die vorliegende Stichprobe festgehalten werden. Eine hohe Ausprägung auf dieser Skala sollte entsprechend mit einem starken Bedürfnis nach sozialer Anerkennung einhergehen.

3.4.2.2 Soziale Vergleiche

Das Phänomen der sozialen Vergleiche und den entsprechenden negativen Emotionen bildet sich durch die Items im zweiten Faktor ab. Die Zusammensetzung deckt sich mit der Operationalisierung im Kapitel 2.1.4.3 Soziale Vergleiche. Die aktuelle Forschung zeigt, dass der ständige Vergleichsdruck negative Auswirkungen auf das Wohlbefinden und den Selbstwert der Nutzer von sozialen Netzwerken hat (Steers, Wickham & Acitelli, 2014). Außerdem zeigen empirische Studien, dass die Verfügbarkeit an Informationen über Erfolge und Erlebnisse anderer Menschen negative Einflüsse auf die Lebenszufriedenheit haben können. Entsprechend deuten mittlerweile einige Studien darauf hin, dass es bei der Betrachtung von Profilen auf Social-Media-Plattformen zu sozialen Vergleichsprozessen nach Festinger (1954) kommt. Diese Vergleichsprozesse führen bei attraktiven Profilfotos zu geringerer Zufriedenheit mit dem eigenen Erscheinungsbild (Haferkamp & Krämer, 2011). Weitere Untersuchungen zeigen, dass Studenten, welche seit längerem Facebook nutzen, eher glauben, dass andere Menschen glücklicher sind als sie selbst (Chou & Edge, 2012). Allerdings verweist Crusius (2022) darauf, dass die sozialen Aufwärtsvergleichsprozesse ebenfalls mit einer gesteigerten Leistungsmotivation einhergehen können. Entsprechend ist zu vermuten, dass eine hohe Ausprägung mit häufigen sozialen Vergleichsprozessen auf sozialen Netzwerken einhergehen.

Zudem lässt sich eine hohe Ausprägung sich mit entsprechenden negativen Emotionen (z. B. Neid, Unzufriedenheit mit der eigenen Person) infolge dieser Vergleichsprozesse in Verbindung setzen.

3.4.2.3 Konformitätsdruck und Gruppendynamik

Auf Basis der Operationalisierung in Kapitel 2.1.5.1 Gruppendynamik lässt sich dieser Faktor und die entsprechenden Items vordergründig auf wirkenden Konformitätsdruck zurückführen (Hewstone & Martin, 2014). Ebenso kann die Skala als eine Form oder Wirkung eines besonderen Bedürfnisses nach Zugehörigkeit verstanden werden. Entsprechend wird eine hohe Ausprägung auf dieser Skala als ein besonderes Streben nach sozialer Zugehörigkeit interpretiert. Weiterhin lässt sich aufgrund einer hohen Ausprägung vermuten, dass durch die Anpassung und konformes Verhalten negatives Feedback vermieden werden bzw. die Zugehörigkeit zur Gruppe gesichert werden soll.

3.4.2.4 Anschlussmotiv

Für die Skala Anschlussmotiv wird ebenfalls auf die Operationalisierung in Kapitel 2.1.3.1 Anschlussmotiv verwiesen. Im Allgemeinen bestimmt die Ausprägung des Anschlussmotivs das Streben nach Anschluss und Kontakt an andere Personen. Entsprechend weisen Personen mit einer höheren Ausprägung eine erhöhte Präferenz dafür auf, mit anderen (auch fremden) Menschen Kontakt aufzunehmen (Neyer & Asendorpf, 2018). Im Zusammenhang mit sozialen Netzwerken wird vermutet, dass sich diese Präferenz auch auf Social-Media-Plattformen beobachten lässt. Entsprechend bildetet diese Skala ein besonderes Streben nach Anschluss und sozialen Interaktionen auf sozialen Netzwerken ab. So wird das Teilen von Informationen auf sozialen Netzwerken von Bedürfnissen bestimmt, die sozialen Kontakt beeinflussen (Ho & Dempsey, 2010). Dies gilt vorwiegend für das Bedürfnis nach Zugehörigkeit, welches besonderen Einfluss auf die Art sowie den Umfang der Nutzung von Social-Media-Plattformen hat (Krämer, Eimler, & Neubaum, 2017). Weiterhin zeigen Deters und Mehl (2013), dass bereits das bloße Teilen einer Statusmeldung dazu führt, dass Probanden sich weniger einsam fühlen, selbst wenn daraus keine Interaktion mit Anderen entsteht. Somit kann ein hoher Testwert als starkes Anschlussmotiv bzw. streben nach Kontakt und Interaktionen auf sozialen Netzwerken interpretiert werden.

3.4.2.5 Selbstkonzept und Selbstwert

Für den Faktor *Selbstkonzept und Selbstwert* bilden die Items vorwiegend negative Einflüsse und Auswirkungen durch Interaktionen auf sozialen Netzwerken ab. Auf Basis der Operationalisierung in Kapitel *2.1.4.4 Selbstkonzept* kann diese Skala darauf hindeuten, dass soziale Netzwerke als wesentliche Informationsquelle für das Selbstkonzept

wirken. Entsprechend wurden negative Auswirkungen auf das Wohlbefinden und den Selbstwert der Nutzer von sozialen Netzwerken bereits in verschiedenen Studien nachgewiesen (Steers, Wickham & Acitelli, 2014). So zeigen empirische Untersuchungen, dass negative Interaktionen und Rückmeldungen aus sozialen Netzwerken zu Wut, Scham sowie Hilflosigkeit und Depressionen führen können (Ortega, Paz Elipe, Mora-Merchán, & Vega, 2009). Außerdem deuten die Ergebnisse von Sumter, Baumgartner, Valkenburg und Peter (2012) darauf hin, dass diese Erfahrungen die Lebenszufriedenheit langfristig negativ beeinflussen. Daraus lässt sich schließen, dass eine hohe Ausprägung daraufhin deuten kann, dass negatives Feedback auf sozialen Netzwerken einen erheblichen Einfluss auf das Selbstwertgefühl hat. Außerdem kann eine hohe Ausprägung auf dieser Skala ebenfalls mit einer erhöhten Verletzlichkeit im Hinblick auf negatives Feedback auf sozialen Netzwerken interpretiert werden.

3.4.2.6 Liebesbeziehungen

Im Rahmen der Faktorenanalyse zeigte sich, dass das Intimitätsmotiv im Hinblick auf soziale Netzwerke in zwei Faktoren teilt. Entsprechend kann die Operationalisierung, die in Kapitel 2.1.3.2 Intimitätsmotiv nur bedingt zur Interpretation dieses Faktors beitragen. Bei Betrachtung der inhaltlichen Formulierung der Items und Trennschärfen bildet diese Skala eher ab, ob eine Person schon mal eine Liebesbeziehung über ein soziales Netzwerk aufgebaut hat oder nicht. Entsprechend bildet diese Skala kein wirkliches Motiv der Selbstdarstellung ab, sondern nimmt eine Kategorisierung vor.

3.4.2.7 Freizügige Selbstdarstellung

Wie bereits erwähnt, stellt die *Freizügige Selbstdarstellung* einen eigenen Faktor dar. Folglich kann auch dieser Faktor nicht vollständig durch die Operationalisierung (siehe Kapitel *2.1.3.2 Intimitätsmotiv*) interpretiert werden. Schließlich muss auch für diesen Faktor nach Betrachtung der inhaltlichen Formulierung und der Trennschärfen davon ausgegangen werden, dass diese Skala eher eine Kategorisierung vornimmt. Somit kann eine hohe Ausprägung lediglich dahin gehend interpretiert werden, ob sich jemand freizügig auf sozialen Netzwerken darstellt oder nicht.

3.4.2.8 Interpretation der korrelativen Zusammenhänge

Wie bereits beschrieben bestehen zwischen allen Faktoren bzw. Skalen verschieden stark ausgeprägte korrelative Zusammenhänge. Im Folgenden sollen diese im Hinblick auf die Operationalisierung im Kontext der Selbstdarstellung auf sozialen Netzwerken interpretiert werden.

Die Skala Soziale Vergleiche weist zu allen anderen Skalen (mit Ausnahme der Skala Liebesbeziehungen) die stärksten Zusammenhänge auf. Diese lassen sich auf Basis der

Theorie sozialer Vergleichsprozesse von Festinger (1954) erklären. Wie bereits im Kapitel 2.1.4.3 Soziale Vergleiche beschrieben, neigt der Mensch dazu, sich ständig mit anderen Menschen zu vergleichen, so offensichtlich auch auf sozialen Netzwerken. Außerdem geht aus der Theorie hervor, dass ein Kernmotiv des sozialen Vergleichs das Streben nach einem positiven Selbstwert ist. Dies verdeutlicht ebenfalls die hohe Korrelation zu den Skalen Streben nach Sozialer Anerkennung und Selbstkonzept und Selbstwert. Wie erwartet fällt auch die Korrelation zwischen den Skalen Konformitätsdruck und Gruppendynamik und Anschlussmotiv hoch aus. Schließlich stellt ein Bedürfnis nach Konformität und Zugehörigkeit gewissermaßen eine Facette des Anschlussmotivs dar. Die Skala Liebesbeziehungen weist die schwächsten Korrelationen zu allen anderen Skalen auf. Lediglich zu den Skalen Freizügige Selbstdarstellung und Anschlussmotiv liegen die Koeffizienten in einem mittleren Bereich. Dies ist ebenfalls im Hinblick auf die theoretische Fundierung im Rahmen des Kapitels 2.1.3 Motive plausibel. Für den Faktor Freizügige Selbstdarstellung lassen sich hauptsächlich hohe Korrelationen zu den Faktoren Gruppendynamik und Konformitätsdruck und Anschlussmotiv beobachten lassen. Das könnte dahin gehend interpretiert werden, dass Personen sich aufgrund vom Konformitätsdruck und dem Bedürfnis nach Anschluss freizügig auf Social Media präsentieren. Allerdings bedarf es für diese Interpretation einer ausführlicheren Untersuchung. Schließlich kann auf Grundlage einer Korrelation keine Aussage über Kausalbeziehungen zweier Variablen getroffen werden (Eid, Gollwitzer, & Schmitt, 2017).

3.4.3 Kritische Reflexion - Studie I

Im Rahmen des folgenden Kapitels soll auf verschiedene Punkte eingegangen, die zu einer Limitierung der Aussagekraft und Übertragbarkeit der Ergebnisse sowie Interpretationen dieser Studie führen könnten. Im Zuge dessen soll eine kritische Reflexion des bisherigen Forschungsprozesses erfolgen. Dabei sollen sowohl die beobachteten Ergebnisse als auch deren Interpretation im Fokus stehen. Des Weiteren soll auch auf Limitierungen und Einschränkungen eingegangen werden.

3.4.3.1 Ergebnisse und Interpretation

Im Hinblick auf die Interpretierbarkeit der Ergebnisse ist festzuhalten, dass durch die Nutzung der Software *Unipark*, die eine Online-Befragung ermöglichte, eine sehr hohe Auswertungs- und Durchführungsobjektivität gewährleistet werden konnte. Weiterhin konnten für die Skalen gute bis sehr gute Reliabilitäten beobachtet werden. Außerdem konnten die Interpretationen der Faktoren *Streben nach Sozialer Anerkennung, Soziale Vergleiche, Konformitätsdruck und Gruppendynamik* sowie *Anschlussmotiv* und *Selbstkonzept und Selbstwert* durch die Operationalisierung gestützt werden. Lediglich die Faktoren *Liebesbeziehungen* und *Freizügige Selbstdarstellung* konnten nicht das im

Hinblick auf die Operationalisierung interpretiert werden. Zudem setzen sich die beiden Faktoren lediglich aus drei Items zusammen, wodurch eine genaue Interpretation deutlich erschwert wird. An dieser Stelle bedarf es einer tiefergehenden Analyse und der Generierung weiterer Items, um die Faktoren im Sinne der Operationalisierung im Zusammenhang mit dem Intimitätsmotiv interpretieren zu können.

3.4.3.2 Limitierungen und Einschränkungen

Der zentralste Punkt, der im Rahmen dieser Untersuchung im Hinblick auf Einschränkungen der Interpretation der Ergebnisse zu nennen ist, ist die Zusammensetzung der Stichprobe. Die vorliegende Stichprobe erlaubt in keine Rückschlüsse auf die Grundgesamtheit. Die im Rahmen dieser Thesis untersuchte Stichprobe kann nicht als repräsentativ für die Grundgesamtheit betrachtet werden. Entsprechend beziehen sich die Ergebnisse lediglich auf die vorliegende Stichprobe. Außerdem wurde der Fragebogen von Burdorf und Lux (2022) für eine Zielgruppe 18 bis 35 Jahren konzipiert. Diese wurde im Rahmen dieser Untersuchung etwas überschritten, um eine möglichst große Stichprobe zu akquirieren. Des Weiteren erfolgte die Stichprobenakquise in erster Linie über persönliche Kontakte des Autors und Studenten der Hochschule Fresenius Köln, da diese sich für eine Teilnahme 0,25 Versuchspersonenstunden anrechnen lassen konnten. Daraus resultiert, dass die Stichprobe größtenteils aus Studierenden sowie Akademikern besteht und verhältnismäßig wenig andere Berufsgruppen oder Bildungsschichten in die Studie einbezogen wurden. In diesem Zusammenhang muss somit davon ausgegangen werden, dass die vorliegende Analysestichprobe als nicht repräsentativ für die angedachte Zielgruppe bewertet werden muss. Entsprechend bedarf es weiterer Analysen anhand einer Stichprobe, die als repräsentativ für die Zielgruppe beurteilt werden kann. Ferner bringt die Erfassung über Selbsteinschätzungen der Probanden grundsätzlich die Gefahr systematischer Urteilsfehler, wie die Verzerrung der Einschätzungen durch soziale Erwünschtheit, mit sich (Eid, Gollwitzer, & Schmitt, 2017). So wurde zur Vermeidung von Verzerrungen durch die Einschätzung, im Sinne der sozialen Erwünschtheit, auf die Anonymität und die vertrauliche Behandlung der Daten hingewiesen.

Im Rahmen der Itemanalyse und -selektion wurde versucht, möglichst klare Kriterien sowie Grenzwerte zu realisieren. Nichtsdestotrotz kann eine gewisse Subjektivität im Hinblick auf die Auswahl und Eliminierung der Items nicht vollständig ausgeschlossen werden. Weiterhin müssen im Rahmen einer Itemanalyse alle Werte und Inhalte stets im Gesamtzusammenhang betrachtet werden, was wiederum eine eindeutige Lösung erschwert. Ähnliches gilt im Zusammenhang mit den durchgeführten Faktorenanalysen. Vor allem für die Wahl der Extraktions- und Rotationmethode sowie der Faktorenanzahl können subjektive Einflüsse auf die Entscheidungen nicht ausgeschlossen werden. Weiterhin ist an dieser Stelle darauf hinzuweisen, dass nach der ersten Reversion des

Fragebogens eine Testung an einer neuen Stichprobe hätte erfolgen müssen. Allerdings musste aus zeitökonomischen Gründen auf eine erneute Stichprobenakquise verzichtet werden. Ferner ist anzumerken, dass durch das explorative Vorgehen eine sehr große Menge an möglicherweise relevanten Informationen verloren gegangen ist.

3.4.3.3 Theoretischer Hintergrund

Im Rahmen der Literaturrecherche stellte sich heraus, dass ein sehr umfangreiches Angebot an Büchern und Forschungsartikeln zur testtheoretischen Basis und Fragebogenkonstruktion vorliegt. Außerdem konnte für die Interpretation der Faktoren auf eine Vielzahl von Fachbüchern sowie Studien und Forschungsartikeln zum Thema Social Media zurückgegriffen werden. Allerdings finden sich in der Literatur überwiegend Studien zur Plattform *Facebook* und kaum Materialien zu anderen Social-Media-Plattformen, wie z. B. *Instagram* oder *TikTok*.

3.5 Fazit - Studie I

Abschließend lässt sich festhalten, dass eine revidierte Version des Fragebogens entwickelt werden konnte, welcher sieben interpretierbare Skalen enthält. Von diesen sieben Skalen konnten wiederum fünf durch die Operationalisierung gestützt werden. Diese können nun im Rahmen einer Validierungsstudie weitergehend untersucht werden. Für die Skalen *Liebesbeziehungen* und *Freizügige Selbstdarstellung* bedarf es einer genaueren Reanalyse sowie der Generierung weiterer Items, damit auch diese Skalen im Sinne der Operationalisierung interpretiert werden können. Außerdem konnten für die Skalen *Modelllernen* und *Minderwertigkeit* im Rahmen der Itemanalyse vielversprechende Ergebnisse hinsichtlich der Reliabilität und der Trennschärfen gewonnen werden. Daher sollten diese Themenbereiche nochmals untersucht und analysiert werden, um das volle Potenzial des Fragebogens auszuschöpfen.

4 Studie II

Dieser Teil der vorliegenden Arbeit beschreibt jegliche Inhalte der zweiten Studie, die primär der Validierung der revidierten Version des Fragebogens dient. Zunächst sollen die testtheoretischen Ansätze und relevante statistische Verfahren erläutert werden, um die Nachvollziehbarkeit der Vorgehensweise zu gewährleisten. Weiterhin wird auf die methodische Herangehensweise eingegangen. Überdies werden die Ergebnisse präsentiert und interpretiert.

4.1 Testtheoretischer Hintergrund – Studie II

Im Folgenden wird die testtheoretische Basis der zweiten Untersuchung weiter ausgeführt. Im Rahmen des Unterkapitels 2.2.5.3 Validität wurden die verschiedenen Formen

der Validität bereits ausführlich thematisiert. Entsprechend wird an dieser Stelle auf eine erneute Ausführung verzichtet. Stattdessen liegt der Fokus auf der Überprüfung der Validität und ihrer besonderen Relevanz in der psychometrische Diagnostik.

