Empirisches Vorgehen

6

Die vorliegende Arbeit hat zum Ziel, herauszufinden, wie sich informativer und unterhaltender Brand Content auf Markenimagedimensionen auswirkt.¹ Im vorherigen Kapitel wurde der aktuelle Forschungsstand mit Bezug zu dieser Frage diskutiert. Bestehende Forschung hat gezeigt, dass die Wirkung von Brand Content auch vom Kommunikationskanal und von der Art Marke, die den Brand Content ausspielt, abhängt. Unterschiedliche Arten von Brand Content sind unterschiedlich erfolgreich, je nachdem über welche Markenpositionierung oder über welches Produktinvolvement eine Marke verfügt. Diverse der entwickelten Hypothesen postulieren kausale Zusammenhänge. Zur Überprüfung *kausaler Zusammenhänge* eignen sich insbesondere wissenschaftliche *Experimente* (Brosius, Haas & Koschel, 2012, S. 197; Kline, 2016, S. 123). Der empirischen Studie liegt ein Experiment zugrunde, für welches der Brand Content (informativ vs. unterhaltend), der Kommunikationskanal (Informations- vs. Unterhaltungserwartungen) und die Marke (funktional vs. emotional und High vs. Low Involvement) manipuliert werden. Das detaillierte Forschungsdesign findet sich in Abschnitt 6.6.

¹ Ergänzende Inhalte zu diesem Kapitel finden sich im elektronischen Zusatzmaterial.

Elektronisches Zusatzmaterial Die elektronische Version dieses Kapitels enthält Zusatzmaterial, das berechtigten Benutzern zur Verfügung steht
https://doi.org/10.1007/978-3-658-35711-5_6.

6.1 Befragung mit experimentellem Design

Das wissenschaftliche Experiment ist eine «Form der Untersuchungsanlage» und «keine Methode der Datenerhebung» (Brosius et al., 2012, S. 197). Das *Experiment* wird deshalb mit einer Erhebungsmethode wie einer *Befragung* kombiniert (ebd.). Mit Experimenten wird der Einfluss einer unabhängigen Variable auf eine abhängige Variable getestet (Antonakis, Bendahan, Jacquart & Lalive, 2010). Bei Experimenten im Rahmen der Kommunikationswirkungsforschung wird ein *Stimulus manipuliert*, d. h., er wird in verschiedenen Ausprägungen gestaltet und unterschiedlichen Experimentalgruppen vorgelegt bzw. bei diesen angewandt. Die unterschiedlichen Ausprägungen des Stimulus dienen als unabhängige Variable, die auf eine abhängige Variable einwirkt. Wird eine abhängige Variable – unter kontrollierten Experimentalbedingungen – auf Basis unterschiedlicher Stimuli unterschiedlich eingeschätzt, dann wird davon ausgegangen, dass dieser Effekt durch die Manipulation zustande gekommen ist. Erhält eine Person (oder Personengruppe) über den Lauf der Zeit mehrere unterschiedliche Stimuli mit jeweils nachgelagerter Messung der abhängigen Variable, wird von einem within-subject-Experiment gesprochen. Erhalten dagegen unterschiedliche Personengruppen je einen Stimulus, handelt es sich um ein between-subject-Experiment. Diejenigen Personen, die denselben Stimulus zu sehen bekommen, bilden eine Experimentalgruppe. Eine Gruppe, die keinen Stimulus zu sehen bekommt, ist eine Kontrollgruppe (Brosius et al., 2012; Diekmann, 2010).

Im Rahmen dieser Arbeit wird ein *between-subject-Experiment mit 16 verschiedenen Experimentalgruppen* durchgeführt. Die 16 Experimentalgruppen bekommen je einen von 16 unterschiedlichen Markenbotschaften (Brand Content) zu sehen. Es gibt keine Kontrollgruppe. Die 16 Markenbotschaften unterscheiden sich in vier Faktoren: Brand Content, Kanal, Markenpositionierung, Produktinvolvement. Es handelt sich entsprechend um ein 2 (informativer Brand Content, unterhaltender Brand Content) $\times 2$ (Kanal mit primären Informationserwartungen, Kanal mit primären Unterhaltungserwartungen) $\times 2$ (funktionale Markenpositionierung, emotionale Markenpositionierung) $\times 2$ (Low-Involvement-Produkt, High-Involvement-Produkt)-faktorielles Design eines between-subject-Experiments (K.-J. Chen et al., 2015; Diekmann, 2010, S. 344; Koch et al., 2019, S. 68–77). Die beiden Faktoren Markenpositionierung und Produktinvolvement sollen durch vier Marken abgebildet werden: Marke A (funktionale Markenpositionierung / tiefes Produktinvolvement), Marke B (funktional / hoch), Marke C (emotional / tief) und Marke D (emotional / hoch).

Das Experiment kann mit einer *bewusst ausgewählten Stichprobe* durchgeführt werden. Wichtig ist, dass sich die Experimentalgruppen nur bezüglich der unabhängigen Variable bzw. der getätigten Manipulation unterscheiden (Brosius et al., 2012, S. 204). Dadurch kann der Einfluss von Störvariablen auf die abhängige Variable ausgeschlossen werden und unterschiedliche Ausprägungen einer abhängigen Variable auf unterschiedliche Ausprägungen der unabhängigen Variable zurückgeführt werden. In diesem Zusammenhang muss auch eine Konfundierung ausgeschlossen werden. Eine solche liegt vor, wenn sich die Manipulation auf zwei Ebenen unterscheidet (Brosius et al., 2012, S. 201). Wird beispielsweise einer Experimentalgruppe ein unterhaltendes Video und einer zweiten Gruppe ein informativer Text vorgelegt, dann können unterschiedliche Bewertungen einer abhängigen Variable nicht eindeutig dem Inhalt oder der Form zugeteilt werden (Brosius et al., 2012). Nach Diekmann (2010) müssen drei Bedingungen erfüllt sein, um von einem Experiment zu sprechen:

1. Es werden mindestens zwei experimentelle Gruppen gebildet.
2. Die Versuchspersonen werden den experimentellen Gruppen nach einem Zufallsverfahren zugewiesen (Randomisierung).
3. Die unabhängige Variable wird vom Forscher manipuliert. (Diekmann, 2010, S. 337)

Die *Randomisierung* neutralisiert den Einfluss von Drittvariablen, weil diese in den Versuchsgruppen dieselbe Verteilung aufweisen (Diekmann, 2010, S. 339). Antonakis et al. (2010) sehen in randomisierten Experimenten eine «fehler-sichere» (S. 1086) Variante, um Kausalbeziehungen zu überprüfen. Wichtig ist schliesslich, dass die Manipulation im Experiment vorkommt, bevor die abhängige Variable erhoben wird (Kline, 2016, S. 123).

6.2 Gütekriterien empirischer Forschung

Wissenschaftliche Forschung ist eine Annäherung an die Realität. Damit diese Annäherung möglichst gut gelingt, muss die Forschung zwei grundlegende Gütekriterien erfüllen: *Reliabilität* und *Validität* (Koch et al., 2019, S. 189).

Bei der *Reliabilität* handelt es sich um die Zuverlässigkeit einer Messung. Wird eine Messung wiederholt, dann muss diese unter denselben Bedingungen zu denselben Resultaten kommen wie die erste Messung (Brosius et al., 2012, S. 49). Nach Brosius et al. (2012, S. 53) muss ein Messinstrument unabhängig vom Forscher und unabhängig von Einflüssen auf die Befragten funktionieren.

Die Unabhängigkeit der Resultate vom durchführenden Forscher wird teilweise auch als eigenständiges Qualitätskriterium der Objektivität geführt (Diekmann, 2010, S. 249). Im Rahmen dieser Arbeit muss insbesondere sichergestellt werden, dass die verwendeten Fragen und Variablen die zu messenden Konstrukte gut abbilden (Koch et al., 2019, S. 171). Detaillierte Ausführungen dazu finden sich im Abschnitt 6.4.4.

Bei der *Validität* handelt es sich um die Gültigkeit der Messung. Eine Messung ist dann valide, wenn «man das misst, was man messen will» (Brosius et al., 2012, S. 63). Bei Experimenten unterscheidet man zwischen *interner und externer Validität* (Koch et al., 2019, S. 52). Ein experimentelles Untersuchungsdesign besitzt hohe interne Validität, wenn der Ablauf des Experiments in hohem Grad standardisiert und kontrolliert ist. Der ungewollte Einfluss von Störvariablen kann dadurch minimiert werden. Bei externer Validität hingegen können die gewonnenen Erkenntnisse zu den Kausalzusammenhängen über die «experimentelle Versuchsanlage» (S. 52) hinaus verallgemeinert werden (Koch et al., 2019). Hohe interne Validität bei hoher Kontrolle geht meist zu Lasten der externen Validität und vice versa. Im Rahmen dieses Experiments wurde aus Gründen der externen Validität mit real existierenden Schweizer Marken gearbeitet. Die Teilnehmenden verfügen jedoch über unterschiedliche Assoziationen mit diesen realen Marken, die über die experimentelle Manipulation (Markenpositionierung und Produktinvolvement) hinausgehen. Dieser Umstand kann nicht vollständig kontrolliert werden und geht zu Lasten der internen Validität. Andererseits wurden die experimentellen Stimuli nicht in einen realen Twitter- oder Instagram-Feed eingebaut, weil andere Posts die Wirkung der Stimuli hätten beeinflussen können. Diese Entscheidung erhöht die interne Validität zu Ungunsten der externen Validität. Koch et al. (2019, S. 53) stellen klar, dass die Balance zwischen interner und externer Validität Erwägungssache des Forschenden ist. Neben diesen grundlegenden Überlegungen zur Validität müssen auch die verwendeten Fragen und Indikatoren (siehe Abschnitt 6.4.4) valide sein. Das heisst, die Variablen müssen denjenigen Sachverhalt messen, der tatsächlich erhoben werden soll. Auf die Validität auf Stufe der Variablen wird ebenfalls ausführlich in Abschnitt 6.8 eingegangen.

6.3 Erfassung latenter Variablen

Die Variablen im Rahmen dieser Arbeit sind *latent*, das heisst, sie können nicht direkt beobachtet oder gemessen werden (z. B. Markenimage). Zu deren Erfassung muss auf *messbare oder beobachtbare Indikatoren* zurückgegriffen werden, welche die latente Variable (Konstrukt) definieren oder bestimmen (Brosius et al., 2012, S. 31).

Indikatoren können in einem formativen oder reflektiven Verhältnis zum Konstrukt stehen (Jarvis, MacKenzie & Podsakoff, 2003). Bei *formativen Modellen* sind die Indikatoren integraler Baustein des Konstrukts. Sie bilden oder formen das Konstrukt bzw. machen dieses aus. Diamantopoulos und Winklhofer (2001) schreiben zu formativen Modellen: «the items used as indicators must cover the entire scope of the latent variable» (S. 271). Die Indikatoren also müssen das Konstrukt in der Gesamtheit abdecken und können nicht ausgetauscht oder entfernt werden. Die Indikatoren wirken einzeln auf das Konstrukt und korrelieren nicht zwingend untereinander (Jarvis et al., 2003). Der «sozioökonomische Status» und dessen Indikatoren Bildung, Beruf, Einkommen und Wohnsitz können als formatives Konstrukt verstanden werden (Diamantopoulos & Winklhofer, 2001). Bei *reflektiven Modellen* steht das Konstrukt am Ausgangspunkt und wirkt auf die Indikatoren ein. Eine Veränderung am Konstrukt führt zu einer Veränderung bei allen Indikatoren. Die Indikatoren korrelieren untereinander. Sie sind nicht Baustein, sondern Resultat des Konstrukts. Dadurch können sie prinzipiell ausgetauscht oder entfernt werden (Jarvis et al., 2003). Die «Kundenzufriedenheit» mit den Indikatoren «Weiterempfehlung» oder «Wiederkauf» kann hier als Beispiel dienen (Eberl, 2004).

Die *latenten Variablen* im Rahmen dieser Arbeit sind allesamt reflektiver Art. Die Einstellung gegenüber dem Brand Content, die Einstellung gegenüber dem Kanal, aber auch das Markenimage oder markenorientierte Verhaltensvariablen sind als «Einstellungen» zu verstehen (siehe Abschnitt 2.3.3). Und Einstellungen werden typischerweise reflektiv konstituiert (Jarvis et al., 2003, S. 201; Teichert, Effertz, Tsoi & Shchekoldin, 2015, S. 700). Auch Gratifikationen auf Ebene des Brand Contents (Jahn & Kunz, 2012) oder auf Ebene des Kanals (Ruehl & Ingenhoff, 2017) werden reflektiv gebildet. Die reflektive Operationalisierung der Konstrukte muss Anforderungen an Reliabilität und Validität genügen. In Abschnitt 6.4.4 wird aufgezeigt, anhand welcher Vorgehensweisen und Kriterien die Konstrukte dieser Arbeit überprüft werden. In Abschnitt 6.8 werden die Konstrukte vorgestellt und überprüft.

6.4 Strukturgleichungsmodell zur Überprüfung der Zusammenhänge

Um die durch die experimentelle Befragung gewonnenen Daten zu analysieren, wird ein *Strukturgleichungsmodell* gerechnet. Strukturgleichungsmodelle vereinen Faktorenanalysen und Regressions-/Pfadanalysen und sind damit speziell geeignet, um Zusammenhänge zwischen latenten Variablen zu untersuchen

(Ingenhoff & Buhmann, 2016; Werner, Schermelleh-Engel, Gerhard & Gäde, 2016). Strukturgleichungsmodelle haben gegenüber traditionellen Auswertungsmethoden (z. B. Varianzanalyse) den Vorteil, dass Messfehler berücksichtigt werden und komplexe Zusammenhänge abgebildet werden können (MacKenzie, 2001).

Wichtig ist zu erwähnen, dass Strukturgleichungsmodelle *keine Aussagen zur Kausalität* ermöglichen (Rost, 2013, S. 166). Kausale Effekte können nur durch ein experimentelles Design überprüft und sichergestellt werden. Wo im Rahmen dieser Arbeit eine abhängige Variable durch eine experimentelle Manipulation beeinflusst wird, kann von Kausalität ausgegangen werden – sofern die in Abschnitt 6.1 dargelegten Erfordernisse erfüllt sind. Bei Zusammenhängen zwischen zwei latenten Variablen, bei denen die unabhängige Variable nicht experimentell manipuliert wurde, kann einzig von einem positiven Zusammenhang im Sinne einer Korrelation ausgegangen werden. In den folgenden Kapiteln wird auf die Eigenheiten von Strukturgleichungsmodellen eingegangen und aufgezeigt, wie Effekte zwischen Variablen getestet werden.

6.4.1 Messmodell und Strukturmodell

Strukturgleichungsmodelle bestehen aus *Messmodellen* und aus *Strukturmodellen* (Ingenhoff & Buhmann, 2016). Messmodelle geben den Zusammenhang zwischen einer latenten Variable (Konstrukt) und den beobachteten Variablen (Indikatoren) wieder, während Strukturmodelle die Zusammenhänge zwischen latenten Variablen abbilden (Werner et al., 2016). Bevor ein Strukturgleichungsmodell gerechnet wird, müssen die verwendeten latenten Variablen mittels einer konfirmatorischen Faktorenanalyse (Confirmatory Factor Analysis, CFA) auf ihre Reliabilität und Validität überprüft werden (Hair, Babin & Krey, 2017).

Im Rahmen dieser Arbeit wird auf die Notation von Mplus zurückgegriffen (Kleinke, Schlüter & Christ, 2017; Muthén & Muthén, 2017; Urban & Mayerl, 2014, S. 235). Die wesentlichen Begriffe sind in Tabelle 6.1 festgehalten.

In Abbildung 6.1 ist ein prototypisches Strukturgleichungsmodell mit drei latenten Variablen skizziert. Dabei sind f_1 und f_2 exogene latente Variablen, f_3 hingegen eine endogene latente Variable. Latente Variablen werden als Ellipsen, beobachtete Variablen als Rechtecke dargestellt. Einseitige Pfeile ($\rightarrow$) sind Faktorladungen oder Pfadkoeffizienten, doppelseitige Pfeile ($\leftrightarrow$) sind Kovarianzen bzw. Korrelationen. Messfehler von Indikatoren geben denjenigen Varianzanteil des Indikators wieder, der nicht durch die latente Variable erklärt wird (Urban & Mayerl, 2014, S. 17). Residuen latenter Variablen geben in analoger Weise den

Tabelle 6.1 Verwendete Symbole bei Strukturgleichungsmodellen

Symbol	Name	Erklärung
f	/	Latente Variable
y	/	Indikator (beobachtete Variable) einer latenten Variable
x	/	Exogene, nicht-latente Variable
λ	lambda	Faktorladung eines Indikators
β	Beta	Pfadkoeffizient zwischen zwei Variablen, wobei β = standardisiert und β_u = unstandardisiert
$\leftarrow / (e)$	/	Messfehler von Indikatoren
$\leftarrow / (d)$	/	Residuen von latenten Variablen (disturbances)
ψ	Psi	– Kovarianz zweier exogener Variablen – Kovarianz zwischen Residuen zweier latenter Variablen
θ	Theta	Kovarianz der Messfehler zweier Indikatoren
R^2		Erklärte Varianz einer latenten Variable

Anmerkungen. Symbole in Anlehnung an Muthén und Muthén (2017), Urban und Mayerl (2014) und Kleinke, Schlüter und Christ (2018).

Varianzanteil einer latenten endogenen Variable an, der nicht durch andere im Modell abgebildete Variablen erklärt wird (Urban & Mayerl, 2014, S. 26).