4.1.1 Bedeutung der Validität für die psychometrische Diagnostik

Die Validität gilt als das wichtigste Gütekriterium zur Beurteilung von Verfahren der psychometrischen Diagnostik. Schließlich kann die Validität im Hinblick auf die Testkonstruktion als das "Ausmaß, in dem empirische Befunde und theoretische Argumente, die die Interpretationen von Testwerten für die beabsichtigten Verwendungen von Tests unterstützen" (AERA, APA & NCME, 2014, S.11, zitiert nach Harting, Frey & Jude, 2020, S. 530), definiert werden. Entsprechend sind die Kriterien Reliabilität und Objektivität der Validität untergeordnet, da diese bei fehlender Validität keine relevante Bedeutung haben. Darüber hinaus ist die Validität das komplexeste Gütekriterium, weil eine vollumfängliche Überprüfung und Nachweis der Validität einen langfristigen sowie vielseitigen Untersuchungsprozess erfordert. Dieser umfasst verschiedene theoretische und empirische Einzelbefunde, die in einer integrierten Beurteilung zusammengeführt werden müssen (Hartig, Frey, & Jude, 2020).

4.1.2 Prüfung der Validität

Das nachfolgende Kapitel beschäftigt sich mit verschiedenen Ansätzen zur Überprüfung der Validität. Diese sollen in ihren Grundzügen beschrieben werden, um einen möglichst breit gefächerten Einblick in den komplexen Prozess der Validierung zu gewinnen. In erster Linie liegt der Fokus auf der Bestimmung der Kriteriumsvalidität sowie auf extrapolierende Testwertinterpretationen. Schließlich sollen diese im Rahmen dieser Untersuchung bestimmt und interpretiert werden. Entsprechend wird auf eine detaillierte Erläuterung aller Facetten der Validität und deren Prüfung, aufgrund des begrenzten Rahmens dieser Arbeit verzichtet.

4.1.2.1 Kriteriumsvalidität und extrapolierende Testwertinterpretationen

Die Kriteriumsvalidität bezieht sich auf die Interpretation und Anwendung von Testwerten auf das Verhalten einer Testperson außerhalb der Testsituation, sogenannte extrapolierende Testwertinterpretationen. Damit ist die Bestimmung der Kriteriumsvalidität darauf ausgerichtet, Bezüge und Zusammenhänge zwischen den Testergebnissen und diagnostisch relevanten Außenkriterien (z. B. Messwerten eines anderen Testverfahrens). Entsprechend lässt die Kriteriumsvalidität am leichtesten über die Feststellung einer Test-Kriterium-Korrelation bestimmen. Bei der Untersuchung der Zusammenhänge zwischen Testwert und Außenkriterien empfiehlt sich, diese vor dem inhalts-theoretischen Hintergrund der Testkonstruktion und dem erfassten Konstrukt abzuleiten. Entsprechend

sollten die Außenkriterien zur Validierung unmittelbare Relevanz aufweisen und hinreichend untersucht sein. Somit sollte bestenfalls die *Konstruktvalidität* (siehe dazu auch Kapitel *2.2.5.3 Validität*) im Hinblick auf das erfasste Konstrukt und die Dimensionalität genau untersucht sein (Pospeschill 2022; Moosbrugger & Kelava, 2020a).

Dennoch ist die Bestimmung und Überprüfung der Kriteriumsvalidität prinzipiell an keine expliziten testtheoretischen Annahmen gebunden. Dem zur Folge können Testwerte, deren inhalts-theoretisches Fundament noch nicht geklärt ist, dennoch eine empirisch nachgewiesene Kriteriumsvalidität aufweisen. Ein Beispiel liefern Goldhammer und Harting (2020) mit der kriteriumsorientierten Testwertinterpretation. Im Allgemeinen erfolgt hier die Testwertinterpretation nicht in Abhängigkeit von der Bezugsgruppe, sondern hinsichtlich eines bestimmten inhaltlichen Kriteriums (Goldhammer & Harting, 2020).

In der Praxis kann die Kriteriumsvalidität eines Testverfahrens natürlich nicht anhand einer einzigen Korrelation bestimmt werden. Wie bereits beschrieben bedarf es einer kontinuierlichen Zusammentragung sowie Integration von Einzelbefunden, um die Validität eines Verfahren hinreichend festzustellen. Dem zur Folge bestimmt das Ausmaß, in dem die Interpretationen durch theoretische Annahmen sowie empirische Belege untermauert sind, die Einschätzung der Validität (Moosbrugger & Kelava, 2020a)

4.1.2.2 Argumentationsbasierter Ansatz der Validierung

Der argumentationsbasierte Validierungsansatz von Testwertinterpretationen erfordert zunächst eine klare Definition, auf welche Interpretationen der Testergebnisse sich die Validierung beziehen soll. Entsprechend werden die zu validierenden Interpretationen als Grundannahmen genau formuliert, welche sich empirisch prüfen lassen müssen. Anschließend soll auf diese Annahmen empirische Evidenz gesammelt werden. Als wesentliche Quellen empirischer Evidenz gelten Testinhalte, die Struktur der Testwerte sowie die Beziehung zwischen Variablen und anderen Konstrukten. Außerdem dienen auch kognitive Prozesse, die bei der Testung ablaufen als empirische Evidenzquellen. In Abhängigkeit davon, ob die Annahmen gestützt oder widerlegt werden, können die Grundannahmen als valide betrachtet werden oder nicht. Diese wird in einer abschließenden und zusammenfassenden Beurteilung der Testwertinterpretationen ausführlich dargelegt (Hartig, Frey, & Jude, 2020; Moosbrugger & Kelava, 2020a).

4.1.3 Zielsetzung

Das Ziel dieser Studie ist es, eine erneute Prüfung der Skalen- und Itemkennwerte durchzuführen. Weiterhin sollen erste Befunde zur Bestimmung der Validität gesammelt werden. Im Zuge dessen sollen folgende Aspekte empirisch geprüft werden:

 Die Itemkennwerte und Reliabilitäten der Skalen des FBMSN-R werden erneut bestimmt, um sie mit den Ergebnissen der ersten Studie zu vergleichen.

- Die Reliabilitätskoeffizienten der Validierungsskalen werden erneut untersucht, um die postulierten Ergebnisse der Skalendokumentationen bestenfalls zu replizieren.
- Die Kriteriumsvalidität der Skalen des FBMSN-R soll anhand verschiedener Außenkriterien bestimmt werden, um erste Befunde für die Validierung des Fragebogens zu sammeln.

Zur Bestimmung der Kriteriumsvalidität des FBMSN-R sollen Zusammenhänge zwischen den einzelnen Skalen und anderen Variablen untersucht werden. Die Außenkriterien weisen verschiedene theoretische Beziehung zu den Interpretationen der einzelnen Faktoren auf. Entsprechend wurden verschiedene Außenkriterien in die Befragung integriert. Somit erfolgt die Bestimmung von Testwert sowie Kriteriumswert parallel bzw. zum selben Messzeitpunkt. Diese Variante der Kriteriumsvalidität wird als Übereinstimmungsvalidität oder auch kongruente Validität bezeichnet (Döring & Bortz, 2016).

4.1.3.1 Empirisch-Inhaltliche Hypothesen

Auf Basis der Interpretation und der Operationalisierung der einzelnen Skalen werden Hypothesen formuliert, um erste Befunde für die Kriteriumsvalidität zu liefern. Im Folgenden sind diese Hypothesen über die Zusammenhänge der einzelnen Subskalen mit bestimmten Außenkriterien inhaltlich dargestellt.

Auf Basis der in Abschnitt 3.4.2.1 Streben nach sozialer Anerkennung geschilderten Interpretation wird vermutet, dass diese Skala einen Zusammenhang zum Grundmotiv Leistung aufweisen sollte. So streben Menschen nach Fogg (2009) stets nach sozialer Anerkennung und der Bestätigung durch andere Personen, um ein positives Selbstwertgefühl aufrechtzuerhalten. Weiterhin versuchen Individuen einen positiven Eindruck zu hinterlassen, der zu sozialen Belohnungen (z. B. Komplimente, Einladungen zu gemeinsamen Aktivitäten) führt (Leary, 1995). Nach Satow (2020) suchen Personen mit einem ausgeprägten Leistungsmotiv ständig nach sozialer Anerkennung für ihre Leistungen. Entsprechend sollte sich dieses Streben nach sozialer Anerkennung auch auf soziale Netzwerke übertragen lassen. Für die Skala Streben nach sozialer Anerkennung wird daher die folgende Hypothese formuliert:

H₁: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Streben nach sozialer Anerkennung und dem Leistungsmotiv.

Im Rahmen der Interpretation in Abschnitt 3.4.2.2 Soziale Vergleiche zeigt sich, dass diese Skala verschiedene Zusammenhänge ausweist. Die aktuelle Forschung zeigt, dass der ständige Vergleichsdruck negative Auswirkungen auf das Wohlbefinden und den Selbstwert der Nutzer von sozialen Netzwerken hat (Steers, Wickham & Acitelli, 2014). Entsprechend sollte ein Zusammenhang mit dem globalen Selbstwertgefühl

bestehen. Daher wird eine Korrelation zwischen der Skala Soziale Vergleiche und der RSES vermutet (Collani & Herzberg, 2003):

H₂: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und dem globalen Selbstwertgefühl.

Des Weiteren deuten mittlerweile einige Studien darauf hin, dass es bei der Betrachtung von Profilen auf Social-Media-Plattformen zu sozialen Vergleichsprozessen nach Festinger (1954) kommt. In diesem Zusammenhang zeigen empirische Studien, dass die Verfügbarkeit an Informationen über Erfolge und Erlebnisse anderer Menschen negative Einflüsse auf die Lebenszufriedenheit haben können (Krämer, Eimler, & Neubaum, 2017). Weitere Studien zeigen, dass Studierende, die seit Längerem auf Facebook aktiv sind, eher glauben, dass andere Personen glücklicher sind als sie selbst (Chou & Edge, 2012). Daher wird die Skala *Neurotizismus* nach Satow (2020) hinzugezogen. Nach Satow (2020) neigen Menschen mit hohen Werten auf der Skala *Neurotizismus* dazu, an sich selbst zu zweifeln und sich selbst zu bemitleiden. Weiterhin neigen sie eher zu depressiven Symptomen (Satow, 2020). Entsprechend wird die folgende Hypothese formuliert:

H₃: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und Neurotizismus.

Dennoch weist Crusius (2022) darauf hin, dass die sozialen Aufwärtsvergleichsprozesse nach Festiger (1954) mit einer gesteigerten Leistungsmotivation einhergehen können. Entsprechend wird ebenfalls Zusammenhang mit dem Leistungsmotiv vermutet (Satow, 2020). Auf dieser Basis wird die folgende Hypothese formuliert:

H₄: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und dem Leistungsmotiv.

Auf Grundlage der Interpretation in Abschnitt 3.4.2.3 Konformitätsdruck und Gruppendynamik werden ebenfalls verschiedene Zusammenhänge vermutet. Zunächst bildet das Streben nach Konformität eine Facette des Anschlussmotivs. Entsprechend wird davon ausgegangen, dass ein Zusammenhang mit dem Bedürfnis nach Zugehörigkeit und Anschluss besteht (Krämer, Eimler, & Neubaum, 2017). Daher wird die Skala Anschluss und Kontaktbereitschaft des TEMEO-R zur Validierung hinzugezogen (Deutscher & Mees, 2012). Es resultiert die folgende Hypothese für den vermuteten Zusammenhang: H₅: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und Bedürfnis nach Zugehörigkeit und Anschluss.

Zur Präzisierung wird ebenfalls ein Zusammenhang zur *indirekten Annäherungsmotivation des Anschlussmotivs* vermutet. So ist diese nach Deutscher und Mees (2012) durch eine Instrumentalisierung des Kontaktes für ein positiv bewertetes übergeordnetes Ziel bestimmt, wie beispielsweise berufliche Vorteile oder, wie in diesem Fall, die

Zugehörigkeit und Anerkennung innerhalb einer sozialen Gruppe. Entsprechend wird eine untergeordnete Hypothese formuliert:

H_{5.1}: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der indirekten Annäherungsmotivation bezüglich Anschluss.

Außerdem muss auch hier davon ausgegangen werden, dass ein Zusammenhang zu der Tendenz sich in sozialen Situationen anderen Personen zu orientieren besteht. Folglich wird die nachfolgende Hypothese für die Skala *Konformitätsdruck und Gruppendynamik* aufgestellt:

H₆: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der Orientierung an Anderen in sozialen Situationen.

Ferner wird vermutet, dass Personen sich dem Konformitätsdruck ergeben und ebenfalls versuchen, durch konformes Verhalten negative Bewertungen durch andere Personen zu vermeiden. Daher wird die folgende Hypothese formuliert:

H₇: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der Skala Angst vor negativer Bewertung.

Für die Skala Selbstkonzept und Selbstwert bilden die Items in erster Linie negative Einflüsse und Auswirkungen auf das Selbstkonzept durch Interaktionen auf Social Media ab. Hier deuten die Ergebnisse von Sumter, Baumgartner, Valkenburg und Peter (2012) darauf hin, dass negative Interaktionen und Feedback die Lebenszufriedenheit langfristig negativ beeinflussen. Entsprechend wird ein Zusammenhang zum globalen Selbstwertgefühl vermutet:

H₈: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und dem globalen Selbstwertgefühl.

Außerdem lässt sich für diese Skala ein Zusammenhang zu einer Angst vor negativer Bewertung annehmen. Entsprechend wird die Skala SANB-5 zur Bestimmung der Bewertungsängstlichkeit hinzugezogen (Kemper, Lutz, & Neuser, 2011) und die folgende Hypothese abgeleitet:

*H*₉: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und der Angst vor negativer Bewertung.

Ebenso wird eine erhöhe Verletzlichkeit im Hinblick auf negatives Feedback auf sozialen Netzwerken vermutet. In diesem Zusammenhang geht aus aktuellen Forschungen hervor, dass negative Rückmeldungen aus sozialen Netzwerken zu Wut, Scham, Hilflosigkeit und Depressionen führen können (Ortega, Paz Elipe, Mora-Merchán, & Vega, 2009). Daher wird ebenfalls von einem Zusammenhang mit der Persönlichkeitsdimension Neurotizismus ausgegangen (Satow, 2020):

H₁₀: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und Neurotizismus.

Auf Basis der Interpretation im Abschnitt 3.4.2.4 Anschlussmotiv wird vermutet, dass ein besonders Streben nach Anschluss und Kontakt sich auch im Zusammenhang mit sozialen Netzwerken nachweisen lassen muss. Entsprechend wird auch hier ein Zusammenhang zur Skala Anschluss und Kontaktbereitschaft des TEMEO-R vermutet (Deutscher & Mees, 2012). Somit wird die folgende Hypothese aufgestellt:

H₁₁: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Anschlussmotiv und dem Bedürfnis nach Anschluss und Kontakt.

Außerdem lässt sich annehmen, dass ein erhöhtes Streben nach Anschluss und Kontakt auf sozialen Netzwerken mit der Persönlichkeitsdimension *Extraversion* zusammenhängt. Diese zeichnet sich nach Satow (2020) unter anderem durch Geselligkeit und Kontaktfreudigkeit aus. Im Hinblick auf die Skala *Anschlussmotiv* wird daher der folgende Zusammenhang vermutet und durch die entsprechende Hypothese abgebildet:

H₁₂: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Anschlussmotiv und Extraversion.

An dieser Stelle wird darauf hingewiesen, dass die Skalen *Liebesbeziehungen* und *Freizügige Selbstdarstellung* nicht in den Validierungsprozess einbezogen werden, da diese nicht das vermutete Intimitätsmotiv abbilden.

4.1.3.2 Statistische Hypothesen

Nachfolgenden sind die zuvor inhaltlich formulierten Hypothesen in statistische Hypothesen übersetzt.

H ₁ : ſ Streben nach soziale Anerkennung; B5-T: Leistungsmotiv	H ₀ : r Streben nach soziale Anerkennung; B5-T: Leistungsmotiv
<i>≠</i> 0	= 0
H ₂ : r _{Soziale Vergleiche; RSES ≠ 0}	H ₀ : r Soziale Vergleiche; RSES = 0
H ₃ : r _{Soziale} Vergleiche; B5-T: Neurotizismus ≠ 0	H_0 : r Soziale Vergleiche; B5-T: Neurotizismus = 0
H ₄ : r _{Soziale} Vergleiche; B5-T: Leistungsmotiv ≠ 0	H ₀ : r Soziale Vergleiche; B5-T: Leistungsmotiv = 0
H ₅ : r Konformitätsdruck und Gruppendynamik; TEMEO-R: An-	Ho: r Konformitätsdruck und Gruppendynamik; TEMEO-R: An-
schluss & Kontaktbereitschaft gesamt ≠ 0	schluss & Kontaktbereitschaft gesamt = 0
$\pmb{\mathcal{H}_{5.1}}$: r Konformitätsdruck und Gruppendynamik; TEMEO-R:	Ho: r Konformitätsdruck und Gruppendynamik; TEMEO-R: An-
Anschluss & Kontaktbereitschaft Annäherung indirekt $ eq 0$	schluss & Kontaktbereitschaft Annäherung indirekt $=0$
H ₆ : I' Konformitätsdruck und Gruppendynamik; SMSC: Orien-	$oldsymbol{\mathcal{H}_0}$: r Konformitätsdruck und Gruppendynamik; SMSC: Orien-
tierung an Anderen $ eq 0$	tierung an Anderen = 0
H ₇ : r Konformitätsdruck und Gruppendynamik; SANB-5 $\neq 0$	H ₀ : r Konformitätsdruck und Gruppendynamik; SANB-5 = 0
H₈: $r_{\text{Selbstkonzept und Selbstwert; RSES}} \neq 0$	H_0 : r Selbstkonzept und Selbstwert; $RSES = 0$
H 9: r Selbstkonzept und Selbstwert; SANB-5 ≠ 0	H ₀ : r Selbstkonzept und Selbstwert; SANB-5 = 0
H ₁₀ : r Selbstkonzept und Selbstwert; B5-T: Neurotizismus	H_0 : r Selbstkonzept und Selbstwert; B5-T: Neurotizismus = 0

≠ 0

H11: 'r Anschlussmotiv: TEMEO-R: Anschluss & Kontaktbereit-

schaft ≠ 0

schaft = 0

H₁₂: $r_{Anschlussmotiv: B5-T: Extraversion} \neq 0$

H₀: $r_{Anschlussmotiv: B5-T: Extraversion} = 0$

4.2 Methode - Studie II

Das nachfolgende Kapitel bildet die methodische Untersuchung und empirische Überprüfung auf Grundlage des beschriebenen testtheoretischen Fundaments für die zweite Studie ab. Im Zuge dessen werden zunächst das Studiendesign, Informationen zur Stichprobenrekrutierung sowie die Datenerhebung dargestellt und das verwendete Messinstrument beschrieben. Abschließend sollen die geplanten statistischen Analysen konkret dargelegt werden.