Abbildung 6.1 Beispielhafte Darstellung eines Strukturgleichungsmodells. (Nach Urban und Mayerl (2014))

6.4.2 Kovarianz- versus varianzbasierte Modelle

Es kann zwischen einem *kovarianzbasierten* und einem *varianzbasierten* Ansatz zur Berechnung von Strukturgleichungsmodellen unterschieden werden (Ingenhoff & Buhmann, 2016). Kovarianzbasierte Ansätze versuchen die «empirische Kovarianzmatrix der Indikatoren» (S. 37) möglichst präzise zu reproduzieren und schätzen so die Modellparameter (Herrmann, Huber & Kressmann, 2006). Varianzbasierte Ansätze versuchen die Indikatorenwerte zu reproduzieren, indem die Varianzen minimiert werden (Herrmann et al., 2006). Varianzbasierte Ansätze finden insbesondere bei formativen Messmodellen breiten Einsatz (Ingenhoff & Sommer, 2010). Nach Herrmann und Kollegen (2006) eignen sich hingegen beide Verfahren gleichermaßen für die Berechnung von Modellen mit rein reflektiven Konstrukten (S. 52). Die Entscheidung zugunsten eines Verfahrens hängt damit von «verfahrensspezifischen Erfordernissen» (S. 52) und der Forschungsmotivation ab. Die Vorteile varianzbasierter Verfahren liegen nach Herrmann et al. (2006) darin, dass (a) eine Schätzung auch mit kleineren Stichproben möglich ist, (b) die analysierten Daten nicht normalverteilt sein müssen und (c) eine identische Verteilung der Residuen (Homoskedastizität) nicht notwendig ist. Herrmann et al. (2006) raten, bei «managementorientierten Problemstellungen» (S. 45), bei denen abhängige Variablen – oder Effekte auf diese – erklärt werden sollen, auf ein varianzbasiertes Verfahren zurückzugreifen. Eine varianzbasierte Methode bietet sich zudem an, wenn wenig theoretisches Vorwissen über die Zusammenhänge der latenten Variablen verfügbar ist (Herrmann et al., 2006, S. 41).

Kline (2016, S. 361) weist hingegen darauf hin, dass varianzbasierte Verfahren gerade bei kleineren Stichproben zu ungenauen und verzerrten Schätzungen führen können im Vergleich zu kovarianzbasierten Verfahren. Verschiedene Forscher sehen gar keine nennenswerten Vorteile eines varianzbasierten Verfahrens gegenüber einem kovarianzbasierten Verfahren und raten generell vom varianzbasierten Verfahren ab (Antonakis et al., 2010; Rönkkö & Evermann, 2013).

Während bei varianzbasierten Verfahren der Fokus der Analyse auf partiellen Beziehungen liegt, steht ein Theoriegefüge im Interesse kovarianzbasierter Verfahren. Sollen theoretisch fundierte Hypothesen geprüft werden, dann eignet sich das kovarianzbasierte Verfahren (Herrmann et al., 2006, S. 45). Im Vergleich zu varianzbasierten Verfahren weisen kovarianzbasierte Verfahren die Gesamtgüte eines postulierten Modells aus (Antonakis et al., 2010, S. 1103). Neuere Schätzmethoden verlangen zudem auch bei kovarianzbasierten Methoden keine Normalverteilung der Daten (Muthén & Muthén, 2017, S. 9). Im Rahmen dieser Arbeit wird ein kovarianzbasiertes Verfahren angewandt, weil nur reflektive latente Variablen vorkommen und ein Gesamtmodell zur Wirkung von

Brand Content (BCBI-Modell) getestet werden soll, das auf einer theoretischen Grundlage hergeleitet wurde.

6.4.3 Modellgüte von Strukturgleichungsmodellen

Bei einem kovarianzbasierten Vorgehen wird die empirische *Kovarianzmatrix*, die sich aus den erhobenen Daten ergibt, mit einer implizierten Kovarianzmatrix, die dem postulierten Modell folgt, gegenübergestellt (Bühner, 2011, S. 382). Je besser das postulierte Modell die Daten abbildet, desto kleiner ist die Diskrepanz zwischen der empirischen und der implizierten Kovarianzmatrix und desto besser die Modellgüte (Bühner, 2011, S. 382). Verschiedene Indikatoren zur *Modellgüte* geben an, wie gut ein postuliertes Modell zu den erhobenen empirischen Daten passt. Gemäss APA sollen bei einer konfirmatorischen Faktorenanalysen sechs Indikatoren in Bezug auf die Modellgüte angegeben werden: df , χ^2 , χ^2/df , AGFI, ECVI und RMSEA (Nicol & Pexman, 2010, S. 129). Bei einem Strukturgleichungsmodell werden df , χ^2 , GFI, CFI, RMSR und IFI empfohlen. Allerdings weichen diese Empfehlungen von denjenigen in den einschlägigen Handbüchern ab. Bereits Hu und Bentler (1999, S. 5) raten beispielsweise davon ab, die Modellgüte bei einer Maximum-Likelihood-Schätzung an AGFI festzumachen, weil der Wert von der Grösse der Stichprobe abhängt.

Kline (2016, S. 269) rät χ^2 (Chi-Quadrat), die dazugehörigen Freiheitsgrade df und den dazugehörigen p -Wert als Resultat der Teststatistik anzugeben. Weiter sollen die Indikatoren RMSEA, CFI und SRMR ausgewiesen werden. Die Beurteilung der Modellgüte folgt im Rahmen dieser Arbeit dieser Empfehlung. Alle Werte werden standardmässig von Mplus berechnet. Zusätzlich wird der Indikator TLI, ebenfalls ein standardmässiger Output in Mplus, genannt. Nach Byrne (2012, S. 70) handelt es sich beim TLI um einen oft genutzten Indikator.

Die Freiheitsgrade (df) geben die Differenz zwischen bekannten und zu schätzenden Parametern an (Bühner, 2011, S. 402). Nur wenn ein postuliertes Modell einen positiven Freiheitsgrad hat, also über mehr bekannte als unbekannte Parameter verfügt, kann das Modell identifiziert und berechnet werden. Die Anzahl der bekannten Parameter kann mit $v(v + 1)/2$ bestimmt werden, wobei v die Anzahl erhobener Variablen angibt (Kline, 2016, S. 127). Als frei gelten diejenigen Parameter, die im Rahmen der Modellberechnung geschätzt oder durch den Forscher a priori fixiert werden. Dazu gehören Pfadkoeffizienten, Faktorladungen, Varianzen der unabhängigen Variablen inklusive der Messfehler und Residuen, sowie Kovarianzen zwischen exogenen Variablen, zwischen Residuen latenter Variablen (disturbances) oder zwischen Messfehlern der Indikatoren (Urban & Mayerl, 2014, S. 25–28).

Dadurch, dass mehr bekannte als unbekannte Parameter vorliegen, kann keine eindeutige Lösung für ein postuliertes Modell gefunden werden. Deshalb werden die zu schätzenden Modellparameter so bestimmt, dass die daraus folgende implizierte Kovarianzmatrix möglichst nahe der empirischen Kovarianzmatrix liegt (Bühner, 2011, S. 405). Je besser das gelingt, desto höher der Modellfit. Um die unbekannten Parameter und dadurch die implizierte Kovarianzmatrix zu schätzen, können verschiedene Verfahren angewandt werden (Bühner, 2011, S. 407). Die bekannteste Schätzmethode ist *Maximum Likelihood* (ebd. S. 408). Diese wird auch im Rahmen dieser Arbeit verwendet.

χ^2 gibt als Gesamtmasse den Unterschied zwischen implizierter und empirischer Kovarianzmatrix an (Bühner, 2011, S. 407). Bei einem signifikanten χ^2 -Wert sind die Unterschiede zwischen implizierter und empirischer Kovarianzmatrix nicht zufällig und das postulierte Modell grundsätzlich ungenügend. Setzt man den χ^2 -Wert ins Verhältnis zu den Freiheitsgraden (χ^2/df), erhält man einen Quotienten, der den χ^2 -Wert als «deskriptives Gütemass» (S. 205) ausgibt (Weiber & Mülhhaus, 2014). Im Idealfall ist der Quotient kleiner als 5.0 (West, Taylor & Wu, 2012, S. 212).

Neuere Studien zeigen, dass der χ^2 -Wert mit Bedacht betrachtet werden muss, weil er bei grossen Stichproben zu einem tiefen p -Wert tendiert und damit Modelle als ungenügend taxiert, die prinzipiell in Ordnung sind (Ciecuch & Davidov, 2016, S. 10; West et al., 2012, S. 211). Byrne (2012, S. 69) hält dazu fest, dass die Güte von Strukturgleichungsmodellen heute beinahe ausschliesslich an elaborierteren Gütemassen ausgemacht wird. Bei den elaborierteren Gütemassen, die nicht von der Stichprobengrösse abhängen, kann zwischen absoluten und komparativen Gütemassen unterschieden werden (Bühner, 2011, S. 424). *Absolute Gütemasse* vergleichen das postulierte Modell mit einem saturierten Modell. Im saturierten Modell korrelieren alle Variablen untereinander (Bühner, 2011, S. 424). Es bildet die erhobenen Daten auf die bestmögliche Art und Weise ab, verfügt über 0 Freiheitsgrade und einen χ^2 -Wert von 0 (Weiber & Mülhhaus, 2014, S. 213). Absolute Gütemasse können wiederum in «Goodness-of-Fit-Indizes» und «Badness-of-Fit-Indizes» unterschieden werden (Bühner, 2011, S. 424). Während Erstere zeigen, wie gut ein Modell ist, geben Letztere darüber Auskunft, wie schlecht dieses ist.

Der RMSEA (Root Mean Square Error of Approximation) ist ein Badness-of-Fit-Index. Während der χ^2 -Wert prüft, ob ein Modell richtig ist, gibt der RMSEA darüber Auskunft, wie gut ein Modell die Realität approximieren kann. Der RMSEA ist dadurch weniger streng als der χ^2 -Wert (Weiber & Mülhhaus, 2014, S. 204). Beim RMSEA wird üblicherweise das Konfidenzintervall mit angegeben. Der SRMR (Standardized Root Mean Square Residual) ist ebenfalls ein Badness-of-Fit-Index, der die Differenz zwischen beobachteten und durch das postulierte Modell vorhergesagten Korrelationen untersucht (Kline, 2016, S. 277).

Der CFI und der TLI sind *komparative Gütemasse* (Byrne, 2012, S. 70). Komparative Gütemasse vergleichen das postulierte Modell mit einem Nullmodell (Bühner, 2011, S. 424). Beim Nullmodell werden alle Kausalbeziehungen und Kovarianzen per Definition auf 0 gesetzt (Urban & Mayerl, 2014, S. 16). Ein postuliertes Modell sollte die Daten zwingend besser abbilden als ein Nullmodell. Entsprechend ist der χ^2 -Wert des Nullmodells im Normalfall höher als derjenige des zu testenden Modells (Byrne, 2012, S. 67). Der CFI (Comparative Fit Index) ist ein standardisierter komparativer Indikator, dessen Werte zwischen 0 (schlechtes Modell) und 1 (gutes Modell) liegen. Der TLI (Tucker Lewis Fit Index) ist nicht standardisiert, seine Werte liegen jedoch meist ebenfalls zwischen 0 und 1 (Byrne, 2012, S. 71). Im Vergleich zum CFI bestraft der TLI überkomplexe Modelle, wenn spezifizierte Abhängigkeiten nur marginal zur Modellgüte beitragen (ebd.). Tabelle 6.2 gibt einen Überblick der in dieser Arbeit verwendeten Gütemasse.

Tabelle 6.2 Kriterien zur Beurteilung der Güte von Strukturgleichungsmodellen

Kriterium der Modellgüte	Schwellenwert
χ^2	je kleiner, desto besser
χ^2/df	$\leq 3.0^a$ bzw. $\leq 5.0^b$
P	$p < 0.05$
CFI	≥ 0.9 (gut) bzw. ≥ 0.95 (sehr gut)
TLI	≥ 0.9 (gut) bzw. ≥ 0.95 (sehr gut)
RMSEA	≤ 0.08 (gut) bzw. ≤ 0.05 (sehr gut)
SRMR	≤ 0.08

Anmerkungen. Die Schwellenwerte basieren auf den Empfehlungen von Hair, Black, Babin und Anderson (2014, S. 584), Weiber und Mühlhaus (2014, S. 203–222) und West, Taylor und Wu (2012, S. 212–213).

^a Weiber und Mühlhaus (2014). ^b West, Taylor und Wu (2012).

6.4.4 Prüfung reflektiver Messmodelle

Zur Überprüfung der Reliabilität und Validität von latenten Variablen wird eine konfirmatorische Faktorenanalyse durchgeführt. Bei einem konfirmatorischen Vorgehen wird eine auf Basis von Vorwissen angenommene Zuordnung von Indikatoren zu latenten Konstrukten geprüft (Brown, 2015, S. 1). Weil die Konstrukte dieser Arbeit aus der Literatur übernommen wurden und sich bei früheren Arbeiten bewährt hatten, ist ein konfirmatorisches Vorgehen angemessen.

Die APA-Vorgaben verlangen bei einer solchen Analyse die Angabe der Modellgüte sowie die Angabe der Faktorladungen (Nicol & Pexman, 2010, S. 127). Die Modellgüte wird über diejenigen Werte ausgemacht, wie sie im vorigen Kapitel zu Strukturgleichungsmodellen generell vorgestellt wurden.

Oft wird zudem empfohlen, spezifische Prüfwerte für Reliabilität und Validität anzugeben. Bei einer Berechnung im Rahmen einer Strukturgleichungsmodellierung bieten sich hierzu Prüfwerte der «zweiten Generation» an (Bruhn et al., 2012; Weiber & Mülhhaus, 2014). Auf diese Prüfwerte wird im Folgenden eingegangen.

Reliabilität

Die gemeinsame Varianz reflektiver Indikatoren kann auf die dahinterliegende latente Variable zurückgeführt werden. Die restliche Varianz der Indikatoren wird als Messfehler bezeichnet (Werner et al., 2016, S. 949). Die Messfehler verschiedener Indikatoren sind zufällig und korrelieren nicht miteinander. Aus der gemeinsamen Varianz der Indikatoren ergibt sich eine messfehlerfreie Varianz des Konstrukts (Werner et al., 2016, S. 949). Ein Indikator ist reliabel, wenn ein hoher Teil seiner Gesamtvarianz durch die latente Variable erklärt wird (Urban & Mayerl, 2014, S. 138). Die *Faktorladung* eines Indikators durch die dazugehörige latente Variable sollte mindestens 0.4 betragen und signifikant sein (Weiber & Mülhhaus, 2014, S. 150). Die quadrierte Faktorladung ergibt die *Indikatorreliabilität* (Urban & Mayerl, 2014, S. 138). Diese gibt den Anteil der Varianz an, die das Konstrukt pro Indikator erklärt (Weiber & Mülhhaus, 2014, S. 150). Je höher die Indikatorreliabilität, desto höher der Anteil der erklärten Varianz und desto höher die Reliabilität des Indikators.

Neben diesen beiden indikatorbezogenen Kriterien werden auf Konstruktebene die *Faktorreliabilität* und die *durchschnittlich extrahierte Varianz* (DEV)² berechnet. Die Faktorreliabilität kann als «Indikatorreliabilität auf Konstruktebene» (S. 150) verstanden werden (Weiber & Mülhhaus, 2014). Sie sollte grösser als 0.6 sein. Die DEV gibt Auskunft darüber, «wie viel Prozent der Streuung des latenten Konstruktes über die Indikatoren durchschnittlich erklärt wird» (Weiber & Mülhhaus, 2014, S. 151). Als Schwellenwert wird hier 0.5 empfohlen (ebd.).

Während die Faktorenladungen direkt Mplus entnommen werden können, müssen die Werte zur Indikatorreliabilität, Faktorreliabilität und die durchschnittlich extrahierte Varianz auf Basis des Mplus-Outputs separat berechnet werden.

² Oft wird auch von durchschnittlich *erfasster* Varianz gesprochen (z. B. Herrmann et al., 2006). Der Begriff durchschnittlich extrahierte Varianz wird hier in Anlehnung an Weiber und Mülhhaus (2014) verwendet.

Die Formeln dazu finden sich beispielsweise bei Weiber und Mühlhaus (2014, S. 150–151). Tabelle 6.3 gibt einen summarischen Überblick über die Kriterien zur Beurteilung der Reliabilität.

Tabelle 6.3 Kriterien zur Beurteilung der Reliabilität latenter Variablen

Kriterium der Reliabilitätsprüfung	Schwellenwert
Faktorladung / Pfadkoeffizienten (λ)	$> 0.4^a$ bzw. $> 0.7^b$; und significant
Indikatorreliabilität	> 0.4
Faktorreliabilität	≥ 0.6
Durchschnittlich extrahierte Varianz	≥ 0.5

Anmerkungen. Die Schwellenwerte basieren auf den Empfehlungen von Weiber und Mühlhaus (2014, S. 150–151) und Hair, Black, Babin und Anderson (2014, S. 605).

^a Weiber und Mühlhaus (2014, S. 155). ^b Hair, Black, Babin und Anderson (2014, S. 605).

Validität

Weiber und Mühlhaus (2014, S. 156) unterscheiden drei Arten der Validität: Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität. *Inhaltsvalidität* ist gegeben, wenn die Indikatoren das zu messende Konstrukt «inhaltlich-semantic» (S. 157) abbilden. Zur Prüfung der Inhaltsvalidität wird beispielsweise auf die Meinungen von Experten zurückgegriffen (Weiber & Mühlhaus, 2014, S. 156). *Kriteriumsvalidität* liegt vor, wenn das Konstrukt mit einem verwandten «Aussenkriterium» (S. 157) übereinstimmt bzw. korreliert. Allerdings, so Weiber und Mühlhaus (2014, S. 168), wird die Kriteriumsvalidität in der Praxis selten erhoben, weil es aufwendig und schwierig ist, passende Aussenkriterien zu definieren. *Konstruktvalidität* schliesslich gliedert sich in *konvergente, diskriminante und nomologische Validität*. Nomologische Validität liegt vor, wenn die erhobenen Konstrukte in einem theoretisch begründeten Verhältnis zueinander stehen. Weil das Strukturgleichungsmodell ein solches theoretisches Verhältnis postuliert, kann bei hoher Modellgüte von nomologischer Validität ausgegangen werden (S. 161). Bei konvergenter Validität liefern zwei Erhebungsmethoden desselben Konstrukts (z. B. Befragung und Beobachtung) übereinstimmende Ergebnisse. Laut Weiber und Mühlhaus (2014, S. 162) wird konvergente Validität aufgrund des damit verbundenen grossen Aufwands in der Praxis selten berücksichtigt. Unterscheidet sich schliesslich das zu untersuchende Konstrukt stark von einem inhaltlich nicht verwandten Konstrukt, kann von diskriminanter Validität ausgegangen werden (S. 164). Hierzu kann das Fornell-Larcker-Kriterium herangezogen werden. Nach dem Fornell-Larcker-Kriterium (Fornell & Larcker,

1981) liegt diskriminante Validität vor, wenn die durchschnittlich extrahierte Varianz (DEV) einer latenten Variable f_1 grösser ist als die quadrierte Korrelation zwischen der latenten Variable f_1 und einer weiteren latenten Variable f_2 (Weiber & Mühlhaus, 2014, S. 165). Das Fornell-Larcker-Kriterium ist allerdings eine strenge Prüfung diskriminanter Validität. Eine weniger strikte Variante der Prüfung auf diskriminante Validität liegt darin zu zeigen, dass die Korrelation zweier latenter Variablen ungleich 1 ist (Weiber & Mühlhaus, 2014, S. 165).