4.2.1 Forschungsdesign

Das Forschungsdesign bildet auch hier den maßgeblichen Rahmen um die Untersuchung. Entsprechend wirkt sich das Design direkt auf den Ablauf der Studie aus. Die vorliegende Untersuchung basiert auf einem nicht-experimentellen, quantitativen Forschungsansatz. Für die Validierung der revidierten Version des Fragebogens (FBMSN-R) wurde eine querschnittliche Befragung anhand einer Validierungsstichprobe durchgeführt. Im Zuge dessen wurde der revidierte Fragebogen inklusive der Validierungsskalen einmalig erhoben (Richter, Paier, & Reiger, 2021; Eid, Gollwitzer, & Schmitt, 2017). Der Erhebungszeitraum startete am 16. September 2022 und endete am 10. Oktober 2022.

4.2.2 Anforderungen an die Stichprobe und Rekrutierung

Die Validierung des FBMSN-R erfolgt anhand einer Validierungsstichprobe, welche aufgrund der zeitlichen und finanziellen Grenzen dieser Arbeit, lediglich einer Gelegenheitsstichprobe entspricht. Bezüglich des Stichprobenumfangs wird eine Größe von n = 100angepeilt, um auf eine angemessene Datenbasis für die Validierung zu gewährleisten. Als Einschlusskriterium wird die Nutzung sozialer Netzwerke festgelegt. Entsprechend sollen ausschließlich Personen für die Stichprobe rekrutiert werden, die zum Zeitpunkt der Erhebung über mindestens einen Social Media Account verfügen und diesen regelmäßig nutzen. Überdies sollte die Stichprobe eine ausgeglichene Geschlechterverteilung aufweisen und eine Altersspanne von 18 bis 40 Jahren nicht überschreiten. Zur Akquirierung der Teilnehmer bietet sich eine Kontaktierung über verschiedene Social-Media-Kanäle sowie E-Mail-Kontakte an, um eine große Anzahl potenzieller Probanden schnell und kostenfrei anzusprechen. Weiterhin wird die Umfrage im digitalen Versuchspersonenstundensystem *Sona* der Hochschule Fresenius eingebunden, um weitere Teilnehmer zu gewinnen, diese können sich für ihre Teilnahme eine Versuchspersonenstunde anrechnen lassen. Außerdem wird der Fragebogen auf den Plattformen *SurveyCircle* und *Thesius* geteilt, um die Untersuchung über möglichst viele verschiedene Kanäle zu verbreiten. Ergänzend werden drei Amazon-Gutscheinkarten unter den Teilnehmenden verlost, um einen zusätzlichen Anreiz zu bieten.

4.2.3 Messinstrumente

Nachfolgend sollen die verwendeten Messinstrumente beschrieben werden. Dabei liegt der Fokus hauptsächlich auf den einzelnen Testverfahren und Skalen, die zur Validierung des FBMSN-R verwendet werden sollen.

4.2.3.1 Revidierte Version des Fragebogens zur Erfassung der zugrundeliegenden Motive der Selbstdarstellung auf sozialen Medien (FBMSN-R)

Alle inhaltlichen sowie formalen Aspekte des FBMSN-R wurden bereits ausführlich im Rahmen der ersten Studie thematisiert und vorgestellt. Daher wird an dieser Stelle auf eine ausführliche Beschreibung des FBMSN-R verzichtet. Die relevanten Kennwerte können im Kapitel 3.3.7 Skalenzusammensetzung und Kennwerte des revidierten Fragebogens nachgelesen oder auch dem Anhang entnommen werden. Bezüglich der Fragen- und Antwortformulierung sowie der Skalierung der Antworten wurden keine Veränderungen vorgenommen. So wurde der formale Aufbau von Burdorf und Lux (2022) beibehalten. Entsprechend stützen sich die Fragen- und Antwortformulierungen auf verschiedene Ich-Aussagen, die eine Meinung oder Einstellung abbilden. Weiterhin werden die Teilnehmer instruiert, den Grad der eigenen Zustimmung auf einer sechsstufigen Likert-Skala mit verbalen Ankern (Stimme überhaupt nicht zu, Stimme nicht zu, Stimme eher zu, Stimme zu, Stimme völlig zu) anzugeben. Die Einstufung erfolgt von 1 (Stimme überhaupt nicht zu) bis 6 (Stimme völlig zu). Die Auswertung erfolgt über Summenwerte der einzelnen Skalen (Burdorf & Lux, 2022).

4.2.3.2 B5-T: Big-Five-Persönlichkeitstest

Der B5-T: Big-Five-Persönlichkeitstest erfasst, neben den grundlegenden Persönlichkeitsdimensionen Neurotizismus (N), Extraversion (E), Gewissenhaftigkeit (C) sowie Verträglichkeit (A) und Offenheit (O), auch die drei Grundmotive Leistungsmotiv (LM) sowie Machtmotiv (MM) und Sicherheitsmotiv (SM). Weiterhin überzeugt der Test sowohl durch seine hohe Skalen-Reliabilität als auch durch seine faktorielle Validität. Zur Ermittlung der Persönlichkeits- und Motivstruktur werden dem Probanden Ich-Aussagen präsentiert, welche anhand einer vierstufigen Likert-Skala mit verbalen Ankern bewertet werden sollen. Die Probanden sollen einstufen, ob die Aussagen auf sie persönlich

zutreffen oder nicht. Die Einstufung erfolgt von 1 ("trifft gar nicht zu") bis 4 ("trifft genau zu") (Satow, 2020). Für die Validierung des FBMSN-R werden die Subskalen *Neurotizismus* sowie *Extraversion* verwendetet. Außerdem wird die Skala *Leistungsmotiv* ebenfalls für die Bestimmung der Validität der Skalen des FBMSN-R herangezogen. Diese sollen im Folgenden kurz vorgestellt werden.

Die Big-Five-Skala *Neurotizismus* umfasst zehn Items und weist mit einer Reliabilität von Cronbachs α = .90 einen ausgezeichneten Wert für die Testgenauigkeit auf. Die Skala *Extraversion* setzt sich ebenfalls aus zehn Items zusammen und erreicht einen ähnlich hohen Wert für die Reliabilität (Cronbachs α = .86). Weiterhin besteht die Motivskala *Leistungsmotiv* aus sechs Items und erreicht auf dem Reliabilitätskoeffizienten Cronbachs α einen ebenfalls sehr guten Wert von .81 (Satow, 2020).

4.2.3.3 TEMEO-R: Fragebogen zur Erfassung telischer und metatelischer Orientierungen in revidierter Form

Der TEMEO-R erfasst insgesamt elf verschiedene Motivbereiche, welche jeweils anhand von vier Skalen zur Erfassung der metatelischen Orientierung (direkte sowie indirekte Annäherungs- und direkte sowie indirekte Vermeidungsmotivation) bestimmt werden. Weiterhin ergibt sich die Gesamtstärke des Motivs aus der Summe der jeweiligen direkten sowie indirekten Annäherungs- und Vermeidungsmotivation. Den Testpersonen werden für die Bestimmung der Skalenwerte selbstbezogene Aussagen präsentiert, welche anhand einer fünfstufigen Likert-Skala mit verbalen Ankern (sehr unzutreffend, unzutreffend, weder noch, zutreffend, sehr zutreffend) bewertet werden sollen. Die Probanden sollen einstufen, ob die Aussagen auf sie persönlich zutreffen oder nicht. Eine Einstufung erfolgt von 1 ("sehr unzutreffend") bis 5 ("sehr zutreffend"). Die Skalendokumentation liefern ausreichende Informationen bezüglich der Kriteriums- und Inhaltsvalidität. Im Hinblick auf die Reliabilität werden akzeptable Werte für die Gesamtskalen der Motivbereiche berichtet. (Deutscher & Mees, 2012). Für die Validierung des FBMSN-R wird die Skala Anschluss und Kontaktbereitschaft des TEMEO-R herangezogen. Diese erreicht mit Cronbachs α = .717 einen guten Wert für die Reliabilität (Deutscher & Mees, 2012).

4.2.3.4 RSES: Rosenberg Self-Esteem-Scale (deutsche Fassung)

Die Rosenberg Self-Esteem-Scale erfasst das globale Selbstwertgefühl anhand von 10 Items, wobei die Items 2, 5, 6, 8 und 9 negativ formuliert sind und entsprechend umkodiert werden müssen. Das globale Selbstwertgefühl repräsentiert eine wesentliche Komponente der allgemeinen Lebenszufriedenheit und gilt als wichtiger Indikator für die psychische Gesundheit. Die Bestimmung des Testwertes erfolgt durch eine Selbstbeurteilung auf den einzelnen Items. Für jedes Item werden 3 Punkte ("stimmt genau"), 2

Punkte ("stimmt überwiegend"), 1 Punkt ("stimmt etwas") oder 0 Punkte ("stimmt gar nicht") vergeben. Die Skalenwerte werden durch Aufsummierung der Itemwerte gebildet. Bezüglich der Validität werden hinreichende Belege geliefert. Außerdem bewegt sich die mittlere interne Konsistenz einem sehr guten Bereich von .83 und .88 (Collani & Herzberg, 2003).

4.2.3.5 SANB-5: Skala Angst vor negativer Bewertung-5

Die Skala Angst vor negativer Bewertung-5 ist ein Selbstbeurteilungsinstrument zur Messung der dispositionellen Angst vor negativer Bewertung (Bewertungsängstlichkeit). Die Skala enthält fünf Items mit vierstufigem Antwortformat von 1 (trifft fast nie zu) bis 4 (trifft fast immer zu), die zu einem Messwert aggregiert werden. Aus der Skalendokumentation gehen hinreichende Belege für die Konstruktvalidität der Skala hervor. Unabhängig von der betrachteten Stichprobe wird für die SANB-5 ein hoher Wert für die Reliabilität erreicht (Cronbachs Alpha zwischen .84 und .94) (Kemper, Lutz, & Neuser, 2011).

4.2.3.6 SMSC: Fragebogen zum Verhalten in sozialen Situationen

Die SMSC erfasst den Grad des Self-Monitorings einer Person. Unter Self-Monitoring wird die Selbstdarstellung einer Person, um einen bestimmten Eindruck bei einem sozialen Umfeld zu kreieren, verstanden. Der Test besteht aus vier Subskalen, welche wie folgt beschrieben sind: (A) Selbstdarstellung vor Publikum (Acting; 11 Items), (O) Orientierung an Anderen (Other directedness; 9 Items) sowie (S) Sensitivität für expressives Verhalten und soziale Hinweisreize (Sensitivity to the reactions of others; 8 Items) und (E) Extraversion (7 Items). Jedes Item wird durch die dichotomen Antwortmöglichkeiten "Trifft zu" und "Trifft nicht zu" bestimmt. Aus dem Testmanual liefert Belege für die Kriteriumsvalidität (Collani & Stürmer, 2009).

Für die Validierungsuntersuchung des FBMSN-R wird lediglich die Subskala (O) Orientierung an Anderen herangezogen. Diese weist eine akzeptable innere Konsistenz von Cronbachs α = .60 auf (Collani & Stürmer, 2009).

4.2.4 Datenerhebung

Die zweite Studie gewinnt ihre Informationen ebenfalls auf Basis von Selbstberichtsdaten. Dazu wurde ein neuer Fragebogen über die Online-Plattform *Unipark* konzipiert. Dieser unterscheidet sich hinsichtlich des formalen Rahmens nur geringfügig von dem Fragebogen, welcher bereits in der ersten Studie konzipiert wurde (siehe Kapitel 3.2.4 Datenerhebung). Daher wird auf eine ausführliche Präsentation des Fragebogens verzichtet und lediglich auf die relevanten Unterschiede eingegangen.

Im Wesentlichen wurden sämtliche Informationen zum Datenschutz und ähnliche Formalitäten im Hinblick auf die Teilnahme beibehalten. Weiterhin wurden sowohl die

Filterfrage als auch die Abfrage der soziodemografischen Daten und die Fragen zum individuellen Nutzungsverhalten von sozialen Netzwerken nicht verändert. Im Unterschied zum ersten Fragebogen beginnt anschließend die Erhebung der revidierten Version des Fragebogens. Außerdem ist der Fragebogen durch die ausgewählten Validierungsskalen, welche im vorangegangen Kapitel ausführlich beschrieben wurden, ergänzt worden. Am Ende der Befragung kann sich der Proband analog zur ersten Studie für das Gewinnspiel anmelden sowie sich zum digitalen Versuchspersonenstundensystem *Sona* der Hochschule Fresenius weiterleiten lassen. An dieser Stelle ist festzuhalten, dass der FBMSN-R inklusive der Validierungsskalen deutlich mehr Zeit in Anspruch nimmt (mindestens 45 – 60 Minuten). Auf diesen Umstand wird auf einer der ersten Seiten verwiesen. Der vollständige Fragebogen inklusive der beschriebenen Validierungsskalen ist dem Anhang A zu entnehmen.

4.2.5 Statistische Analysen

Die statistische Auswertung der zweiten Studie setzt sich im Wesentlichen aus zwei Schritten zusammen. Zunächst erfolgt eine Reliabilitätsanalyse für die einzelnen Subtests. Bestenfalls können an dieser Stelle die Ergebnisse aus Studie I bzw. aus den jeweiligen Testmanualen der Validierungsskalen bestätigt und repliziert werden. Im zweiten Schritt wird die Kriteriumsvalidität der einzelnen Subskalen des Fragebogens anhand von Zusammenhangsmaßen geprüft und die zuvor formulierten Hypothesen getestet. Dazu werden die entsprechenden bivariaten Korrelationen betrachtet. Alle Berechnungen und Aufbereitungen der Daten wurden mithilfe der Software *IBM SPSS Statistics* durchgeführt.

4.3 Ergebnisse – Studie II

Dem nachfolgenden Kapitel sind die Ergebnisse der zweiten Studie zu entnehmen. Die nachfolgenden Kapitel behandeln die Zusammensetzung der Stichprobe, die Analyse der Itemkennwerte und der Reliabilitätskoeffizienten sowie die Bestimmung der Kriteriumsvalidität der Skalen des FBMSN-R. Sämtliche Ergebnisse der statistischen Auswertung dieser Untersuchung sind im Anhang C hinterlegt.

4.3.1 Zusammensetzung der Stichprobe

In der 25-tägigen Erhebungsphase nahmen insgesamt 149 Personen an der Befragung teil. Davon beendeten 112 Teilnehmer alle Fragen vollständig. So ergibt sich eine Abbruchquote von ca. 24 %. Alle unvollständigen Fragebögen wurden nicht in die weiteren Berechnungen und deskriptiven Statistiken einbezogen. Entsprechend umfasst die betrachtete Stichprobe 112 Versuchspersonen (*N*).

Bei Betrachtung der Häufigkeitsverteilung des Geschlechts zeigt sich, dass 62,5 % (n = 70) der Probanden weiblichen und 37,5 % (n = 42) männlichen Geschlechts sind (siehe Abbildung 7). Im Durchschnitt sind die Testpersonen 22,69 Jahre alt, bei einer Standardabweichung von 3,73 Jahren. Bezüglich der Altersspanne umfasst die Stichprobe einen Bereich von 18 bis 38 und beträgt entsprechend 20 Jahre. Die

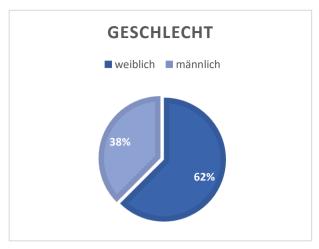


Abb. 7: Verteilung nach Geschlecht - Studie II

Hälfte der Probanden (50 %) sind zwischen 21 und 24 Jahre alt. Das Verhältnis aus Singles und Personen, die sich zum Zeitpunkt der Erhebung in einer Beziehung befinden, ist sehr ausgeglichen (51 % Single; 49 % in einer Beziehung). Auffällig ist weiterhin, dass 74 % das Abitur als höchsten Bildungsabschluss vorweisen können (n = 83). Entsprechend ist daher nicht überraschend, dass 73 % der Probanden sich zum Zeitpunkt der Datenerfassung im Studium befinden (n = 82). Davon gehen 33 % einer Werkstudententätigkeit nach. Außerdem zeigt die Analyse der soziodemografischen Daten, dass 53,6 % (n = 60) der Befragten sich keiner Religion zugehörig fühlen oder diese praktizieren.

Das folgende Säulendiagramm (siehe Abb. 8) bietet eine Übersicht über die relevanten Social-Media-Plattformen. Es vergleicht die einzelnen Plattformen hinsichtlich ihrer Nutzerzahlen innerhalb der Stichprobe. Die meisten Teilnehmer nutzen Instagram (86,6 %; n = 97). Auffällig ist vor allem, dass lediglich ein Viertel der Teilnehmer ein Facebook-Profil haben (25 %; n = 28). Zusätzlich ist die noch relativ junge Plattform TikTok mit 51,8 % (n = 100

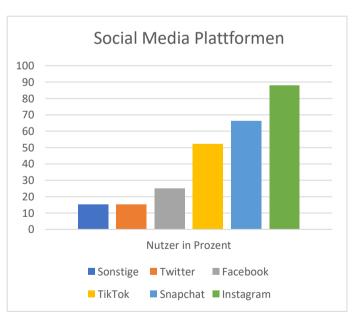


Abb. 8: Nutzung von Social Media Plattformen – Studie II.

58) sehr stark repräsentiert. In Bezug auf die tägliche Verweildauer kann festgehalten werden, dass 56 % der Teilnehmer (n = 63) zwei bis vier Stunden täglich auf verschiedenen Social-Media-Plattformen aktiv sind.

4.3.2 Itemkennwerte des FBMSN-R

Im Rahmen dieser Validierungsstudie wurden die Itemkennwerte des FBMSN-R erneut untersucht. Im Zuge dessen liegt der Fokus primär auf den Trennschärfen. Die folgende Tabelle bietet eine Übersicht über die Itemkennwerte der Skalen des FBMSN-R im Vergleich zu den Ergebnissen für die Trennschärfekoeffizienten der ersten Studie. Alle weiteren Ergebnisse finden sich im Anhang C.

Tab. 9: Itemanalyse Studie II inklusive der Trennschärfekoeffizienten aus Studie I

		\overline{X}	SD	Var(x)	r _{it} (Studie I)	r it	Pi
Streben nach so-	SA_01	3,63	1,363	1,858	,76	,62	52,6
zialer Anerken-	SA_02	3,06	1,490	2,221	,73	,80	41,2
nung	SA_03	3,82	1,310	1,716	,53	,74	56,4
	SA_04	4,42	1,373	1,885	,64	,53	68,4
	SA_05	3,38	1,538	2,365	,72	,66	47,6
	SA_06	2,75	1,516	2,297	,67	,75	35
	SA_08	3,79	1,525	2,327	,62	,53	55,8
	AM_11	2,98	1,439	2,072	,64	,70	39,6
		\overline{X}	SD	Var(x)	r _{it} (Studie I)	r _{it}	Pi
Sozialer Verglei-	SV_01	2,87	1,527	2,333	,59	,72	37,4
che	SV_05	2,56	1,374	1,888	,78	,80	31,2
	SV_06	2,22	1,347	1,815	,76	,73	24,4
	SV_07	2,68	1,396	1,950	,73	,80	33,6
	SV_08	2,37	1,294	1,676	,73	,70	27,4
	AM_02	2,50	1,336	1,784	,61	,53	30
		X	SD	Var(x)	r _{it} (Studie I)	r _{it}	Pi
Konformitätsdruck	GD_01	2,54	1,208	1,458	,54	,54	30,8
und Gruppendy-	GD_02	2,10	1,139	1,297	,71	,65	22
namik	GD_03	1,77	1,057	1,117	,74	,69	15,4
	GD_05	1,97	1,166	1,360	,69	,69	19,4
	GD_06	2,96	1,565	2,449	,54	,29	39,2
	GD_07	1,94	1,225	1,501	,61	,57	18,8
	GD_08	1,51	,977	,955	,63	,63	10,2
		\overline{X}	SD	Var(x)	r _{it} (Studie I)	r _{it}	Pi
Anschlussmotiv	AM_03	2,87	1,430	2,045	,54	,48	37,4
	AM_04	2,37	1,356	1,838	,54	,46	27,4
	AM_06	3,13	1,574	2,477	,48	,39	42,6
	AM_08	2,45	1,334	1,781	,53	,60	29
	AM_09	2,84	1,498	2,244	,60	,61	36,8
	AM_14	2,63	1,389	1,930	,65	,60	32,6
	AM_15	2,51	1,427	2,036	,65	,67	30,2

		X	SD	Var(x)	r _{it} (Studie I)	r it	Pi
Selbstkonzept	SK_01	2,90	1,554	2,414	,73	,65	38
und Selbstwert	SK_02	3,11	1,515	2,295	,60	,61	42,2
	SK_03	3,76	1,526	2,329	,48	,53	55,2
	SK_04	2,58	1,431	2,048	,77	,76	31,6
	AM_13	2,28	1,310	1,715	,53	,28	25,6
	SV_04	3,98	1,470	2,162	,61	,44	59,6
		\overline{X}	SD	Var(x)	r _{it} (Studie I)	r it	Pi
Liebesbeziehun-	IM_04	3,54	1,874	3,512	,71	,40	50,8
gen	IM_05	2,81	1,789	3,199	,76	,57	36,2
	IM_06	1,91	1,220	1,487	,59	,40	18,2
Freizügige Selbst-	IM_01	1,88	1,191	1,419	,75	,74	17,6
darstellung	IM_02	1,72	1,084	1,175	,74	,73	14,4
	IM_07	1,77	1,107	1,225	,59	,65	15,4

Anmerkungen: Mittelwert (\bar{x}), Standardabweichung (SD), Varianz (Var(x)), Trennschärfe (r_{it}), Schwierigkeitsindex (P_i)

Nach Betrachtung der Trennschärfekoeffizienten zeigt sich, dass die erreichten Werte zum Teil sehr stark von den Ergebnissen aus der ersten Studie abweichen. Auffällig sind vordergründig die niedrigen Werte für die Items *GD_06*, *AM_06* und *SV_04*. Alle anderen Werte bewegen sich in einem "guten" bis "sehr guten" Bereich (Pospeschill, 2022).

4.3.3 Reliabilitätsanalyse

Nach der Durchführung wurde für alle erhobenen Skalen die Reliabilität geprüft. Zur Bestimmung wurde der Reliabilitätskoeffizient Cronbachs Alpha (α) herangezogen. Im Zuge dessen wurden die Ergebnisse mit den Werten aus Studie I verglichen. Außerdem sind auch die Reliabilitätskoeffizienten aus den Skalendokumentationen hinzugezogen worden. Die nachfolgende Tabelle stellt die Ergebnisse übersichtlich dar.

Tab. 10: Übersicht und Vergleich der Reliabilitätskoeffizienten

Reliabilitätskoeffizienten (α)

	α (Studie I bzw. Ska- lendokumentation)	α (Studie II)
Streben nach sozialer Anerkennung	.89	.89
Soziale Vergleiche	.88	.89
Konformitätsdruck und Gruppendynamik	.86	.82
Anschlussmotiv	.82	.81
Selbstkonzept und Selbstwert	.84	.79
Liebesbeziehungen	.82	.63
Freizügige Selbstdarstellung	.83	.84
B5-T: Neurotizismus	.90	.88

B5-T: Extraversion	.86	.84
B5-T: Leistungsmotiv	.81	.74
SMSC: Orientierung an Anderen	.60	.45
RSES	.8388	.89
SANB-5	.8494	.87
TEMEO-R: Anschluss & Kontaktbereitschaft	.72	.67

Die Tabelle zeigt zum Teil deutliche Abweichungen der Koeffizienten. Auffällig sind vor allem die niedrigen Werte aus dieser Erhebung für die Skala *Liebesbeziehungen* und *SMSC: Orientierung an Anderen*. Ebenso zeigt die Skala des B5-T, welche das Leistungsmotiv abbilden soll, eine relativ starke Abweichung vom berichteten Wert. Nichtsdestotrotz erreichen alle Skalen, mit Ausnahme der Skala *SMSC: Orientierung an Anderen* einen zumindest "akzeptablen" Wert von .60 (Bühner, 2011).

4.3.3 Auswertung der Hypothesen zur Bestimmung der Kriteriumsvalidität

Im Rahmen des nachfolgenden Kapitels liegt der Fokus auf der Prüfung der Hypothesen, welche im Hinblick auf die Zusammenhänge der Subskalen des FBMSN-R mit bestimmten Außenkriterien formuliert wurden. Zunächst wird die Vorgehensweise bei der Untersuchung der Hypothesen kurz dargestellt. Anschließend werden die statistischen Daten der Erhebung analysiert und ausgewertet. Abschließend werden die Befunde zur Quantifizierung der Kriteriumsvalidität zusammenfassend und übersichtlich zusammengetragen.

4.3.3.1 Darstellung der Vorgehensweise

Um die vorliegenden Hypothesen zu prüfen, wurden bivariate Korrelationen berechnet. Im Zuge dessen wurde jede Subskala des FBMSN-R mit der Skala, die das entsprechende Außenkriterium repräsentiert, korreliert. Dazu wird jeweils die Korrelationsmatrix für die Subskala und die Kriterien angegeben. Anschließend werden die Ergebnisse im Hinblick auf die Hypothesen ausgewertet. Um die untersuchten Kriterien miteinander in Beziehung zu setzen und eine Produkt-Moment-Korrelation nach Pearson bestimmen zu können, wird zunächst eine z-Standardisierung der Variablen vorgenommen (Eid, Gollwitzer, & Schmitt, 2017).

4.3.3.2 Korrelationsanalyse für die Skala Streben nach Sozialer Anerkennung

H₁: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Streben nach sozialer Anerkennung und dem Leistungsmotiv.

Zur Prüfung dieser Hypothese zeigt die nachfolgende Tabelle die Korrelationsmatrix für die Skala Streben nach Sozialer Anerkennung und dem Leistungsmotiv.

Tab. 11: Korrelationsmatrix für die Skala Streben nach Sozialer Anerkennung

Korrelationen

		Z_SA	Z_Leistung_ges
Z_SA	r	1	
	Sig. (2-seitig)		
	N	112	
Z_Leistung_ges	r	,525**	1
	Sig. (2-seitig)	<,001	
	N	112	112

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Anmerkungen: r = Korrelationskoeffizient nach Pearson, Z_SA = Streben nach Sozialer Anerkennung, Z_Leistung_ges = Leistungsmotiv.

Auf Basis der in Tabelle 11 dargestellten Ergebnisse lässt sich die Hypothese H_1 bestätigen, da ein signifikant positiver linearer Zusammenhang zwischen der Skala *Streben nach sozialer Anerkennung* und dem Leistungsmotiv besteht. Der Korrelationskoeffizienten (r = .525) deutet nach Cohen (1988) auf einen starken Zusammenhang hin.

4.3.3.3 Korrelationsanalyse für die Skala Soziale Vergleiche

H₂: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und dem globalen Selbstwertgefühl.

*H*₃: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und Neurotizismus.

*H*₄: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und dem Leistungsmotiv.

Um die vorliegenden Hypothesen zu untersuchen, wird in der nachfolgenden Tabelle die Korrelationsmatrix für die Skala *Soziale Vergleiche* dargestellt und die Ergebnisse entsprechend interpretiert.

Tab. 12: Korrelationsmatrix für die Skala Soziale Vergleiche

Korrelationen

		Z_SV	Z_RSES	Z_Neuro_ges	Z_Leistung_ges
Z_SV	r	1			
	Sig. (2-seitig)				
	N	112			
Z_RSES	r	-,430 ^{**}	1		
	Sig. (2-seitig)	<,001			
	N	112	112		
Z_Neuro_ges	r	,579**	-,644**	1	
	Sig. (2-seitig)	<,001	<,001		
	N	112	112	112	

Z_Leistung_ges	r	,500**	-,294**	,382**	1
	Sig. (2-seitig)	<,001	,002	<,001	
	N	112	112	112	112

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Anmerkungen: r = Korrelationskoeffizient nach Pearson, Z_SV = Soziale Vergleiche, Z_RSES = Rosenberg Self-Esteem-Scale, Z_Neuro_ges = Neurotizismus, Z_Leistung_ges = Leistungs-motiv.

Bei Betrachtung der Korrelationskoeffizienten für die Skala *Soziale Vergleiche* (siehe Tab. 12) lässt sich feststellen, dass alle vier Hypothesen im Hinblick auf die Zusammenhänge zwischen der Skala *Soziale Vergleiche* und den Außenkriterien bestätigt werden können. Der Korrelationskoeffizient für die Hypothese H_2 (r = -.430) deutet auf einen mittleren negativen Zusammenhang hin (Cohen, 1988). Weiterhin ist die Korrelation signifikant geworden. Entsprechend kann ein signifikanter linearer Zusammenhang beobachtet werden. Auch die Hypothese H_3 wird auf Basis des Korrelationskoeffizienten (r = .579) angenommen. Entsprechend besteht ein starker signifikanter positiver linearer Zusammenhang zwischen der Skala *Soziale Vergleiche* und Neurotizismus. Des Weiteren lässt sich im Hinblick auf die Hypothese H_4 festhalten, dass eine signifikante positive lineare Korrelation (r = .500) für die Skala *Soziale Vergleiche* und Leistungsmotiv festzustellen ist. Nach Cohen (1988) kann auch dieser Zusammenhang als Mittel bis Stark beschrieben werden.

4.3.3.4 Korrelationsanalyse für die Skala Konformitätsdruck und Gruppendynamik

H₅: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und Bedürfnis nach Zugehörigkeit und Anschluss.

H_{5.1}: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der indirekten Annäherungsmotivation bezüglich Anschluss.

Die formulierten Hypothesen lassen sich anhand der Korrelationsmatrix, welche in der folgenden Tabelle angegeben ist prüfen. Dazu werden die entsprechenden Korrelationskoeffizienten betrachtet und interpretiert.

Tab. 13: Korrelationsmatrix für die Skala Konformitätsdruck und Gruppendynamik und die Skala Anschluss & Kontaktbereitschaft (TEMEO-R)

Korrelation	en						
		Z_GD	Z_A_dir_K	Z_A_ind_K	ZV_dir_k	ZV_ind_K	Z_K_ges
ZGD	r	1					
	Sig.						
	N	112					
Z_A_dir_K	r	-,001	1				
	_Sig.	,994					

	Ν	112	112				
Z_A_ind_k	(r	,330**	,625**	1			
	Sig	<,001	<,001				
	N	112	112	112			
Z_V_dir_K	r	,295**	,156	,346**	1		
	Sig.	,002	,100	<,001			
	N	112	112	112	112		
Z_V_ind_k	(r	,211*	-,227*	,097	-,003	1	
	Sig.	,026	,016	,307	,975		
	N	112	112	112	112	112	
Z_K_ges	r	,354**	,576**	,806**	,655**	,400**	1
	Sig.	<,001	<,001	<,001	<,001	<,001	
-	N	112	112	112	112	112	112

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

 $\label{eq:continuous} \textit{Anmerkungen}: r = Korrelationskoeffizient nach Pearson, Z_GD = Konformitätsdruck und Gruppendynamik, Z_A_dir_K = Anschluss und Kontaktbereitschaft (direkte Annäherung),$

Z_A_ind_K = Anschluss und Kontaktbereitschaft (indirekte Annäherung), Z_V_dir_K = Anschluss und Kontaktbereitschaft (direkte Vermeidung), Z_V_ind_K = Anschluss und Kontaktbereitschaft (indirekte Vermeidung), Z_K_ges = Anschluss und Kontaktbereitschaft (gesamt).

Nach der Korrelationsanalyse für die Skala *Konformitätsdruck und Gruppendynamik* lässt sich festhalten, dass beide Hypothesen hinsichtlich der vermuteten Zusammenhänge mit den Außenkriterien bestätigt werden können (siehe Tab. 13). Entsprechend lassen sich signifikant positive lineare Zusammenhänge für die Hypothese H_5 (r = .354) beobachtet. Der Koeffizienten deutet nach Cohen (1988) für die Hypothese H_5 auf einen mittleren hin. Weiterhin kann die untergeordnete Hypothese $H_{5.1}$ ebenfalls angenommen werden. Somit besteht ein mittelstarker signifikant positiver linearer Zusammenhang (r = .330) im Hinblick auf die Hypothese $H_{5.1}$.

H₆: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der Orientierung an Anderen in sozialen Situationen.

H₇: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der Skala Angst vor negativer Bewertung.

Die Hypothesen H_6 und H_7 lassen sich anhand der nachfolgenden Korrelationsmatrix prüfen. Folglich werden die entsprechenden Korrelationskoeffizienten betrachtet.

^{*.} Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Tab. 14: Korrelationsmatrix für Skala Konformitätsdruck und Gruppendynamik, SMSC: Orientierung an Anderen und SANB-5

Kο	$rr \triangle l$	latı	nη	αr

Nonelationen				
		Z_GD	Z_O_A	Z_SANB
Z_GD	r	1		
	Sig. (2-seitig)			
	N	112		<u></u>
Z_O_A	r	,258**	1	
	Sig. (2-seitig)	,006		
	N	112	112	
Z_SANB	r	,287**	,613**	1
	Sig. (2-seitig)	,002	<,001	
	N	111	111	111

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Anmerkungen: r = Korrelationskoeffizient nach Pearson, Z_GD = Konformitätsdruck und Gruppendynamik, Z_O_A = SMSC: Orientierung an Anderen, Z_SANB = Skala Angst vor negativer Bewertung.

Für beide Hypothesen konnte die Alternativhypothese angenommen werden. Entsprechend der Hypothese H_6 besteht ein signifikanter positiver linearer Zusammenhang zwischen den beiden Skalen (r = .258). Der Koeffizient deutet auf einen mittleren Zusammenhang hin (Cohen, 1988). Ferner konnte die Hypothese H_7 ebenso bestätigt werden. Folglich besteht ein signifikanter linearer Zusammenhang zwischen den beiden Skalen (r = .287). Überdies lässt sich dieser nach Cohen (1988) ebenfalls als mittelstarker positiver Zusammenhang beschreiben.

4.3.3.5 Korrelationsanalyse für die Skala Selbstkonzept und Selbstwert

H₈: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und dem globalen Selbstwertgefühl.

H₉: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und der Angst vor negativer Bewertung.

*H*₁₀: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und Neurotizismus.

Um die Hypothesen für die Skala *Selbstkonzept und Selbstwert* zu überprüfen, werden die Skalenwerte mit den Werten der Skalen *FSES, SANB-5* und *Neurotizismus* korreliert. Die Ergebnisse der Korrelationsanalyse für die Skala *Selbstkonzept und Selbstwert* sind in der folgenden Tabelle dargestellt.