Zusammenfassend raten die Autoren (Weiber & Mühlhaus, 2014, S. 168), die Inhaltsvalidität anhand von fundierter Operationalisierung und Rücksprache mit Experten sicherzustellen und anschliessend die Konstruktvalidität über die diskriminante Validität anhand des Fornell-Larcker-Kriteriums zu beurteilen (siehe Tabelle 6.4).

Tabelle 6.4 Kriterien zur Beurteilung der Validität latenter Variablen

Kriterium der Validitätsprüfung	Schwellenwert
Inhaltsvalidität	Operationalisierung und Expertenmeinung
Diskriminante Validität	$DEV(f) > \psi f_1 f_2$

Anmerkungen. Die Empfehlungen basieren auf Weiber und Mühlhaus (2014, S. 167).

6.4.5 Prüfung struktureller Zusammenhänge

Zur Beurteilung der Effektstärke einer unabhängigen Variable f_1 auf eine abhängige Variable f_2 dient der *Pfadkoeffizient* β zwischen den Variablen. Es kann zwischen unstandardisierten (β_u) und standardisierten (β) Werten unterschieden werden (Urban & Mayerl, 2014, S. 48). Unstandardisierte Koeffizienten müssen mit Blick auf die den Variablen zugrunde liegenden Skalen interpretiert werden (Weiber & Mühlhaus, 2014, S. 229). Sollen Pfadkoeffizienten unterschiedlicher Modelle (z. B. unterschiedliche Gruppen) verglichen werden, dann müssen die unstandardisierten Koeffizienten verwendet werden (Urban & Mayerl, 2014, S. 223). Im Normalfall werden die standardisierten Koeffizienten angegeben, weil deren Interpretation einfacher fällt. Steigt die unabhängige Variable f_1 um eine Standardeinheit (Standardabweichung), dann verändert sich die abhängige Variable um die durch den standardisierten Pfadkoeffizienten β angegebene (multiplizierte) Standardeinheit, wenn alle anderen Einflussvariablen auf die Variable f_2 konstant gehalten werden (Kleinke et al., 2017, S. 15; Urban & Mayerl, 2014, S. 48). Standardisierte Pfadkoeffizienten nehmen Werte zwischen +1 und -1 an,

wobei ab Werten von $+0.2$ bzw. -0.2 von einem bedeutungsvollen Zusammenhang gesprochen werden kann (Chin, 1998, S. xiii; Weiber & Mülhhaus, 2014, S. 229).

Wenn unterschiedliche Variablen unterschiedliche Varianzen – und damit Standardabweichungen – haben, erschwert das die Interpretation der Effekte. So könnte eine abhängige Variable $f3$ bei höherer Varianz deutlich stärker steigen als eine abhängige Variable $f4$, obwohl beide mit demselben Pfadkoeffizienten durch die unabhängige Variable $f1$ beeinflusst werden (Urban & Mayerl, 2014, S. 49). Weil die zentralen Variablen im Rahmen dieser Arbeit alle auf einer fünfstufigen Skala gemessen wurden (siehe Abschnitt 6.8.2), scheint diese Gefahr für diese Arbeit weniger relevant.

Unabhängige Variablen können auch indirekte Effekte auf eine abhängige Variable haben. In diesem Fall wird ein *Mediator* durch die unabhängige Variable beeinflusst, der sich wiederum auf die abhängige Variable auswirkt. Der Gesamteffekt einer unabhängigen Variable auf die abhängige Variable ergibt sich aus der Summe des direkten und aller indirekten Effekte (Werner et al., 2016, S. 952).

Zu den Pfadkoeffizienten wird ein *p-Wert* ausgewiesen. Ein mit einem erhaltenen Resultat verbundener *p-Wert* soll Auskunft darüber geben, ob das erhaltene Resultat *statistisch signifikant* ist oder nicht. Wichtig ist die korrekte Interpretation der *p-Werte* als diejenige Wahrscheinlichkeit, mit der ein erhaltenes (oder noch extremeres) Resultat unter den Annahmen der Nullhypothese auftritt (Kline, 2016, S. 55). Zur Signifikanzprüfung werden typischerweise drei Niveaus herangezogen. Diese werden mit α bezeichnet.

- $\alpha = 0.05$
- $\alpha = 0.01$
- $\alpha = 0.001$

Erhält man zu einem getesteten Effekt einen *p-Wert*, der unter dem stärksten Niveau signifikant ist (z. B. $p < 0.001$), dann ist die Wahrscheinlichkeit kleiner als 0.1 %, dass der gefundene Effekt (oder ein noch stärkerer Effekt) unter den Annahmen der Nullhypothese auftritt. Die *Nullhypothese* wird dabei durch den Forschenden festgelegt (Kühnel & Krebs, 2007, S. 277) und besagt für Pfadkoeffizienten, dass deren Effekt in der Population 0 ist (Muthén & Muthén, 2017, S. 471). Als *Population* kann bei Aussagen über Signifikanz und Konfidenzintervalle immer nur diejenige Gruppe verstanden werden, für welche die gezogene Stichprobe repräsentativ steht (siehe Brosius et al., 2012, S. 59; Rost, 2013, S. 102).

P-Werte werden unter Zuhilfenahme des *Standardfehlers* der Schätzung (SE = Standard Error) berechnet. Der Standardfehler gibt Auskunft darüber, «mit welcher Streuung bei den jeweiligen Parameterschätzungen zu rechnen ist» (Weiber & Mühlhaus, 2014, S. 229). Dividiert man den unstandardisierten Pfadkoeffizienten durch den Standardfehler, erhält man den t-Wert (C. R. = Critical Ratio). Die Verteilung dieser Werte (t-Verteilung) nähert sich ab einer Stichprobe von ca. 120 Fällen der z-Verteilung (Urban & Mayerl, 2014, S. 59). Anhand der z-Verteilung lässt sich die Signifikanz der Pfadkoeffizienten bewerten. Ein t-Wert von ± 1.96 zeugt von Signifikanz auf dem 5 %-Level.

Standardfehler sinken mit der Stichprobengrösse und steigen mit der Varianz des zu schätzenden Koeffizienten oder Parameters (Kuckartz, Rädiker, Ebert & Schehl, 2013, S. 141; Rost, 2013, S. 237). Signifikanz muss unter Berücksichtigung der Stichprobengrösse beachtet werden. Bei grossen Stichproben tendieren kleine (nicht vorhandene) Effekte zu Signifikanz (Typ-1-Fehler), während bei kleineren Stichproben tatsächlich vorhandene (starke) Effekte keine Signifikanz erreichen (Typ-2-Fehler; Geuens & Pelsmacker, 2017, S. 87).

Um verschiedene Parameter innerhalb eines Strukturgleichungsmodells oder verschiedene Parameter über verschiedene Gruppen hinweg zu vergleichen (z. B. $\beta_a = \beta_b$) und eine allfällige Differenz auf Signifikanz zu testen, wird ein χ^2 -Differenztest angewandt. Dazu wird ein unrestringiertes Modell (β_a und β_b werden frei geschätzt) einem restringierten Modell ($\beta_a = \beta_b$) gegenübergestellt und geschaut, ob sich die χ^2 -Werte der beiden Modelle signifikant unterscheiden. Verschlechtert sich das Modell durch die Restriktionen signifikant (χ^2 -Wert wird grösser), dann muss die Nullhypothese, wonach sich die Parameter nicht unterscheiden, abgelehnt werden (Weiber & Mühlhaus, 2014, S. 164–165). Bei Mplus kann dazu ein χ^2 -Differenz-Test definiert und durchgeführt werden (Muthén & Muthén, 2017, S. 772). Bei Multigruppenanalysen werden bei Mplus die Faktorladungen und Achsenabschnitte der Indikatoren über die Gruppen hinweg konstant gehalten, um Messinvarianz (siehe Abschnitt 6.4.6) sicherzustellen (Muthén & Muthén, 2017, S. 81).

Ein *Konfidenzintervall* umfasst diejenigen Resultate oder Effekte, die unter Berücksichtigung des Stichprobenfehlers als äquivalent zum erhaltenen Resultat angesehen werden können (Kline, 2016, S. 57). Im Gegensatz zu Signifikanztests wird bei der Berechnung von Konfidenzintervallen keine Nullhypothese vorausgesetzt. Das Konfidenzintervall eines Parameters hängt vom Standardfehler ab und ergibt sich, indem der Standardfehler mit 1.96 multipliziert wird und das Ergebnis zum Schätzwert addiert (obere Grenze) bzw. vom Schätzwert subtrahiert (untere Grenze) wird (Urban & Mayerl, 2014, S. 61). Entsprechend ist davon auszugehen, dass sich die Endpunkte des Intervalls bei einer anderen

(zweiten) Stichprobe ändern (Kline, 2016). Typischerweise werden 90 %- oder 95 %-Konfidenzintervalle verwendet. Das 95 %-Konfidenzintervall besagt, dass das Intervall den gesuchten Populationswert mit einer Wahrscheinlichkeit von 95 % enthält (Kühnel & Krebs, 2007, S. 244).