Tab. 15: Korrelationsmatrix für die Skala Selbstkonzept und Selbstwert

Korrelationen

	Z_SK	Z_RSES	Z_SANB	Z_Neuro_ges
r	1			
Sig. (2-seitig)				
N	112			
r	-,297**	1		
Sig. (2-seitig)	,001			
N	112	112		
r	,458**	-,676 ^{**}	1	
Sig. (2-seitig)	<,001	<,001		
N	111	111	111	
r	,530**	-,644**	,691**	1
Sig. (2-seitig)	<,001	<,001	<,001	
N	112	112	111	112
	Sig. (2-seitig) N r Sig. (2-seitig) N r Sig. (2-seitig) N r Sig. (2-seitig) N	r 1 Sig. (2-seitig) N 112 r -,297** Sig. (2-seitig) ,001 N 112 r ,458** Sig. (2-seitig) <,001 N 111 r ,530** Sig. (2-seitig) <,001	r 1 Sig. (2-seitig) N 112 r -,297** 1 Sig. (2-seitig) ,001 N 112 112 r ,458** -,676** Sig. (2-seitig) <,001 <,001 N 111 111 r ,530** -,644** Sig. (2-seitig) <,001 <,001	r 1 Sig. (2-seitig) N 112 r -,297** 1 Sig. (2-seitig) ,001 N 112 112 r ,458** -,676** 1 Sig. (2-seitig) <,001 <,001 N 111 111 111 r ,530** -,644** ,691** Sig. (2-seitig) <,001 <,001 <,001

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Anmerkungen: r = Korrelationskoeffizient nach Pearson, Z_SK = Selbstkonzept und Selbstwert, Z_RSES = Rosenberg Self-Esteem-Scale, Z_SANB = Skala Angst vor negativer Bewertung, Z_Neuro_ges = Neurotizismus.

Wie aus der dargestellten Korrelationsmatrix für die Skala *Selbstkonzept und Selbstwert* hervorgeht, können die Hypothesen H_8 , H_9 und H_{10} bestätigt werden, da der Korrelationskoeffizient jeweils die vermuteten Zusammenhänge widerspiegelt (siehe Tabelle 14). So ergibt sich für die Hypothese H_8 ein signifikant negativer linearer Zusammenhang zwischen der Skala *Selbstkonzept und Selbstwert* und dem globalen Selbstwertgefühl (r = -.297). Der Zusammenhang kann nach Cohen (1988) als Korrelation mittlerer Stärke bezeichnet werden. Außerdem lässt sich im Hinblick auf die Hypothese H_9 ein signifikanter positiver linearer Zusammenhang (r = .458) zwischen der Skala *Selbstkonzept und Selbstwert* und der Skala *Angst vor negativer Bewertung* beobachten. Nach Cohen (1988) kann von einem starkem linearem Zusammenhang gesprochen werden. Weiterhin kann ein signifikanter linearer Zusammenhang hinsichtlich der Hypothese H_{10} beobachtet werden (r = .530). Entsprechend der Klassifikation nach Cohen (1988) kann auch hier ein starker linearer Zusammenhang bestimmt werden.

4.3.3.6 Korrelationsanalyse für die Skala Anschlussmotiv

H₁₁: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Anschlussmotiv und dem Bedürfnis nach Anschluss und Kontakt.

H₁₂: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Anschlussmotiv und Extraversion.

Zur Prüfung der Hypothesen für die Skala *Anschlussmotiv* wurden ebenfalls die entsprechenden Korrelationen bestimmt. In der nachfolgenden Tabelle ist die Korrelationsmatrix für die Skala Anschlussmotiv aufgeführt.

Tab. 16: Korrelationsmatrix für die Skala Anschlussmotiv

	•				•			
ĸ	\sim	rre	M2	١tı	1	n	Δ	n
<i>,</i> ,	v	,,,	10	L	v	,,	ᆫ	,

		Z_AM	Z_Kontakt_ges	Z_Extra_ges
Z_AM	r	1		
	Sig. (2-seitig)			
	N	112		
Z_Kontakt_ges	r	,003	1	
	Sig. (2-seitig)	,973		
	N	112	112	
Z_Extra_ges	r	-,099	,480**	1
	Sig. (2-seitig)	,300	<,001	
	N	112	112	112

^{**.} Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Anmerkungen: r = Korrelationskoeffizient nach Pearson, Z_AM = Anschlussmotiv, Z_Kontakt_ges = Anschluss und Kontaktbereitschaft, Z_Extra_ges = Extraversion

Auf Grundlage der Korrelationsanalyse müssen sowohl die Hypothese H_{10} als auch die Hypothese H_{11} verworfen werden. Schließlich zeigt sich bei Betrachtung des Korrelationskoeffizienten für die Skalen *Anschlussmotiv* und *Anschluss und Kontaktbereitschaft*, dass kein linearer Zusammenhang besteht (r = .003). Außerdem geht aus dem Koeffizienten für die Skalen *Anschlussmotiv* und *Extraversion* hervor, dass kein signifikanter linearer Zusammenhang zwischen diesen Skalen besteht (r = .009). Zusätzlich wurden alternative Korrelationsmaße (Spearman-Rho und Kendall-Tau) betrachtet. Allerdings konnten auch hier keine signifikanten Zusammenhänge beobachtet werden. Die entsprechenden Korrelationsmatrizen sind im Anhang C hinterlegt.

4.3.3.6 Bestimmung der Koeffizienten zur Einschätzung der Kriteriumsvalidität

Im Rahmen dieses Abschnittes werden die Ergebnisse der Korrelationsanalysen und Hypothesentestungen zusammengetragen sowie übersichtlich zusammengeführt, um eine Einschätzung der Kriteriumsvalidität im Sinne einer Übereinstimmungsvalidität vorzunehmen. Entsprechend bietet die nachfolgende Tabelle eine Übersicht über die gesammelten Befunde. Der *Validitätskoeffizienten* (r_{tc}) ergibt sich aus der Korrelation der jeweiligen Skalenwerte. Laut Bühner (2011) gelten Werte unter .40 als niedrig und über .60 als hoch. Es ist anzumerken, dass diese Beurteilungsrichtlinien lediglich als Faustformeln dienen. So geben Validitätskoeffizienten allein wenig Aufschluss über die Bedeutung eines Tests (Bühner, 2011).

Tab. 17: Übersicht über die Validitätskoeffizienten $Validitätskoeffizienten (r_{tc})$

		Validierungsskala	r _{tc}
Skala	Streben nach sozialer	B5-T: Leistungsmotiv	.53
	Anerkennung		
	Soziale Vergleiche	RSES	43
		B5-T: Neurotizismus	.58
		B5-T: Leistungsmotiv	.50
	Konformitätsdruck und	TEMEO-R: Anschluss & Kontakt-	.35
	Gruppendynamik	bereitschaft (Gesamt)	
		TEMEO-R: Anschluss & Kontakt-	.33
		bereitschaft (Direkte Annäherungs-	
		motivation)	
		SMSC: Orientierung an Anderen	.26
		SANB-5	.28
	Selbstkonzept und	RSES	.30
	Selbstwert	SANB-5	.46
		B5-T: Neurotizismus	.53
	Anschlussmotiv	TEMEO-R: Anschluss & Kontakt-	-
		bereitschaft (Gesamt)	
		B5-T: Extraversion	-

Auf Basis der Beurteilungsrichtlinien (Bühner, 2011) lässt sich festhalten, dass für die Skalen Streben nach sozialer Anerkennung und Soziale Vergleiche sowie Selbstkonzept und Selbstwert insgesamt mittlere Werte für die Kriteriumsvalidität nachgewiesen können. Für die Skala Konformitätsdruck und Gruppendynamik konnten nur niedrige Werte beobachtet werden. Ferner konnten keine signifikanten Zusammenhänge für die Skala Anschlussmotiv nachgewiesen werden. Daher konnten keine Befunde hinsichtlich der Validierung gesammelt werden.

4.4 Diskussion und Interpretation der Ergebnisse – Studie II

Im folgenden Abschnitt werden die Ergebnisse der zweiten Studie inhaltlich interpretiert und im Hinblick auf die theoretische Basis sowie Literatur diskutiert. Im Fokus steht dabei die Validierung der Skalen. Weiterhin sollen auch Limitierungen dieser zweiten Erhebung thematisiert werden.

4.4.1 Zusammenfassung und Interpretation der empirischen Befunde

Im Wesentlichen besteht die zweite Studie aus zwei Teilen. So sollten im ersten Schritt die relevanten Item- und Skalenwerte sowie die Reliabilität des FBMSN-R erneut

untersucht werden. Im zweiten Schritt erfolgte die Bestimmung der Kriteriumsvalidität für den revidierten Fragebogen.

4.4.1.1 Reanalyse des FBMSN-R

Zunächst wurde als ein Ziel dieser zweiten Studie ausgegeben, die Itemkennwerte und Reliabilitäten der Skalen des FBMSN-R erneut zu untersuchen, um die Ergebnisse der ersten Untersuchung bestenfalls zu replizieren. Im Hinblick auf die Itemkennwerte lagt der Fokus hauptsächlich auf den Trennschärfen der einzelnen Items jeder Subskala. Hier konnten die Ergebnisse aus der ersten Studie weitestgehend bestätigt werden. Lediglich Items GD_06, AM_06 und SV_04 wurden Trennschärfen unterhalb des Akzeptanzbereichs von .40 festgestellt (Kelava & Moosbrugger, 2020a). Laut Pospeschill (2022) kennzeichnen Trennschärfekoeffizienten Bereich von .30 bis .39 Items, die durch eine Anpassung verbessert werden könnten. Des Weiteren deuten Werte im Bereich von .20 und .29 darauf hin, dass diese Items definitiv angepasst werden sollten. Allerdings wurde zunächst keine Eliminierung oder Überarbeitung der genannten Items vorgenommen. Hinsichtlich der Reliabilität wurde die interne Konsistenz der Skalen mithilfe des Koeffizienten Cronbachs Alpha erneut bestimmt. Die Reliabilitätskoeffizienten der Skalen wiesen insgesamt akzeptable Werte auf und fallen überwiegend sogar gut bis sehr qut aus. Allerdings ist darauf hinzuweisen, dass die Skala Liebesbeziehungen mit α = .63 zwar einen akzeptablen Wert aufweist, im Vergleich zur ersten Untersuchung (α = .82) jedoch eine extreme Schwankung beobachtet wurde. Dennoch konnten auch hier die Ergebnisse der ersten Untersuchung im Wesentlichen bestätigt werden. Entsprechend deuten die Ergebnisse auf eine insgesamt angemessene und reliable Skalenstruktur des Messinstrumentes hin.

4.4.1.2 Validierung des FBMSN-R

Im nächsten Schritt erfolgte die Validierung des Fragebogens. Im Zuge dessen sollten erste Befunde zur Bestimmung der Kriteriumsvalidität, im Sinne einer Übereinstimmungsvalidität anhand von parallel erhobenen Außenkriterien zusammengetragen werden. Um die Validität des FBMSN-R zu prüfen, wurden Zusammenhänge zwischen den einzelnen Skalen und anderen Variablen untersucht, die entsprechende theoretische Bezugspunkte aufweisen. Auf Grundlage der Beurteilungsrichtlinien nach Bühner (2011) lässt sich dokumentieren, dass für die Skalen Streben nach sozialer Anerkennung und Soziale Vergleiche sowie Selbstkonzept und Selbstwert insgesamt mittlere Werte für die Kriteriumsvalidität nachgewiesen können. Wobei die Skala Soziale Vergleiche mit einem Wert von .58 für die Skala Neurotizismus den insgesamt höchsten beobachteten Wert erreicht. Für die Skala Konformitätsdruck und Gruppendynamik konnten nur niedrige Werte beobachtet werden. Außerdem konnten keinerlei signifikanten Zusammenhänge

für die Skala *Anschlussmotiv* beobachtet werden. Daher bleibt die Validierung dieser Skala ohne Befund.

Auf Basis der theoretischen Bezüge wurden verschiedene Hypothesen über die Zusammenhänge formuliert, welche im Rahmen der Validierung getestet wurden. Wie bereits erwähnt, können genau genommen nur korrelative Aussagen getroffen werden. Dennoch orientieren sich die folgenden Aussagen an der theoretisch vermuteten Einflussrichtung, um die Befunde zielgerichtet interpretieren zu können. Allerdings kann eine entgegengesetzte Wirkrichtung ebenfalls nicht ausgeschlossen werden.

Im Rahmen der empirischen Untersuchung der Hypothesen konnte die erste Hypothese (H₁: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Streben nach sozialer Anerkennung und dem Leistungsmotiv) für die vorliegende Stichprobe angenommen werden. Daher besteht ein signifikanter positiver linearer Zusammenhang (r = .525) zwischen dem Streben nach sozialer Anerkennung auf sozialen Netzwerken und dem Leistungsmotiv. Nach Satow (2020) suchen Menschen mit einem stark ausgeprägten Leistungsmotiv häufig nach sozialer Anerkennung für ihre Leistungen. Entsprechend wird auf Basis der Ergebnisse davon ausgegangen, dass ein ausgeprägtes Leistungsmotiv mit einem erhöhten Streben nach sozialer Anerkennung auf sozialen Netzwerken einhergeht. Auf sozialen Netzwerken sind Äußerungen des sozialen Feedbacks allgegenwärtig. Sie können anhand eines einfachen Klicks auf den "Gefällt mir" - Button erfolgen, um die positive Bewertung eines Posts oder Bildes auszudrücken. Weiterhin können Rückmeldungen von einem deutlich erweiterten Personenkreis empfangen werden, als es bei einer Face-to-Face-Interaktion der Fall wäre (Manago, Graham, Greenfield, & Salimkhan, 2008; Siibak, 2009). Im Zusammenhang mit einem besonderen Streben nach sozialer Anerkennung könnte vermutet werden, dass Menschen mehr Zeit auf sozialen Netzwerken verbringen, da soziales Feedback und Anerkennung leichter und in größerer Maße verfügbar ist als in der Realität.

Im Zuge der Untersuchung der zweiten Hypothese (H_2 : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und dem globalen Selbstwertgefühl), konnte ebenfalls ein signifikanter linearer Zusammenhang (r = -.430) beobachtet werden. Die negative Wirkungsrichtung des Zusammenhangs könnte dahin gehend interpretiert werden, dass häufige soziale Vergleiche auf sozialen Netzwerken zu einer Minderung des globalen Selbstwertgefühls führen. Die aktuelle Forschung stützt diese Interpretation: So zeigen Steers, Wickham und Acitelli (2014), dass der ständige Vergleichsdruck negative Auswirkungen auf das Wohlbefinden sowie den Selbstwert der Nutzer von sozialen Netzwerken hat. Für die zweite Hypothese (H_3 : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und Neurotizismus) im Hinblick auf die Skala Soziale Vergleiche wurde der Zusammenhang mit der Skala Neurotizismus untersucht. Nach Betrachtung des Korrelationskoeffizienten (r = 1)

.579) lässt sich auch diese Hypothese für die vorliegende Stichprobe annehmen. Im Hinblick auf die aktuelle Forschung deuten mittlerweile einige Studien darauf hin, dass es bei der Betrachtung von Profilen auf Social-Media-Plattformen zu sozialen Vergleichsprozessen nach Festinger (1954) kommt. Die Vergleichsprozesse können zum Teil negative Auswirkungen haben, so führt die Nutzung von Facebook zu einer Steigerung der Unzufriedenheit und Neid, indem die Nutzer ihre eigene Lebenssituation mit denen der anderen Nutzern auf Grundlage der geteilten Beiträge vergleichen (Kneidinger-Müller, 2017). Weitere Untersuchungen zeigen, dass Studierende, die seit Längerem auf Facebook aktiv sind, eher glauben, dass andere Personen glücklicher sind als sie selbst (Chou & Edge, 2012). Weiterhin konnte eine signifikante lineare Korrelation (r = .500) im Sinne der Hypothese H₄ (Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Soziale Vergleiche und dem Leistungsmotiv) beobachtet werden. Nach Festiger (1954) wird die Selbstverbesserung als ein Motiv des sozialen Vergleichs beschrieben. In der Praxis kann ein Vergleichsprozess mit einer gewählten Referenzgruppe zu einer gesteigerten Leistungsmotivation führen (Morf & Koole, 2014). Entsprechend kann auch diese Hypothese durch die theoretische Basis gestützt werden.

Für die Validierung der Skala Konformitätsdruck und Gruppendynamik wurden zwei Hypothesen formuliert. Zur empirischen Prüfung der Hypothese H₅ wurde der korrelative Zusammenhang zwischen der Skala Konformitätsdruck und Gruppendynamik und dem Bedürfnis nach Zugehörigkeit und Anschluss untersucht. Im Rahmen der Hypothesentestung ist die Korrelation signifikant geworden und es konnte ein linearer positiver Zusammenhang beobachtet werden (r = .354). Entsprechend wird die Hypothese (H_5 : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und Bedürfnis nach Zugehörigkeit und Anschluss) für die vorliegende Verteilung angenommen. So bestimmt das Anschlussmotiv eine ausgeprägte Präferenz für soziale Interaktionen sowie Kontakt mit unbekannten oder bekannten Personen. Im Hinblick auf digitale soziale Medien zeigt Kim (2014), dass soziale Kontakte und das Gefühl von Zugehörigkeit zu einer (virtuellen) Gruppe wichtige Treiber für die Verwendung verschiedener Online-Angebote sind. Weiterhin konnte die untergeordnete Hypothese $H_{5,1}$ ebenfalls bestätigt werden. So wird ein Wert von r = .330 erreicht, was auf einen positiven linearen Zusammenhang hindeutet. Außerdem ist die Korrelation signifikant geworden. Daraus kann geschlossen werden, dass ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der indirekten Annäherungsmotivation bezüglich des Anschlussmotivs besteht. Dieses zeichnet sich dadurch aus, dass die Kontaktaufnahme für ein übergeordnetes Ziel instrumentalisiert wird (Deutscher & Mees, 2012). Im Hinblick auf die vorliegende Skala könnte das Bedürfnis nach Zugehörigkeit und Anerkennung innerhalb der sozialen Gruppe als übergeordnetes Ziel interpretiert werden. Bei Betrachtung der Korrelationsmatrix in

Kapitel 4.3.3.4 Korrelationsanalyse für die Skala Konformitätsdruck und Gruppendynamik (oder Anhang C) zeigt sich, dass diese Subskala den höchsten Zusammenhang im Vergleich zu den anderen Subskalen des TEMEO-R: Anschluss & Kontaktbereitschaft aufweist. Entsprechend kann dieser Subskala für die Skala Gruppendynamik und Konformitätsdruck eine besondere Bedeutung zuteilwerden, um eine genauere Untersuchung im Hinblick auf die Validierung dieser Skala vorzunehmen. Des Weiteren konnte die sechste Hypothese (H_6 : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der Orientierung an Anderen in sozialen Situationen) für die vorliegende Stichprobe angenommen werden. Somit besteht ein signifikanter positiver linearer Zusammenhang (r = .258) zwischen der Skala Gruppendynamik und Konformitätsdruck und der Orientierung an Anderen in sozialen Situationen. Diese Annahme lässt sich durch die theoretischen Konzepte im Zusammenhang mit Konformitätsdruck und dem Streben nach Zugehörigkeit zu einer bestimmten sozialen Gruppe stützen (Hewstone & Martin, 2014). Weiterhin stützen diese Annahmen die Vermutung, dass ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und Bewertungsängstlichkeit besteht. Nach der Untersuchung lässt sich für die Hypothese (H_7 : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Gruppendynamik und Konformitätsdruck und der Skala Angst vor negativer Bewertung) ein signifikant positiver linearer Zusammenhang feststellen (r = .287).

Zur Validierung der Skala Selbstkonzept und Selbstwert wurden weitere drei Hypothesen im Hinblick auf den Zusammenhang mit verschiedenen Außenkriterien formuliert. So wurde ein signifikanter linearer Zusammenhang zwischen der Skala mit dem globalen Selbstwertgefühl vermutet. Entsprechend erfolgte eine Untersuchung der korrelativen Beziehung der Skala Selbstkonzept und Selbstwert und der RSES, welche das globale Selbstwertgefühl abbildet. Im Zuge dessen konnte die Hypothese (H_8 : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und dem globalen Selbstwertgefühl) angenommen werden, da ein signifikant negativer linearer Zusammenhang (r = -.297) zwischen den Skalen besteht. Entsprechend wird vermutet, dass eine hohe Ausprägung auf der Skala Selbstkonzept und Selbstwert mit einem niedrigem globalem Selbstwertgefühl einhergeht. Diese Annahme lässt sich durch aktuelle Forschungen stützen. So sind negative Auswirkungen auf das subjektive Wohlbefinden sowie den Selbstwert von Nutzern sozialer Netzwerke bereits in verschiedensten Studien nachgewiesen worden (Steers, Wickham & Acitelli, 2014). Ferner zeigen die Ergebnisse von Sumter und Kollegen (2012), dass diese negativen Erfahrungen sozialer Interaktionen die Lebenszufriedenheit langfristig negativ beeinflussen können. Weiterhin wurde im Rahmen der Hypothese H₉ ein linearer Zusammenhang der Skala Selbstkonzept und Selbstwert und der Bewertungsängstlichkeit vermutet. Entsprechend wurde

eine Korrelationsanalyse für die vorliegende Skala und die Skala Angst vor negativer Bewertung (SANB-5) untersucht. So konnte auch diese Hypothese (H₉: Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und der Angst vor negativer Bewertung) für die vorliegende Stichprobe angenommen werden, da eine signifikante Korrelation beobachtet werden konnte (r = .458). Folglich besteht ein positiver, linearer Zusammenhang. Überdies wird ebenfalls eine erhöhe Sensitivität im Hinblick auf negative Rückmeldungen aus sozialen Netzwerken vermutet. In diesem Zusammenhang geht aus aktuellen Studien hervor, dass negatives Feedback aus sozialen Netzwerken zu Wutausbrüchen, Scham sowie Hilflosigkeit und Depressionen führen können (Ortega et al., 2009). Diese Ergebnisse stützen ebenfalls die nächste Hypothese (H_{10} : Es besteht ein signifikanter linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und Neurotizismus). Diese konnte ebenso für die vorliegende Stichprobe angenommen werden, da die Produkt-Moment-Korrelation mit einem Korrelationskoeffizient von r = .530 signifikant geworden ist. Entsprechend besteht ein positiver linearer Zusammenhang zwischen der Skala Selbstkonzept und Selbstwert und der Persönlichkeitsdimension Neurotizismus.

Aus der aktuellen Forschung geht hervor, dass das Teilen von Informationen auf sozialen Netzwerken von bestimmten Bedürfnissen geleitet wird, die sozialen Kontakt beeinflussen (Ho & Dempsey, 2010). Dies gilt in erster Linie für das Bedürfnis nach Zugehörigkeit. Dieses Bedürfnis hat wiederum einen besonderen Einfluss darauf, in welchem Umfang sowie auf welche Art und Weise Menschen soziale Netzwerke nutzen (Krämer, Eimler, & Neubaum, 2017). Entsprechend wurde ein signifikanter linearer Zusammenhang zwischen der Skala Anschlussmotiv und dem Bedürfnis nach Anschluss und Kontakt (H_{11}) . Jedoch musste nach der Untersuchung der Produkt-Moment-Korrelation sowie weiterer Korrelationsmaße (Kendall-Tau und Spearman-Rho) die Hypothese H_{11} verworfen werden und die Nullhypothese H_0 für die vorliegende Stichprobe beibehalten werden. Entsprechend besteht kein signifikanter Zusammenhang zwischen den beiden Variablen. Weiterhin wurde für die Validierung der Skala Anschlussmotiv ein signifikanter linearer Zusammenhang mit der Persönlichkeitsdimension Extraversion vermutet. Nach Satow (2020) sind sowohl Geselligkeit als auch Kontaktfreudigkeit charakteristisch für dieses Merkmal. Außerdem kann das Bedürfnis nach Zugehörigkeit und Anschluss als Aspekt von Extraversion aufgefasst werden (Jha, 2010). Allerdings musste auch für diese Hypothese die Nullhypothese (H_0) beibehalten und die Alternativhypothese (H_{12}) verworfen werden. Dementsprechend konnte ebenfalls keine signifikante Korrelation sowohl für die Produkt-Moment-Korrelation als auch für weiterer Korrelationsmaße (Kendall-Tau und Spearman-Rho) beobachtet werden.

4.4.2 Kritische Reflexion

An dieser Stelle soll auf verschiedene Punkte eingegangen, die zu einer Limitierung der Aussagekraft und Übertragbarkeit der Ergebnisse sowie Interpretationen dieser Studie führen könnten. Im Zuge dessen soll eine kritische Reflexion des bisherigen Forschungsprozesses erfolgen. Dabei sollen sowohl die beobachteten Ergebnisse als auch deren Interpretation im Fokus stehen. Außerdem soll auch auf Limitierungen und Einschränkungen eingegangen werden.

4.4.2.1 Gütekriterien der verwendeten Messinstrumente

Durch die Verwendung der Software Unipark, mit deren Hilfe die Erhebung als Online-Befragung durchgeführt werden konnte, sind sowohl eine sehr hohe Durchführungs- als auch eine sehr hohe Auswertungsobjektivität gewährleistet. Entsprechend erhielten alle Probanden die gleichen Instruktionen bei der Durchführung der Studie. Weiterhin wurde die Reliabilität der verwendeten Validierungsskalen analysiert, um den Grad der Zuverlässigkeit der Messungen zu bestimmen. Nach Bühner (2011) gilt eine Reliabilität von .6 als akzeptabel. Basierend auf dieser Annahme, können die berechneten Reliabilitätskoeffizienten für die Skalen B5-T: Neurotizismus ($\alpha = .88$), B5-T: Extraversion ($\alpha = .84$), B5-T: Leistungsmotiv (α = .74), RSES (α = .89) und SANB-5 (α = .87) für gut bis sehr gut befunden werden. Für die Skala *TEMEO-R: Anschluss & Kontaktbereitschaft* (α = .67) konnte ein akzeptabler Wert beobachtet werden. Lediglich für die Skala SMSC: Orientierung an Anderen ($\alpha = .45$) wurde ein sehr schlechter Wert erreicht. Dies ist wahrscheinlich auf die Zusammensetzung der Stichprobe zurückzuführen. So hängt die Reliabilität vom Ausmaß der wahren Unterschiede ab. Dem zur Folge ist nicht auszuschließen, dass die vorliegende Stichprobe im Hinblick auf einzelne Merkmale homogen ist und somit eine geringere Varianz der wahren Werte aufweist (Eid, Gollwitzer, & Schmitt, 2017). Die Reliabilität konnte von Collani und Stürmer (2003) anhand einer größeren Normstichprobe (N = 324) nachgewiesen werden. Entsprechend können die Ergebnisse dieser Validierungsstudie relativ zuverlässig interpretiert werden.

4.4.2.2 Limitierungen und Einschränkungen

Ein wesentlicher Punkt, der im Rahmen dieser Untersuchung im Hinblick auf Einschränkungen der Interpretation der Ergebnisse zu nennen ist, ist die Zusammensetzung der Stichprobe. Die vorliegende Stichprobe kann nicht als repräsentativ für die Grundgesamtheit betrachtet werden. Somit erlaubt die untersuchte Stichprobe keine Rückschlüsse auf die Grundgesamtheit. Entsprechend beziehen sich die Ergebnisse lediglich auf die vorliegende Stichprobe. Weiterhin wurde der Fragebogen von Burdorf und Lux (2022) für eine Zielgruppe 18 bis 35 Jahren konzipiert. Diese wurde im Rahmen dieser Untersuchung überschritten, um eine möglichst große Validierungsstichprobe zu

akquirieren. Außerdem erfolgte die Akquirierung der Probanden in erster Linie über persönliche Kontakte des Versuchsleiters. Daraus resultiert, dass viele Teilnehmer aus dem Kreis Düren kommen und zum engeren Bekanntenkreis des Versuchsleiters gehören. Weiterhin besteht die Stichprobe überwiegend aus Studierenden der Hochschule Fresenius Köln, da diese sich für eine Teilnahme eine Versuchspersonenstunde anrechnen lassen konnten. Daraus resultiert, dass die Stichprobe größtenteils aus Studierenden und Akademikern besteht. Dem zur Folge sind andere Berufsgruppen oder Bildungsschichten in die Studie unterrepräsentiert. Des Weiteren bringt die Erfassung über Selbsteinschätzung der Teilnehmer grundsätzlich die Gefahr systematischer Urteilsfehler, beispielsweise die Verzerrung der Einschätzungen durch soziale Erwünschtheit, mit sich (Eid, Gollwitzer, & Schmitt, 2017). Um dies vorzubeugen, wurde auf die Anonymität sowie dem verantwortungsbewussten Umgang mit den erhobenen Daten hingewiesen. In diesem Zusammenhang muss weiterhin davon ausgegangen werden, dass die vorliegende Validierungsstichprobe als nicht repräsentativ für die Zielgruppe bewertet werden muss. Entsprechend bedarf es weiterer Analysen anhand einer Stichprobe, die als repräsentativ für die Zielgruppe beurteilt werden kann.

Bezüglich der Interpretation der Korrelationsanalysen zur Validierung der Skalen des FBMSN-R ist an dieser Stelle darauf hinzuweisen, dass eine signifikante Korrelation lediglich als Hinweis für die Kriteriumsvalidität der Skala zu verstehen ist. Schließlich kann auf Grundlage einer Korrelation keine Aussage über die Kausalbeziehung zweier Variablen getroffen werden (Eid, Gollwitzer, & Schmitt, 2017). Somit sollten die Zusammenhänge im Rahmen weiterer Forschungen genauer untersucht werden. So könnten die Zusammenhänge durch eine Regressionsanalyse weiter untersucht werden, um die Befunde im in Blick auf die Vorhersagekraft zu bestätigen.

4.4.2.3 Theoretischer Hintergrund

Im Rahmen der Literaturrecherche konnten einige passende Testverfahren und Skalen gefunden werden, um erste Befunde für die Bestimmung der Kriteriumsvalidität zu sammeln. Allerdings wiesen die meisten Verfahren, die frei im Internet zur Verfügung stehen, Mengel in der Konstruktvalidität auf. Außerdem steht eine vollständige Normierung der meisten Verfahren ebenfalls noch aus.

4.5 Fazit - Studie II

Zusammenfassend lässt sich festhalten, dass für die Subskalen Streben nach sozialer Anerkennung, Soziale Vergleiche, Konformitätsdruck und Gruppendynamik sowie Anschlussmotiv und Selbstkonzept und Selbstwert die guten bis sehr guten Werte für die Reliabilitätskoeffizienten repliziert werden konnten. Im Hinblick auf die Skala Freizügige Selbstdarstellung konnte ebenfalls ein hoher Wert für die Reliabilität beobachtet.

Lediglich für die Skala Liebesbeziehungen konnte im Rahmen der zweiten Erhebung nur ein akzeptabler Wert für die interne Konsistenz nach Cronbachs Alpha beobachtet werden. Weiterhin lässt sich dokumentieren, dass für vier der fünf Subskalen des FBMSN-R erste Hinweise und Befunde für die Kriteriumsvalidität gesammelt werden konnten. Im Zuge dessen wurden verschiedene Zusammenhänge zwischen den einzelnen Subskalen und weiteren Außenkriterien untersucht. Die Korrelationsanalyse für die einzelnen Subskalen mit den entsprechenden Kriterien lieferte mittlere Werte für die Validitätskoeffizienten für die Skalen Streben nach sozialer Anerkennung, Soziale Vergleiche sowie Selbstkonzept und Selbstwert. Für die die Skala Konformitätsdruck und Gruppendynamik konnten lediglich niedrige Werte für die beobachteten Skalen ermittelt werden. Dennoch konnten hinreichende theoretische Anknüpfungspunkte und aktuelle Forschungen in den Validierungsprozess einbezogen werden, um weitere inhaltliche Befunde für die Validierung des FBMSN-R zusammenzutragen. Lediglich für die Skala Anschlussmotiv ließen sich keine Zusammenhänge in Bezug auf die erhobenen Kriterien nachweisen. Entsprechend konnten noch keine Befunde für die Validierung der Skala Anschlussmotiv gesammelt werden.

5 Diskussion

Im Rahmen des nachfolgenden Kapitels erfolgt die eine abschließende Zusammenfassung und Interpretation der Ergebnisse beider Studien, die in den vorangegangenen Kapiteln geschildert wurden. Darauf folgt eine kritische Schlussbetrachtung der Vorgehensweise sowie der verwendeten Methoden und der Ergebnisse. Abschließend sollen Schlussfolgerungen sowie ein Ausblick für weiterführende Untersuchungen gegeben werden.

5.1 Zusammenfassung und Interpretation der Ergebnisse

Im Folgenden soll sowohl der gesamte Forschungsprozess als auch die Ergebnisse und Meilensteine beider Studien zusammenfassend dargestellt werden. Auf eine detaillierte Interpretation einzelner Ergebnisse wird an dieser Stelle verzichtet, da diese bereits im Rahmen der Kapitel 3.4 Diskussion und Interpretation der Ergebnisse – Studie I und 4.4 Diskussion und Interpretation der Ergebnisse – Studie II stattgefunden hat.

Im Rahmen der ersten Studie konnte auf Basis einer umfangreichen Item- und Skalenanalyse eine erste Revision des Fragebogens vorgenommen werden. Weiterhin wurden erste Befunde für die Reliabilitätskoeffizienten der Skalen gesammelt. Anschließend konnte die faktorielle Struktur des Fragebogens analysiert und entsprechend revidiert werden. Dazu erfolgte zunächst eine Hauptkomponentenanalyse mit Promax-Rotation. Anschließend wurden verschiedene Methoden zur Bestimmung der Extraktionskriterien herangezogen, um die optimale Faktorenanzahl zu bestimmen. Die finale

Faktorenlösung setzt sich aus sieben Faktoren zusammen, welche mittels Hauptachsenanalyse mit Promax-Rotation extrahiert wurden. Entsprechend lieferten die Analysen insgesamt sieben Faktoren, die sinnvoll interpretiert werden konnten. Nach einer entsprechenden Umstrukturierung des Fragebogens und der Selektion weiterer unpassender oder redundanter Items wurde eine Skalenstruktur gemäß diesen sieben Faktoren realisiert. Daher konnten die folgenden Skalen abgeleitet werden: Streben nach sozialer Anerkennung, Soziale Vergleiche, Konformitätsdruck und Gruppendynamik, Anschlussmotiv. Selbstkonzept und Selbstwert sowie Liebesbeziehungen und Freizügige Selbstdarstellung. Im Rahmen der Interpretation der einzelnen Faktoren zeigte sich, dass die Skalen Liebesbeziehungen und Freizügige Selbstdarstellung nicht im Sinne der Operationalisierung interpretiert werden konnten. Allerdings konnte aus zeitökonomischen Gründen keine Überarbeitung oder Generierung neuer Items erfolgen. Entsprechenden lagt der Fokus im Rahmen der weiteren Analysen und Interpretationen auf den verbliebenen fünf Skalen (Streben nach sozialer Anerkennung, Soziale Vergleiche, Konformitätsdruck und Gruppendynamik, Anschlussmotiv, Selbstkonzept und Selbstwert), da diese hinsichtlich der theoretischen Fundierung und Operationalisierung sowie in Bezug auf die Zielsetzung des Fragebogens untersucht werden konnten. Als Ergebnis der ersten Studie konnte somit die revidierte Version des Fragebogens (FBMSN-R) berichtet werden. Dieser basiert auf einer faktoranalytisch geprüften Struktur und konnte durch die Kürzung wesentlich ökonomischer gestaltet werden. Weiterhin liefert der FBMSN-R gute Ergebnisse für die Itemkennwerte und für die interne Konsistenz der einzelnen Skalen.