6.4.6 Modellierung experimenteller Manipulationen

Um Populationen oder Gruppen, die unterschiedlichen experimentellen Stimuli zugeordnet wurden, zu vergleichen, gibt es innerhalb eines Strukturgleichungsmodells zwei Möglichkeiten: MIMIC (multiple-indicator, multiple-cause) und SMM (Structured Means Modeling; Breitsohl, 2019; Hancock, 2004).

MIMIC (multiple-indicator, multiple-cause)

Bei MIMIC-Auswertungen werden *Dummy-Variablen* als exogene Variablen verwendet, um die Gruppenzugehörigkeit zu berücksichtigen (Hancock, 2004). Dummy-Variablen sind dichotome, kategoriale Variablen und verfügen über zwei Werte: 0 (Kategorie 1) und 1 (Kategorie 2). Eine Dummy-Variable (z. B. 0 = informativer Brand Content, 1 = unterhaltender Brand Content) beeinflusst bei erfolgreicher Manipulation die Ausprägung einer unabhängigen Experimentalvariable (MacKenzie, 2001). Ist der Pfadkoeffizient der Dummy-Variable auf die unabhängige Experimentalvariable signifikant höher als 0, kann von einem Effekt der Manipulation gesprochen werden (Hancock, 2004, S. 321). Die unabhängigen Experimentalvariablen wirken sich schliesslich auf die abhängigen Variablen aus (MacKenzie, 2001, S. 161–162). Die Daten unterschiedlicher (Experimental-) Gruppen werden dadurch gemeinsam in einem Modell analysiert.

Erreicht der Pfadkoeffizient von einer Dummy-Variable (z. B. Gruppenzugehörigkeit, Manipulationsscheck) hin zu einer latenten Variable (z. B. Globalimage) statistische Signifikanz, dann kann davon ausgegangen werden, dass sich die Mittelwerte dieser latenten Variablen zwischen den getesteten Gruppen oder Manipulationen unterscheiden. Die Nullhypothese hingegen, wonach sich die Mittelwerte nicht unterscheiden, kann verworfen werden (Hancock, 2004, S. 322). Der Effekt einer Manipulation zugrundeliegenden Dummy-Variablen auf eine abhängige Variable kann als experimenteller *Haupteffekt* bezeichnet werden. Haupteffekte kennzeichnen die Wirkung der unterschiedlichen experimentellen Faktoren (Koch et al., 2019, S. 79).

Mit MIMIC wird davon ausgegangen, dass in allen Gruppen dasselbe Messmodell gültig ist (Hancock, 2004, S. 332). Das heisst, unterschiedliche Ausprägungen einer beobachteten Variable können auf unterschiedliche Ausprägungen

der latenten Variable zurückgeführt werden. Das ist der Fall, wenn keine weitere Variable, die in den verschiedenen Gruppen unterschiedlich ausgeprägt ist, auf die abhängige, latente Variable einwirkt. Bei einem experimentellen Design und einer randomisierten Zuordnung der Teilnehmenden in die Experimentalgruppen kann davon ausgegangen werden, dass das Messmodell in den Gruppen identisch ist. MIMIC berücksichtigt keine y-Achsenabschnitte (intercepts). Die Variablen haben einen Mittelwert von 0, die Werte können als Abweichungen vom Mittelwert im Gesamtsample verstanden werden (Hancock, 2004, S. 321).

Multigruppenvergleiche

Bei SMM (Structured Means Modeling) werden im Gegensatz zu MIMIC die Daten unterschiedlicher (*Experimental-*)Gruppen *getrennt analysiert* (Hancock, 2004, S. 326). Dummy-Variablen sind dadurch nicht länger nötig. Bei SMM handelt es sich entsprechend um eine Multigruppenanalyse, mit der kategoriale *Moderations- oder Interaktionseffekte* überprüft werden können (Urban & Mayerl, 2014, S. 40). Die Begriffe Moderation und Interaktion werden oft bedeutungsgleich verwendet (Koch et al., 2019, S. 79) und drücken aus, dass der Effekt einer unabhängigen Variable auf eine abhängige Variable (z. B. ein Haupteffekt) von der Ausprägung einer Drittvariable, der Moderations- oder Interaktionsvariable, abhängt (Baron & Kenny, 1986, S. 1174). Während der Begriff Moderation der Korrelationsanalyse entstammt, fand der Begriff der Interaktion vor allem in Zusammenhang mit Varianzanalysen Verwendung (Baron & Kenny, 1986). Baron und Kenny (1986) verwenden den Begriff Moderator als Überbegriff und sprechen von Interaktion (als eine Art der Moderation), wenn sowohl die unabhängige Variable als auch die Moderatorvariable kategorial sind. Bei der Beschreibung der Hypothesen wurde dieser Unterschied ebenfalls angewandt (siehe Abschnitt 5.1).

Bei SMM wird ebenfalls davon ausgegangen, dass Unterschiede zwischen den Gruppen bei den beobachteten Variablen auf unterschiedliche Ausprägungen der latenten Variablen zurückzuführen sind. Damit diese Folgerung allerdings zulässig ist, müssen die Messmodelle zwischen den Gruppen invariant sein (Hancock, 2004). Das ist der Fall, wenn eine latente Variable über verschiedene zu vergleichende Gruppen hinweg gleich gemessen wird und sogenannte Messinvarianz gegeben ist (Cieciuch & Davidov, 2016). Messinvarianz kann auf verschiedenen Stufen vorliegen: *konfigurale Invarianz*, *metrische Invarianz* und *Skalar-Invarianz* (Cieciuch & Davidov, 2016). Eine höhere Stufe der Invarianz baut jeweils auf den Annahmen der davorgehenden Stufe auf. Skalar-Invarianz erfordert entsprechend metrische Invarianz.

Die Prüfung auf Messinvarianz erfolgt mittels einer konfirmatorischen Faktorenanalyse in mehreren Schritten (Cieciuch & Davidov, 2016). Es können einzelne latente Variablen eigenständig oder ausgewählte, zusammenhängende latente Variablen simultan auf Messinvarianz überprüft werden (Bruhn et al., 2012). Zu Beginn erfolgt der Test auf *konfigurale Invarianz*, der darüber Auskunft gibt, ob latente Variablen in verschiedenen Gruppen über dieselben Indikatoren gemessen werden (Cieciuch & Davidov, 2016). Bei *metrischer Invarianz* werden die Faktorladungen der Indikatoren über alle Gruppen hinweg konstant gehalten (Weiber & Mülhhaus, 2014, S. 300). Bei gegebener metrischer Invarianz können unstandardisierte Pfadkoeffizienten zwischen Gruppen verglichen werden (Cieciuch & Davidov, 2016, S. 2). Solche Vergleiche sind auch möglich, wenn nur partielle metrische Invarianz vorliegt (Steenkamp & Baumgartner, 1998). In diesem Fall müssen nicht alle, aber mindestens zwei Indikatoren pro latenter Variable über die Gruppen hinweg konstant sein (Steenkamp & Baumgartner, 1998, S. 81). Bei *Skalar-Invarianz* werden neben den Faktorladungen auch die Achsenabschnitte (intercepts) der Indikatoren über die Gruppen hinweg gleichgesetzt (Weiber & Mülhhaus, 2014, S. 300). Skalar-Invarianz erlaubt den Vergleich von Mittelwerten (Cieciuch & Davidov, 2016, S. 2). Auch hier genügt zu Vergleichszwecken partielle Skalar-Invarianz (Steenkamp & Baumgartner, 1998).

Der Test auf vorliegende Invarianz sieht vor, dass die Güte eines Modells höherer Invarianz nicht signifikant schlechter sein darf als die Güte des Modells auf der vorhergehenden Stufe (Cieciuch & Davidov, 2016; Steinmetz, Schmidt, Tina-Booh, Wieczorek & Schwartz, 2009). Dazu wird ein χ^2 -Differenztest durchgeführt. Bei einer MLR-Schätzung (Maximum Likelihood With Robust Standard Errors), wie sie im Rahmen dieser Arbeit angewandt wird, müssen die χ^2 -Werte für den Vergleich um den durch Satorra und Bentler (2001) entwickelten «Scaling Correction Factor (SCL)» bereinigt werden (Kleinke et al., 2017, S. 46–48). Allerdings reagiert der χ^2 -Text, wie schon in Abschnitt 6.4.3 ausgeführt, sensibel auf die Stichprobengröße. Bei größeren Stichproben sind Unterschiede zwischen den Gruppen tendenziell signifikant (Cheung & Rensvold, 2002, S. 234). Deshalb wird die Modellgüte unterschiedlicher Invarianz-Levels anhand von CFI-, RMSEA- und SRMR-Unterschieden beurteilt (Cieciuch & Davidov, 2016). Nach F. F. Chen (2007, S. 501) können folgende Schwellenwerte zur Beurteilung von Invarianz verwendet werden bei Sample-Größen über 300: $\Delta\text{CFI} < -0.01$, $\Delta\text{RMSEA} < +0.015$ und $\Delta\text{SRMR} < +0.03$ (von konfiguraler zu metrischer Invarianz) bzw. $< +0.01$ (von metrischer Invarianz zu Skalar-Invarianz). Während der χ^2 -Wert bei einer MLR-Schätzung für den Modellvergleich angepasst

werden muss, können Gütekriterien wie der CFI oder RMSEA bei einer MLR-Schätzung ohne Anpassung für den Modellvergleich genutzt werden (Chiorri, Day & Malmberg, 2014, S. 52). Die Kriterien zur Prüfung von Messinvarianz sind in Tabelle 6.5 zusammengefasst.