Entsprechend konnte die optimierte Version anhand einer neuen Stichprobe erhoben werden. Weiterhin wurden in die Erhebung weitere Skalen integriert, um die Kriteriumsvalidität (bzw. Übereinstimmungsvalidität) zu bestimmen. Die Außenkriterien sollten dann im Hinblick auf ihre theoretischen Bezüge und Zusammenhänge mit den Skalen des FBMSN-R untersucht werden. Zunächst wurden allerdings die Itemwerte sowie die Reliabilitätskoeffizienten erneut bestimmt, um die Ergebnisse der ersten Studie bestenfalls zu replizieren. Insgesamt konnten für die Itemkennwerte und die Reliabilitätskoeffizienten gute bis sehr gute Werte beobachtet werden. Einzelne Werte wischen zwar zum Teil deutlich von der ersten Studie ab, blieben jedoch in einem akzeptablen Bereich. Entsprechend weist der FBMSN-R eine angemessen reliable Skalenstruktur auf. Die Skalen Liebesbeziehung und Freizügige Selbstdarstellung wurden nicht in die Validierung einbezogen, da diese nicht im Hinblick auf die theoretische Operationalisierung interpretiert werden konnten. Zur Bestimmung der Kriteriumsvalidität wurden die Skalen B5-T: Neurotizismus, B5-T: Extraversion, B5-T: Leistungsmotiv, SMSC: Orientierung an Anderen, RSES sowie SANB-5 und TEMEO-R: Anschluss & Kontaktbereitschaft herangezogen. Auf Basis der theoretischen Hintergründe und aktuellen Forschungen wurden zwölf Haupthypothesen sowie eine untergeordnete Hypothese über die vermuteten Zusammenhängen zwischen den Skalen des FBMSN-R und den Außenkriterien formuliert. Diese wurden anschließend geprüft und die Ergebnisse interpretiert, um erste Befunde für die Validierungen des Fragebogens zusammenzutragen. Im Rahmen der Hypothesenprüfung konnten elf der insgesamt 13 Hypothesen für die vorliegende Stichprobe angenommen werden. Die beiden Hypothesen (H_{11} und H_{12}) für die Skala Anschlussmotiv konnten nicht bestätigt werden. Daher können Befunde für die Validität der Skala Anschlussmotiv berichtet werden. Nichtsdestotrotz konnten für die Skalen Streben nach sozialer Anerkennung, Soziale Vergleiche sowie Selbstkonzept und Selbstwert mittlere Werte im Hinblick auf die Kriteriumsvalidität nachgewiesen werden. Für die Skala Konformitätsdruck und Gruppendynamik erreichen die Validitätskoeffizienten lediglich niedrige Werte. Dennoch konnten erste Befunde für die Kriteriumsvalidität gesammelt werden. Diese konnten wiederum durch aktuelle Forschungsergebnisse gestützt sowie auf Basis des theoretischen Fundaments untermauert werden.

5.2 Kritische Schlussbetrachtung

Der nachfolgende Abschnitt umfasst die kritische Würdigung aller Untersuchungen, die im Rahmen dieser Thesis durchgeführt wurden. Im Zuge dessen werden sowohl die Datenerhebung sowie die Stichproben als auch die verwendeten Methoden und die gewonnenen Ergebnisse kritisch reflektiert. Ferner sollen Limitierungen und Einschränkungen des Geltungsbereichs der Ergebnisse beleuchtet werden.

5.2.1 Datenerhebung und Stichproben

Die Datenerhebung wurde in beiden Studien als schriftliche Befragung mithilfe der Online-Plattform *Unipark* durchgeführt. Die Fragebogenmethode ist äußerst ökonomisch, obwohl sie einen hohen Grad der Strukturierung erfordert. So bedarf es für diese Methode kaum Material oder besondere Kosten. Weiterhin ist die Durchführung besonderen Schulungsaufwand und in Gruppen möglich. Des Weiteren kann durch diese Methode eine hohe Durchführungs- und Auswertungsobjektivität gewährleistet werden. Außerdem ist eine hohe Vergleichbarkeit der Ergebnisse über mehrere Teilnehmer, Messzeitpunkte sowie Durchführungskontexte gegeben. Entsprechend konnten Verzerrungen wie Versuchsleitereffekte vermieden werden (Eid, Gollwitzer, & Schmitt, 2017). Weiterhin bietet die Nutzung eines computergestützten Online-Verfahrens die Möglichkeit, alle Items als Pflichtitems zu markieren, um Missing-Data-Probleme vorzubeugen. Überdies bietet die Optimierung für mobile Endgeräte (z. B. Smartphones oder Tablets) den Teilnehmern die Möglichkeit, zu jeder Zeit sowie in einem beliebigen Umfeld an der Untersuchung teilzunehmen und die Befragung gegebenenfalls zu pausieren (Eid, Gollwitzer, & Schmitt, 2017). Folglich kann eine sehr große Reichweite erzielt werden. Dennoch

ergeben sich bei Online-Befragungen oft hohe Abbruchquoten. Um dies vorzubeugen, wurden die ethischen Richtlinien der Deutschen Gesellschaft für Psychologie (DGP) beachtet. Zudem wurden die Teilnehmer über ihre Anonymität und vertrauliche Behandlung ihrer Angaben aufgeklärt.

Für die jeweiligen Stichproben wurden bereits umfangreiche Überlegungen hinsichtlich verschiedener Selektions- und Verzerrungseffekte gemacht (siehe Kapitel 3.4.3.2 Limitierungen und Einschränkungen bzw. Kapitel 4.4.2.2 Limitierungen und Einschränkungen). Entsprechend wird an dieser Stelle auf eine Wiederholung verzichtet. Nichtsdestotrotz ist an dieser Stelle darauf hinzuweisen, dass hinsichtlich der Fragebogenrevision nicht ausgeschlossen werden kann, dass aufgrund von Verzerrungen der Ersterhebung Items eliminiert wurden, die sich in der zweiten Erhebung möglicherweise als geeignet erwiesen hätten.

5.2.2 Methodische Vorgehensweise

Wie bereits im Rahmen der ersten Studie kurz angeschnitten wurde, ist ein gewisses Maß an Subjektivität sowohl bei Itemanalyse und -selektion als auch bei den anschließenden Faktorenanalysen nicht auszuschließen. Dennoch wurde versucht, möglichst klare Kriterien und Beurteilungsrichtlinien auszuwählen, um eine eindeutige Nachvollziehbarkeit der einzelnen Schritte zu gewährleisten. Dennoch sind gewisse Kompromisse im Rahmen der Itemanalyse aufgrund wechselseitigen Abhängigkeiten der Itemkennwerte nicht zu vermeiden. Im Rahmen der Validierung hätte zunächst der vollständige SMSC: Fragebogen zum Verhalten in sozialen Situationen und nicht nur die Skala Orientierung an Anderen erhoben werden sollen. Allerdings ist dies auf einen persönlichen Fehler des Versuchsleiters zurückzuführen. Dennoch wird an dieser Stelle darauf hingewiesen, dass sich die vollständige Skala aufgrund ihrer inhaltlichen Bezüge als ein geeignetes Validierungsinstrument gewesen wäre. Außerdem eigneten sich die verwendeten Validierungsskalen nur bedingt zur Bestimmung der Kriteriumsvalidität, da eine Konstruktvalidierung oder Normierung für einzelne Skalen zum Zeitpunkt der Erhebung noch ausstand. Allerdings musste aufgrund limitierten finanziellen Ressourcen auf die Verfahren zurückgegriffen werden, die kostenfrei verfügbar waren. Ferner wäre eine Konstruktvalidierung auf Basis der Multitrait-Multimethod-Analyse nach Campell und Fiske (1959) die beste Methode, um anschließend nachhaltige Aussagen über die Validität des Fragebogens treffen zu können. Allerdings ist dieses Verfahren sehr aufwendig und zeitintensiv. Außerdem sollte diese Masterthesis lediglich als Erstüberprüfung des Fragebogens dienen.

5.2.3 Reflexion der Ergebnisse

Im Rahmen der ersten Studie wurde der Fragebogen zunächst analysiert und anschließend auf Basis der Ergebnisse der Itemanalyse und der Faktorenanalysen gekürzt und umstrukturiert. Im Zuge dessen wurden sehr harte Kriterien für die Item- und Skalenselektion durchgesetzt, da eine Prüfung der faktoriellen Struktur zunächst keine interpretierbaren Ergebnisse lieferte. In diesem Zusammenhang kann nicht ausgeschlossen werden, dass potenziell geeignete Items entfernt wurden. An dieser Stelle ist darauf hinzuweisen, dass für die Skalen Modelllernen und Minderwertigkeit im Rahmen der Itemanalyse relativ gute Ergebnisse für die einzelnen Item- und Skalenwerte erzielt wurden. Allerdings mussten die Skalen aufgrund mangelnder Operationalisierung und entsprechend limitierten inhaltlichen Bezügen selektiert werden. Dennoch könnten die Items weiterentwickelt und angepasst werden, um sich doch noch als zweckdienlich zu erweisen. Selbiges gilt für die Skala Eskapismus. Allerdings ist in diesem Zusammenhang darauf hinzuweisen, dass Eskapismus ein kaum zu operationalisierendes Konstrukt darstellt. Außerdem wurden die Skalen Freizügige Selbstdarstellung und Liebesbeziehungen nicht weitergehend untersucht, da sie nicht im Hinblick auf die Operationalisierung interpretiert werden konnten. Entsprechend steht auch für diese Skalen eine Anpassung sowie Überarbeitung noch aus. So müssten vor allem für diese beiden Skalen weitere Items generiert werden, da die Interpretierbar der beiden Skalen deutlich limitiert ist. Im Rahmen der zweiten Studie wurden lediglich die Itemkennwerte und die Reliabilitätskoeffizienten erneut überprüft. Entsprechend steht eine erneute Prüfung der Skalenstruktur noch aus. Allerdings konnte auf Grundlage der Itemkennwerte und der Reliabilitätskoeffizienten eine angemessene Skalenstruktur angenommen werden. Ferner musste aufgrund des limitierten zeitlichen Rahmens dieser Thesis auf eine erneute Überprüfung und Umstrukturierung, welche sich eventuell hätte, ergeben können, verzichtet werden. Anschließend wurden in der zweiten Studie Zusammenhängen mit bestimmten Außenkriterien empirisch untersucht. Aufgrund der Tatsache, dass die einzelnen Skalen teilweise auf unterschiedlichen Bewertungssystemen und Testwertverteilungen basieren wurde eine z-Transformation durchgeführt, um die Vergleichbarkeit der einzelnen Skalenwerte zu gewährleisten. Entsprechend konnte für die Validierung auf die Produkt-Moment-Korrelation zurückgegriffen werden. An dieser Stelle wird jedoch darauf hingewiesen, dass anhand einer Korrelation keine Aussage über die Kausalbeziehung von Variablen getroffen werden kann (Eid, Gollwitzer, & Schmitt, 2017).

Außerdem wurde der Begriff Social Media Plattform für beide Studien relativ breit gefasst, um jeweils eine möglichst große Stichprobe zu akquirieren. So wurde WhatsApp aufgrund der Story-Funktion und der Möglichkeit unmittelbar auf diese zu reagieren sowie YouTube aufgrund der Kommentarfunktion und der Möglichkeit einen eigenen

Channel zu unterhalten ebenfalls als soziale Netzwerke behandelt. Allerdings ist fraglich, welche Probanden die Plattformen tatsächlich in dieser Hinsicht nutzen.

Weiterhin war die Vorgehensweise für beide Studien überwiegend explorativer Natur. Entsprechend wurden die Hypothesen in der zweiten Studie sehr unspezifisch formuliert und lediglich Vermutungen über allgemeine Zusammenhänge der einzelnen Skalen dokumentiert. Abschließend ist festzuhalten, dass sich alle Ergebnisse dieser Studie ausschließlich auf die jeweiligen Stichproben beziehen. Es können somit auch keine Aussagen über die Grundgesamtheit oder über das genaue Ausmaß der Validität getroffen werden. Die Ergebnisse können lediglich als erste empirische Befunde für eine vollständige Validierung betrachtet werden.

5.3 Schlussfolgerungen und Ausblick

Für das übergeordnete vorliegende Thesis und den überprüften Fragebogen lässt sich festhalten, dass keine vollständig optimierte und fertige Version des Fragebogens von Burdorf und Lux (2022) konturiert werden konnte. Allerdings war diese Zielsetzung auch sehr utopisch und bezog sich eher auf einen sehr unwahrscheinlichen Optimalfall. Dennoch konnte, auf Grundlage einer ausführlichen Item- und Skalenanalyse eine erste revidierte Version des Fragebogens (FBMSN-R) konstruiert werden. Der FBMSN-R basiert auf einer faktoranalytisch ermittelten Struktur und liefert in seiner Skalenstruktur gute bis sehr gute Kennwerte im Hinblick auf die Skalenwerte und die innere Konsistenz. Außerdem konnten für die Skalen Streben nach sozialer Anerkennung, Soziale Vergleiche sowie Selbstkonzept und Selbstwert und Skala Konformitätsdruck und Gruppendynamik erste Befunde zur Bestimmung der Kriteriumsvalidität des FBMSN-R gesammelt werden. Allerdings stehen eine weitere Überprüfung und Optimierung der Skalenstruktur anhand einer weiteren Stichprobe noch aus. Zudem sollten zukünftige Untersuchungen den Fokus auf die Optimierung der Subskalen sowie der Generierung weiterer passender Items, um die genaue Interpretierbarkeit der Skalen zu fördern. Ferner sollten auch die Skalen, die im Rahmen der ersten Studie aufgrund mangelnder Operationalisierung und fehlender direkter thematischer Bezüge nicht vollständig verworfen werden. So wird auch hinsichtlich dieser Itemsätze eine tiefergehende Analyse sowie inhaltliche Überarbeitung und Anpassung empfohlen. Schließlich lässt sich in diesem Zusammenhang aufgrund der Themenbereiche, aber auch aufgrund der ersten Ergebnisse der Itemanalyse ein Potenzial vermuten, welches nicht vollständig ausgeschöpft worden ist. Entsprechend sollten die Untersuchungen und Weiterentwicklungen dieses Messinstrumentes auf Grundlage der Ergebnisse dieser Forschungsarbeit weiter vorangetrieben werden. Im Idealfall kann eine finale Version des Fragebogens konturiert werden, welche schließlich in der Medienerziehung und zu Präventionszwecken eingesetzt werden kann.

6 Gesamtfazit

Zusammenfassend lässt sich festhalten, dass der im Rahmen dieser Thesis überprüfte und überarbeitete Fragebogen noch keine fertige Version darstellt, welche für Forschungszwecke freigegeben werden kann. Allerdings konnte im Rahmen der ersten Studie eine umfassende Itemanalyse der Rohversion von Burdorf und Lux (2022) erfolgen, welche Aufschluss über die Trennschärfen, Itemvarianzen und Itemschwierigkeit gibt. Weiterhin konnten erste Ergebnisse für die interne Konsistenz bestimmt werden. Nach der ausführlichen Analyse und Überarbeitung des Fragebogens konnte anhand der exploratorischen Faktorenanalysen eine interpretierbare faktorielle Struktur ermittelt werden. Entsprechend wurden die folgenden sieben Faktoren extrahiert: Streben nach sozialer Anerkennung, Soziale Vergleiche, Konformitätsdruck und Gruppendynamik, Anschlussmotiv. Selbstkonzept und Selbstwert sowie Liebesbeziehungen und Freizügige Selbstdarstellung. Es stellte sich heraus, dass lediglich die Faktoren Streben nach sozialer Anerkennung, Soziale Vergleiche, Konformitätsdruck und Gruppendynamik, Anschlussmotiv sowie Selbstkonzept und Selbstwert im Hinblick auf die Operationalisierung interpretiert werden konnten. Für die Skalen Liebesbeziehungen und Freizügige Selbstdarstellung wurde eine vorläufige Interpretation vorgenommen, welche im Rahmen dieser Masterthesis nicht tiefergehender untersucht wurde. Stattdessen wird eine Reanalyse und Anpassung der Skalen sowie die Generierung neuer passender Items empfohlen. Im Rahmen der zweiten Studie konnten die guten Kennwerte der Itemanalyse sowie für die Reliabilitätskoeffizienten Cronbachs Alpha weitestgehend repliziert werden. Im nächsten Schritt erfolgte die Validierung der einzelnen Skalen anhand verschiedener Außenkriterien, welche parallel erhoben wurden. Die Kriterien wurden anhand theoretischer und inhaltlicher Bezüge zum FBMSN-R ausgewählt, um die Kriteriumsvalidität im Sinne einer Übereinstimmungsvalidität oder auch kongruenten Validität zu bestimmen. Entsprechend wurden verschiedene Hypothesen über die korrelativen Zusammenhänge aufgestellt und untersucht. Zusammenfassend lässt sich im Hinblick auf die Validierung der Skalen berichten, dass für die Skalen Streben nach sozialer Anerkennung, Soziale Vergleiche und Selbstkonzept und Selbstwert mittlere Werte für den Validitätskoeffizienten rtc feststellen lassen. Für die Skala Konformitätsdruck und Gruppendynamik konnten lediglich niedrige Werte erreicht werden. Außerdem konnten im Hinblick auf die Skala Anschlussmotiv keine Befunde zur Validierung ermitteltet werden. Abschließend lässt sich festhalten, dass eine erste revidierte Version des Fragebogens konstruiert werden konnte, welche nun im Rahmen weiterer Studien weiterentwickelt werden sollte. Weitere Forschungen sollten kurzfristig darauf ausgerichtet sein, die einzelnen Subskalen zu optimieren und weitere passende Items zu generieren. Mittelfristig sollte der Fokus auf Prüfung der faktoriellen Struktur und im Zuge dessen auf der

Untersuchung der Validität liegen. Langfristig sollte dann eine finale Version des Fragebogens zur Erfassung von zugrundeliegenden Motiven der Selbstdarstellung auf sozialen Netzwerken entwickelt werden, welcher in der Präventionsarbeit sowie in der Medienaufklärung praktische Anwendung finden kann.

Literaturverzeichnis

- Abel, J., Buff, C., & Burr, S. (2016). Social media an the fear of missing out: Scale development and assessment. *Journal of Business & Economics Research* (*JBER*), 14(1), 33-44.
- American Educational Research Association (AERA), American Psychological Association (APA) & National Council on Measurement in Education (NCME). (2014). Standards for educational and psychological testing. Washington, DC: American Psychological Association.
- Ardelt-Gattinger, E., Lechner, H., & Schlögl, W. (1998). *Gruppendynamik: Anspruch und Wirklichkeit der Arbeit in Gruppen*. Göttingen: Verlag für Angewandte Psychologie.
- Arte. (18. September 2019). *Instagram 1 Süchtig nach Dopamin*. Abgerufen am 4. November 2022 von https://www.arte.tv/de/videos/085801-004-A/dopamin/
- Asch, S. E. (1955). Options and social pressure. Scientific American, 193(5), 31-35.
- Asendorpf, J. B. (1989). Soziale Gehemmtheit und ihre Entwicklung. Berlin: Springer Verlag.
- Asendorpf, J. B., & van Aken, M. A. (1993). Deutsche Versionen der Selbstkonzeptskalen von Harter. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 25, 64-86.
- Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall.
- Benecke, C., & Brauner, F. (2017). *Motivation und Emotion*. Stuttgart: Kohlhammer Verlag.
- Bonfadelli, H., & Friemel, T. N. (2017). *Medienwirkungsforschung (6. Aufl.)*. Stuttgart: UTB.
- Bradlee, P. M., & Emmos, R. A. (1992). Locating narcissism within the interpersonal circumplex and the five-factor model. *Personality and Individual Differences*, 13, S. 821–830.
- Brandstätter, V., Schüler, J., Puca, R. M., & Lozo, L. (2018). *Motivation und Emotion.*Wiesbaden: Springer.
- Brandt, H. (2020). Exploratorische Faktorenanalyse. In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 575-612). Berlin: Springer.
- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Personality and Social Psychology*, *54*(1), 106-148.

- Bühner, M. (2011). Einführung in die Test- und Fragebogenkonstruktion (3. Aufl.). München: Pearson Studium.
- Burdorf, L. S., & Lux, L. (2022). Entwicklung und erste Überprüfung eines Fragebogens zur Motivation der Darstellung auf sozialen Medien. Bachelorarbeit. Köln: Hochschule Fresenius.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *103*(56), 2.
- Chou, H. G., & Edge, N. (2012). "They are happier and having better lives than I am": the impact of using Facebook on perceptions of others' lives. Cyberpsychology, Behavior and Social Networking, 15(2), 117–121.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.
- Collani, G. v., & Stürmer, S. (2009). SMSC. Fragebogen zum Verhalten in sozialen Situationen [Verfahrensdokumentation und Fragebogen]. In Leibniz-Institut für Psychologie (ZPID) (Hrsg.), *Open Test Archive*. Trier: ZPID.
- Collani, G., & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. Zeitschrift für Differentielle und Diagnostische Psychologie, 24, 3-7.
- Crusius, J. (2022). Soziale Vergleiche. In P. Ozimek, H.-W. Bierhoff, E. Rohmann, & S. Hanke (Hrsg.), *Angewandte Sozialpsychologie: Ein Lehrbuch.* Stuttgart: Kohlhammer.
- Deters, F., & Mehl, M. R. (2013). Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. *Social Psychological and Personality Science*, *4*(5), 579–586.
- Deutscher, S., & Mees, U. (2012). TEMEO-R. Fragebogen zur Erfassung telischer und metatelischer Orientierungen in revidierter Form [Verfahrensdokumentation, Autorenbeschreibung, Fragebogen und Auswertungshinweise]. In Leibniz-Institut für Psychologie (ZPID) (Hrsg.), Open Test Archive. Trier: ZPID.
- Dimbath, O. (2015). Juvenilität als Eskapismus. In S. Lessenich (Hrsg.), Routinen der Krise Krise der Routinen. Verhandlungen des 37. Kongresses der Deutschen Gesellschaft für Soziologie (S. 379-388). Trier: Universität Trier.
- Döring, N., & Bortz, J. (2016). Forschungsmethoden und Evaluation in den Sozialund Humanwissenschaften (5. Aufl.). Berlin-Heidelberg: Springer.
- Dufner, M., Rauthmann, J. F., Czarna, A. Z., & Denissen, J. J. (2013). Are narcissists sexy? Zeroing in on the effect of narcissism on shortterm mate appeal. *Personality and Social Psychology Bulletin*, 39, 870-882.

- Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden.*Weinheim: Beltz.
- Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2), 117-140.
- Fiedler, P. (2011). Persönlichkeitsstörungen. In H.-U. Wittchen, & J. Hoyer (Hrsg.), Klinische Psychologie und Psychotherapie (2. Aufl.) (S. 1101-1119). Berlin Heidelberg: Springer.
- Fisseni, H.-J. (2004). Lehrbuch der psychologischen Diagnostik: Mit Hinweisen zur Intervention. Göttingen: Hogrefe.
- Fogg, B. J. (2009). *A bahavoiral model for persuasive design.* Stanford: Persuasive Technology Lab.
- Gable, S. L. (2006). Approach and avoidance social motives and goals. *Journal of Personality*, 71, 175–222.
- Gäde, J. C., Schermelleh-Engel, K., & Werner, C. S. (2020). Klassische Methoden der Reliabilitätsschätzung. In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (3. Aufl.)* (S. 305-333). Berlin: Springer.
- Gerber, J., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years. *Psychological bulletin*, 144(2), 177-197.
- Germar, M. U. (2016). Wie die Mehrheit unsere Sicht auf die Ding verändern kann. Untersuchungen zu den Auswirkungen von Mehrheitseinfluss auf die neuro-kognitiven Prozesse visuellen Entscheidens. Hildesheim: Universitätsverlag.
- Gerrig, R. J. (2015). *Psychologie*. Hallbergmoos: Pearson.
- Goldhammer, F., & Harting, J. (2020). Testwertinterpretation, Testnormen und Testeichung. In H. Moosbruger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (3. Aufl.)* (S. 172-195). Berlin: Springer.
- Gonzales, A. L., & Hancock, J. T. (2011). Mirror, mirror on my Facebook wall: Effects of exposure to Facebook on self-esteem. *Cyberpsychology, Behavior and Social Networking*, *14*, 79-83.
- Guenther, C. L., & Alicke, M. D. (2010). Deconstructing the Better-Than-Average Effect. *Journal of Personality and Social Psychology*, 99(5).
- Haferkamp, N., & Krämer, N. C. (2011). Social comparison 2.0: Examining the effects of online profiles on social-networking sites. *Cyberpsychology, Behavior and Social Networking, 14*(5), 309–314.
- Hannich, H.-J. (2018). Individualpsychologie nach Adler. Stuttgart: Kohlhammer.

- Hartig, J., Frey, A., & Jude, N. (2020). Validität von Testwertinterpretationen. In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (3. Aufl.) (S. 529-546). Berlin: Springer.
- Heckhausen, H. (1980). Motivation und Handeln. Berlin: Springer.
- Heckhausen, J., & Heckhausen, H. (2018a). *Motivation und Handeln (5. Aufl.)*. Berlin Heidelberg: Springer.
- Heckhausen, J., & Heckhausen, H. (2018b). Motivation und Handeln: Einführung und Überblick. In J. Heckhausen, & H. Heckhausen (Hrsg.), *Motivation und Handeln (5. Aufl.)* (S. 1-12). Berlin: Springer.
- Hewstone, M., & Martin, R. (2014). Sozialer Einfluss. In K. Jonas, W. Stroebe, & M. Hewstone (Hrsg.), *Sozial Psychologie* (S. 270-312). Berlin Heidelberg: Springer.
- Ho, J. Y., & Dempsey, M. (2010). Viral marketing: Motivations to forward online content. *Journal of Business Research*, *63*(9-10), 1000–1006.
- Hofer, J., & Hagemeyer, B. (2018). Soziale Anschlussmotivation: Affiliation und Intimität. In J. Heckhausen, & H. Heckhausen (Hrsg.), *Motivation und Handeln* (5.Aufl.) (S. 223-244). Berlin: Springer.
- IBM. (o.J.). IBM SPSS-Software Bessere Geschäftsergebnisse durch Vorhersageanalysen. Abgerufen am 27. September 2022 von https://www.ibm.com/de-de/analytics/spss-statistics-software
- Impett, E. A., Strachman, A., Finkel, E. J., & Gable, S. L. (2008). Maintaining sexual desire in intimate relationships: The importance of approach goals. *Journal of Personality and Social Psychology*, *94*, 808–823.
- Janis, I. (1991). Groupthink. In E. Griffin (Hrsg.), *A First Look at Communication Theory*. New York: McGrawHill.
- Jha, S. (2010). Need for growth, achievement, power and affiliation: Determinants of psychological empowerment. *Global Business Review, 11*(3), 379-393.
- Kelava, A., & Moosbrugger, H. (2020a). Deskriptivstatistische Itemanalyse und Testwertbestimmung. In H. Moosbrugger, & A. Kelava (Hrsg.), Testtheorie und Fragebogenkonstruktion (3. Aufl.) (S. 143-158). Heidelberg: Springer.
- Kelava, A., & Moosbrugger, H. (2020b). Einführung in die Item-Response-Theorie (IRT). In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (3. Aufl.)* (S. 369-407). Berlin: Springer.
- Kemper, C. J., Lutz, J., & Neuser, J. (2011). Konstruktion und Validierung einer Kurzform der Skala Angst vor negativer Bewertung (SANB-5). Klinische Diagnostik und Evaluation, 4, S. 342-359.

- Kiesel, A., & Koch, I. (2012). *Lernen. Grundlagen der Lernpsychologie.* Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kim, J. W. (2014). Scan and click: The uses and gratifications of social recommendation systems. *Computers in Human Behavior*, 33, 184-191.
- Kneidinger-Müller, B. (2017). Identitätsbildung in sozialen Medien. In J.-H. Schmidt,
 & M. Taddicken (Hrsg.), *Handbuch Soziale Medien* (S. 62-77). Wiesbaden:
 Springer Fachmedien.
- Krämer, N. C., Eimler, S. E., & Neubaum, G. (2017). Selbstpräsentation und Beziehungsmanagement in sozialen Medien. In J.-H. Schmidt, & M. Taddicken (Hrsg.), *Handbuch Soziale Medien* (S. 41-55). Wiesbaden: Springer Fachmedien.
- Kuhlmann, C., & Gehrau, V. (2011). Auf der Flucht vor dem Tod? Eskapistische Mediennutzung und Narkotische Dysfunktion. *Publizistik*, *3*, 305–326.
- Lammers, C.-H., Vater, A., & Roepke, S. (2013). Narzisstische Persönlichkeitsstörung. *Der Nervenarzt 84*(7), 879-888.
- Leary, M. R. (1995). *Impression management and interpersonal behavior*. Madison: Brown & Benchmark Publishers.
- Leuchter, E., Ardelt-Gattinger, E. E., & Schlögl, W. (1998). *Gruppendynamik.*Anspruchund Wirklichkeit der Arbeit in Gruppen. Göttingen: Verlag für Angewandte Psychologie.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99.
- Manago, A., Graham, M. B., Greenfield, P., & Salimkhan, G. (2008). Self-presentation and gender on MySpace . *Journal of Applied Development Psychology*, 29, 446–458.
- Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross.cultural generalizability of the Internal/External Frame of Reference predictions across 26 countries. *Journal of Educational Psychology*, 96, 56-67.
- Maslow, A. H. (1981). *Motivation und Persönlichkeit.* Hamburg: Rowohlt Taschenbuch.
- McClelland, D. C. (1985). How motives, skills, and values determine what people do. *American Psychologist, 40*, 812-825.
- Milyavskaya, M., Saffran, M., Hope, N., & Koestner, R. (2018). Fear of missing out: prevalence, dynamics, and consequences of experiencing FOMO. *Motivation and Emotion*, *42*(3), 725–737.

- Moosbrugger, H., & Harting, J. (2003). Faktorenanalyse. In K. Kubinger, & R. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 137-145). Weinheim: Beltz.
- Moosbrugger, H., & Kelava, A. (2020a). Qualitätsanforderungen an Tests und Fragebogen ("Gütekriterien"). In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (3. Aufl.)* (S. 15-38). Berlin: Springer.
- Moosbrugger, H., & Kelava, A. (2020b). *Testtheorie und Fragebogenkonstruktion (3. Aufl.)*. Berlin: Spinger.
- Moosbrugger, H., Gäde, J. C., Schermelleh-Engel, K., & Rauch, W. (2020). Klassische Testtheorie (KTT). In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (3. Aufl.)* (S. 275-303). Berlin: Springer.
- Moosbrugger, H., Schermelleh-Engel, K., Gäde, J. C., & Kelava, A. (2020). Testtheorien im Überblick. In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion (3. Aufl.)* (S. 251-271). Berlin: Springer.
- Morf, C. C., & Koole, S. L. (2014). Das Selbst. In K. Jonas, W. Stroebe, & M. Hewstone (Hrsg.), *Sozialpsychologie* (S. 141-195). Berlin Heidelberg: Springer.
- Mulaik, S. A. (2010). Foundations of factor analysis (2nd ed.). Boca Raton, FL: Chapman & Hall.
- Murray, H. A. (1938). Explorations in personality. New York: Oxford University Press.
- Neyer, F. J., & Asendorpf, J. B. (2018). *Psychologie der Persönlichkeit.* Berlin: Springer.
- Nijstad, B. A., & van Knippenberg, D. (2014). Gruppendynamik. In K. Jonas, W. Stroebe, & M. Hewstone (Hrsg.), *Sozialpsychologie* (S. 439-467). Berlin Heidelberg: Springer.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402.
- Orlowski, J., Coombe, D., Curtis, V. (Produzenten), & Orlowski, J. (Regisseur). (2020). *The Social Dilemma* [Dokumentarfilm]. USA: Netflix.
- Ortega, R., Paz Elipe, J. A., Mora-Merchán, J., & Vega, E. (2009). The emotional impact on victims of traditional bullying and cyberbullying. A study of Spanish adolescents. *Journal of Psychology*, *217*(4), 197–204.
- Peter, C., Fahr, A., & Früh, H. (2012). Im Spiegel der Anderen. *Publizistik 57*, 161-178.

- Pospeschill, M. (2022). *Testtheorie, Testkonstruktion, Testevaluation.* München: Ernst Reinhardt Verlag.
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29, S. 1814-1848.
- Rammsayer, T., & Weber, H. (2016). Differentielle Psychologie Persönlichkeitstheorien. Göttingen: Hogrefe.
- Renner, K.-H., Heydasch, T., & Ströhlein, G. (2012). Forschungsmethode der Psychologie: Von der Fragestellung zur Präsentation. Wiesbaden: Springer.
- Richter, L., Paier, D., & Reiger, H. (2021). *Qualitative Sozialforschung Eine Einführung.* Wien: Facultas.
- Satow, L. (2020). *B5T*® *Big-Five-Persönlichkeitstest: Test- und Skalendokumentation.* Abgerufen am 15. August 2022 von Dr. Satow: https://www.drsatow.de/tests/persoenlichkeitstest/
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretation. *Review of Education Research*, 407-441.
- Siibak, A. (2009). Constructing the self through the photo selection Visual impression management on social networking websites. *Cyberpsychology:*Journal of Psychosocial Research on Cyberspace, 3(1), 1.
- Sokolowski, K., & Heckhausen, H. (2006). Soziale Bindung: Anschlussmotivation und Intimitätsmotivation. In J. Heckhausen, & H. Heckhausen (Hrsg.), *Motivation und Handeln*. Berlin, Berlin: Springer.
- Steers, M., Wickham, R., & Acitelli, L. (2014). Seeing Everyone Else's Highlight Reels: How Facebook Usage Is Linked to Depressive Symptoms. *Journal of Social and Clinical Psychology*, *33*(8), 701-731.
- Stiller, J., & Alfermann, D. (2005). Selbstkonzept im Sport. Zeitschrift für Sportpsychologie, 12(4), 119-126.
- Sumter, S. R., Baumgartner, S. E., Valkenburg, P. M., & Peter, J. (2012). Developmental trajectories of peer victimization: Off-line and online experiences during adolescence. *Journal of Adolescent Health*, *50*(6), 607–613.
- Valkenburg, P. M., & Peter, J. (2009). Social consequences of the Internet for adolescents: A decade of research. *Current Directions in Psychological Science*, 18(1), 1-5.
- Walden, T. (2020). Theorieansätze und Hypothesen in der Medienpädagogik: Das Eskapismus-Konzept. In U. Sander, F. von Gross, & K.-U. Hugger (Hrsg.), Handbuch Medienpädagogik. Wiesbaden: Springer.

Wittchen, H.-U., & Jacobi, F. (2011). Epidemiologische Beiträge zur Klinischen Psychologie. In H.-U. Wittchen, & J. Hoyer (Hrsg.), *Klinische Psychologie und Psychotherapie (2. Aufl.)* (S. 57-89). Berlin Heidelberg: Springer.

Anhang

Anhangsverzeichnis

Anhang A: Items und Fragebögen
Anhang B: Ergebnisse – Studie I
Anhang C: Ergebnisse - Studie II
Anhang D: Exploratorische Faktorenanalysen: Rohversion des FBMSN

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe. Die Stellen der Arbeit, einschließlich der Tabellen und Abbildungen, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen sind, habe ich in jedem einzelnen Fall kenntlich gemacht und die Herkunft nachgewiesen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen und wurde noch nicht veröffentlicht. Die im Rahmen der Arbeit ausgewerteten Daten habe ich weder erfunden noch gefälscht. Sie wurden wie dokumentiert eigenständig erhoben bzw. mir aus zuverlässiger und nachvollziehbar kenntlich gemachter Quelle für diese Neuoder Reanalyse zur Verfügung gestellt. Ich verpflichte mich, die Originaldaten den Prüferinnen auf Rückfrage für einen Zeitraum von fünf Jahren ab Abgabedatum zugänglich zu halten.

Köln, 21.11.2022

Luca Mario Derjung