Tabelle 6.5 Kriterien zur Beurteilung der Messinvarianz zwischen Gruppen

Kriterium der Invarianzprüfung	Schwellenwert
χ^2 -Differenztest (Δp)	nicht significant
ΔCFI	< -0.01
$\Delta RMSEA$	$< +0.015$
$\Delta SRMR$	$< +0.03$ für metrische Invarianz $< +0.01$ für Skalar-Invarianz

Anmerkungen. Die Schwellenwerte basieren auf Kleinke, Schlüter und Christ (2017), Cieciuch und Davidov (2016) und F. F. Chen (2007).

6.4.7 Kontrollvariablen und demografische Moderationseffekte

Kontrollvariablen dienen der Prüfung, ob eine Korrelation zwischen zwei Variablen durch eine allfällige Drittvariable (Kontrollvariable) erklärt wird (Rost, 2013, S. 163). Kontrollvariablen werden als unabhängige Dummy-Variablen in das Strukturgleichungsmodell integriert. Sämtliche abhängigen latenten Variablen des Strukturgleichungsmodells werden dazu zusätzlich auf die Kontrollvariablen regrediert. Ist der Pfadkoeffizient der Kontrollvariable hin zu einer abhängigen latenten Variable signifikant positiv (negativ), dann ist die abhängige Variable bei Personen mit Ausprägung 1 bei der Dummy-Variable höher (tiefer). Bei diesem Vorgehen müssen immer auch die postulierten Zusammenhänge zwischen den latenten Variablen auf sich ändernde Signifikanz überprüft werden (Rost, 2013, S. 163). Verliert beispielsweise der Pfad von Variable f1 zu Variable f2 statistische Signifikanz durch die Hinzunahme einer Kontrollvariable, dann muss davon ausgegangen werden, dass die Kontrollvariable als Drittvariable f2 positiv beeinflusst. Es wird empfohlen, Kontrollvariablen als zusätzliche Analysen im Nachgang der Hypothesenprüfung zu besprechen (Deutsche Gesellschaft für Psychologie, 2019, S. 38).

Auch allfällige Moderationsanalysen auf Basis demografischer Variablen ohne spezifische Hypothesen (z. B. Unterschiede zwischen Geschlechtern) sollen als zusätzliche Analysen am Ende des Resultatekapitels diskutiert werden (Deutsche Gesellschaft für Psychologie, 2019, S. 38). Dazu wird der Datensatz anhand der Kontrollvariablen in mehrere Gruppen eingeteilt und eine Multigruppenanalyse gerechnet. Damit können Interaktionseffekte der Gruppenzugehörigkeit geprüft werden (wie in Abschnitt 6.4.6 beschrieben). Unterschiedlich ausgeprägte Pfadkoeffizienten zwischen zwei latenten Variablen f_1 und f_2 bei zwei Gruppen zeigen, dass der Effekt durch die demografische Variable moderiert wird.

6.4.8 Multikollinearitätsprüfung

Multikollinearität liegt vor, wenn in einem multiplen Regressionsmodell zwei oder mehrere unabhängige Variablen stark korrelieren (Kühnel & Krebs, 2007, S. 545). Kühnel und Krebs sehen eine *Korrelation* ab einem *Schwellenwert* von $r \geq 0.9$ als problematisch an. Im Rahmen dieser Arbeit wird ein Strukturgleichungsmodell mit direkten und indirekten Effekten gerechnet. Verschiedene Variablen sind als Mediatoren sowohl abhängige als auch unabhängige Variablen: Informationsgratifikation des Brand Contents, Unterhaltungsgratifikation des Brand Contents, Einstellung gegenüber dem Brand Content, funktionales Markenimage, emotionales Markenimage, Globalimage der Marke, Informationsgratifikation des Kanals, Unterhaltungsgratifikation des Kanals und Einstellung gegenüber dem Kanal. Es folgt in Abschnitt 6.8.1 im Rahmen der konfirmatorischen Faktorenanalyse eine Prüfung auf Multikollinearität. Multikollinearität ist allerdings insbesondere bei formativen Messmodellen ein Problem (Weiber & Mülhhaus, 2014, S. 262).

6.4.9 Normalverteilung der Daten

Ursprünglich wurde bei kovarianzbasierten Verfahren der Strukturgleichungsanalyse eine Normalverteilung der Daten vorausgesetzt (Herrmann et al., 2006). Um von einer Normalverteilung auszugehen, müssen einerseits die einzelnen Variablen normalverteilt sein, andererseits muss eine multivariate Normalverteilung vorliegen (Weiber & Mülhhaus, 2014, S. 180).

Prüfung einzelner Variablen auf Normalverteilung: Die Daten einer Variable sind normalverteilt, wenn Schiefe (Skewness) und Wölbung (Kurtosis) der Verteilung gleich 0 sind (Weiber & Mühlhaus, 2014, S. 180). Ab Werten von $|>2|$ bei der Schiefe und $|>7|$ bei der Wölbung ist von einer Verletzung der Normalverteilung auszugehen. Der Kolmogorov-Smirnoff-Test und Shapiro-Wilk-Test können zur Beurteilung der Normalverteilung zugezogen werden. Allerdings würden Ratingskalen den (für SEM zu) strengen Anforderungen der Tests oftmals nicht gerecht, weshalb zu deren Beurteilung auf die Schiefe- und Wölbe-Koeffizienten zurückgegriffen werden soll (Weiber & Mühlhaus, 2014, S. 180–181).

Prüfung auf multivariate Normalverteilung: Zur Prüfung der multivariaten Normalverteilung dient unter anderem «Mardia's Koeffizient der multivarianten Wölbung» (Weiber & Mühlhaus, 2014, S. 181). Allerdings ist der Koeffizient nur bedingt hilfreich, weil er bei grösseren Samples schon kleine Abweichungen als signifikante Verletzung der multivariaten Normalverteilung taxiert (Kline, 2016, S. 74). Zudem ist die Berechnung des Mardias Koeffizienten nur möglich, wenn keine fehlenden Werte vorliegen.

Fortgeschrittene Schätzmethoden (z. B. Maximum Likelihood with Robust Standard Errors, MLR) moderner Berechnungsprogramme wie Mplus verlangen keine Normalverteilung mehr (Muthén & Muthén, 2017, S. 9). In Abschnitt 6.8.1 wird dennoch auf die Normalverteilung der Daten eingegangen, um die Entscheidung zugunsten eines allfälligen, auf Normalverteilung korrigierten Schätzers zu fällen.

6.5 Vorstudien und Entwicklung der experimentellen Faktoren

Von April bis Dezember 2018 wurden insgesamt vier Vorstudien durchgeführt, um einerseits die für das Hauptexperiment erforderlichen Manipulationen (bzw. Faktoren) auf Ebene der Kanäle, der Marken und des Brand Contents zu ermitteln und zu entwickeln und andererseits die Messmodelle für die Erfassung der Konstrukte und Variablen zu testen (siehe Tabelle 6.6).

Tabelle 6.6 Übersicht über die durchgeführten Vorstudien

	Zeitraum	Sample	Inhalt
Vorstudie Nr. 1	April bis Mai 2018	181 Studierende	Kanäle/Marken
Vorstudie Nr. 2	Juni bis Juli 2018	82 Personen eines Online-Convenience-Samples	Marken
Vorstudie Nr. 3	Oktober 2018	101 Studierende	Marken
Vorstudie Nr. 4	Dezember 2018	142 Studierende	Brand Content

Anmerkungen. Bei den Studierenden handelt es sich um Studentinnen und Studenten in der Ausbildung (Bachelor / Master) und in der Weiterbildung der Hochschule Luzern – Wirtschaft.

6.5.1 Digitale Kanäle

In Abschnitt 3.8.2 wurde aufgezeigt, dass *unterschiedliche soziale Medien* unterschiedliche Gratifikationen erbringen und für unterschiedliche Zwecke genutzt werden (Voorveld et al., 2018). Die Forschung hat ebenfalls gezeigt, dass das Medium als Kontext einer Markenbotschaft einen Einfluss auf die Wahrnehmung der Botschaft hat. Im Rahmen dieser Arbeit soll untersucht werden, inwiefern die Wahrnehmung von Brand Content durch die Gratifikationserwartungen an ein soziales Medium beeinflusst wird. Aus diesem Grund sollten für die Untersuchung zwei digitale Kanäle ausgewählt werden, wobei ein Kanal primär mit Informations- und der andere Kanal primär mit Unterhaltungserwartungen verbunden wird. Als weitere Kriterien mussten die sozialen Medien in der Schweiz bekannt sein und sowohl auf User- als auch auf Unternehmensseite eine gewisse Relevanz haben.

Laut einer Studie des Content Marketing Forums (Koob, 2018) fliessen in Deutschland, Österreich und der Schweiz 40 % der digitalen Content-Marketing-Investitionen in Webseiten, Microsites und Apps. 26 % der Investitionen werden für Social-Media-Aktivitäten ausgegeben. Die am häufigsten genutzte Plattform ist Facebook, auf den Rängen 2 und 3 finden sich YouTube und Twitter. Es folgen WhatsApp, Facebook Messenger, Google Plus und Instagram (Koob, 2018).

Einen Blick auf die Nutzung sozialer Medien in der Schweizer Bevölkerung erlaubt der Media Use Index, der alle zwei Jahre durch die Agentur Y&R herausgegeben wird (Y&R Group Switzerland, 2018). Dem aktuellen Report nach sind WhatsApp, Instagram, Snapchat, Facebook, Facebook Messenger und Twitter die wichtigsten sozialen Medien. Stellt man die Unternehmenssicht der Usersicht entgegen, ergibt sich folgendes Bild (siehe Tabelle 6.7):

Tabelle 6.7 Nutzung von sozialen Medien bei Schweizer Unternehmen und Usern

Rang-platz	Unternehmen	User
1	Facebook	WhatsApp
2	YouTube	Instagram
3	Twitter	Snapchat
4	WhatsApp	Facebook
5	Facebook Messenger	Facebook Messenger
6	(Google Plus)	Twitter
7	Instagram	

Anmerkungen. Quellen sind Koob 2018 für die Unternehmenssicht und Y&R Group Switzerland 2018 für die Usersicht.

Weil Google Plus 2019 eingestellt wurde und WhatsApp keine Werbung bzw. Brand Pages anbietet, fielen diese beide Plattformen für die Analyse weg. Weil nicht zwei gleiche Kanäle verglichen werden sollen, wurde auch der Facebook Messenger zugunsten des Hauptportals Facebook zurückgestellt. YouTube fiel ebenso aus dem engeren Analysekreis, weil Bewegtbilder im Rahmen dieser Studie nicht untersucht wurden. Snapchat ist schliesslich für Schweizer Firmen (noch) wenig relevant.

Für die Vorstudie wurden deshalb die drei digitalen Kanäle Facebook, Twitter und Instagram auf die *Gratifikationserwartungen* (LaRose & Eastin, 2004) bezüglich Information und Unterhaltung und bezüglich der Einstellung gegenüber dem Kanal untersucht. Informationserwartungen, Unterhaltungserwartungen und Einstellung gegenüber dem Kanal wurden jeweils mit drei Indikatoren gemessen (siehe Abschnitt 6.8.4). Die Vorstudie fand von April bis Mai 2018 mit einem Studierenden-Sample statt. Alle 181 Studierenden bewerteten jeweils alle drei Kanäle.

Eine konfirmatorische Faktorenanalyse (Mplus 8.0, MLR-Schätzung) mit allen drei Variablen ergab eine akzeptable Modellgüte ($df = 24$, $\chi^2 = 89.198$, $\chi^2/df = 3.72$, $p = 0.00$, CFI = 0.970, TLI = 0.955, RMSEA = 0.073 mit 90 % K.I. zwischen 0.057 und 0.090, SRMR = 0.044).³ Die drei latenten Variablen können entsprechend mit den verwendeten Indikatoren operationalisiert werden. Eine Prüfung auf (partielle) Skalar-Invarianz, um die Mittelwerte der latenten Faktoren zu vergleichen, war nicht erfolgreich (siehe elektronisches Zusatzmaterial

³ Die Modellgüte für die jeweiligen Gruppen aller Prüfungen auf Messinvarianz wird im elektronischen Zusatzmaterial aufgeführt (siehe hier beispielsweise Tabelle A-2).

Tabelle A-1). Zur Etablierung der partiellen Skalar-Invarianz wurden die Faktorladungen und Achsenabschnitte eines Indikators pro latenter Variable über die Gruppen hinweg frei geschätzt. Allerdings wurde auch partielle Skalar-Invarianz nicht erreicht. Bei fehlender (partieller) Skalar-Invarianz müssen die Mittelwerte mit Vorsicht interpretiert und verglichen werden. Es werden keine Signifikanztests durchgeführt (Steinmetz et al., 2009, S. 608–610). Tabelle 6.8 fasst die Mittelwerte der drei Kanäle bezüglich Informations- und Unterhaltungserwartungen sowie zur Einstellung gegenüber dem Kanal zusammen.

Tabelle 6.8 Mittelwerte der kanalbezogenen latenten Variablen für Kanäle der Vorstudie

Kanal	Informations- erwartungen	Unterhaltungs- erwartungen	Einstellung gegenüber Kanal
Facebook (<i>n</i> = 181)	0	0	0
Twitter (<i>n</i> = 160)	1.055	−0.671	0.065
Instagram (<i>n</i> = 168)	−0.062	0.633	0.933

Anmerkung: Information (K_IE), Unterhaltung (K_UE), Einstellung gegenüber Kanal (K_ES; siehe Abschnitt 6.8.4); Berechnung erfolgte auf Basis partieller Skalar-Invarianz (siehe elektronisches Zusatzmaterial Tabelle A-1).

Die Mittelwerte bei Facebook wurden auf null gesetzt und dienen als Referenz. Die Angaben zu den Mittelwerten bei Twitter und Instagram beziehen sich auf die Abweichung gegenüber diesem Referenzwert. *Instagram* schürt höhere *Unterhaltungserwartungen* als Facebook und Twitter, Facebook schürt höhere *Unterhaltungserwartungen* als Twitter. Ein umgekehrtes Bild ergibt sich bei den *Informationserwartungen*. Dort erreicht *Twitter* einen höheren Wert als Facebook und Instagram. Die beiden Letzteren verfügen über ähnlich hohe *Informationserwartungen*. Schliesslich erreicht Instagram eine höhere *Einstellung gegenüber dem Kanal* als Twitter und Facebook. Auf Basis des Pretests wurde Twitter als informativer Kanal und Instagram als unterhaltender Kanal für das Hauptexperiment gewählt.

6.5.2 Marken

Die Schweiz ist ein Land mit grosser Markentradition. Trotz der hohen Bedeutung von Marken gibt es in der Schweiz vergleichsweise wenig wissenschaftliche Markenforschung. Die Arbeit soll dazu beitragen, diesem Defizit entgegenzuwirken. Aus diesen Überlegungen heraus wurde der Fokus der empirischen Untersuchung

auf *reale Schweizer Marken* gelegt. Auf fiktive Marken wurde aus Gründen der externen Validität (siehe Abschnitt 6.2) verzichtet. Die Resultate sollten Markenverantwortlichen in der Schweiz eine realitätsnahe Hilfestellung zum Einsatz von Brand Content in der digitalen Markenkommunikation geben. Weil Kommunikationseffekte bei Dienstleistungen anderen Automatismen unterliegen (M. R. Stafford & Day, 1995), wurden diese nicht berücksichtigt. Ziel der Vorstudie war es, vier Marken zu identifizieren, die sich bezüglich der Markenpositionierung (funktional vs. emotional) und bezüglich des Produktinvolvements (High vs. Low Involvement) unterscheiden.

Die Auswahl von Marken für die Vorstudien (siehe Tabelle 6.9) wurde auf der Basis von zwei Quellen getroffen. Einerseits wurden *praxisorientierte Markenstudien* konsultiert, die bekannte Schweizer Marken identifizieren (GfK Switzerland, 2018; Interbrand, 2016; Y&R Group Switzerland, 2015). Andererseits wurde auf *einschlägige Literatur* zurückgegriffen, die unterschiedliche Produkte anhand des primären funktionalen oder emotionalen Kundennutzens sowie des Produktinvolvements kategorisieren (Weinberger et al., 1995).

Tabelle 6.9 Untersuchte Marken im Rahmen der Vorstudien

Marke	Branche / Produkt	Involvement (angenommen)	Positionierung (angenommen)
<i>Vorstudie 1: April und Mai 2018</i>			
Elmex	Zahnpasta	Tief	Funktional
Rivella	Süssgetränke	Tief	Emotional
Freitag	Freizeittaschen	Hoch	Emotional
Odlo	Sportkleidung	Hoch	Funktional
<i>Vorstudie 2: Juli und August 2018</i>			
Lindt	Schokolade	Tief	Emotional
Rolex	Uhren	Hoch	Emotional
Zweifel	Pommes-Chips	Tief	Emotional
<i>Vorstudie 3: Oktober 2018</i>			
V-Zug	Haushaltsgeräte	Hoch	Funktional
Flyer	Elektrofahrräder	Hoch	Funktional
Swatch	Uhren	Hoch	Emotional

Anmerkungen. Annahmen zu Involvement und Positionierung durch den Autor u. a. auf Basis von Voss, Spangenberg und Grohmann (2003), Weinberger, Spotts, Campbell und Parsons (1995) und Vaughn (1986).

Auf Basis dieser Grundlage wurden insgesamt zehn Schweizer Marken in den Vorstudien 1, 2 und 3 auf die Markenpositionierung und das Produktinvolvement getestet.

Die *Markenpositionierung* wurde anhand eines einzelnen Items auf einer fünfstufigen Likert-Skala erhoben mit *Bitte geben Sie an, ob es aus Ihrer persönlichen Sicht eine komplett rationale/kopforientierte (1) oder eine komplett emotionale/gefühlorientierte Entscheidung (5) ist*, die Produkte der jeweiligen Marke zu kaufen (Kempf, 1999, S. 43). Um zu prüfen, ob die Mittelwerte über die drei Vorstudien hinweg vergleichbar sind, wurden die Marken Elmex, Rivella, Freitag, Odlo und Rolex in allen drei Vorstudien einbezogen. Es zeigte sich, dass die Mittelwerte über die drei Erhebungen zwar variieren, die Einteilung der Marken auf dem Kontinuum funktional – emotional aber grösstenteils konstant blieb. Für die zehn getesteten Marken ergaben sich gemäss Tabelle 6.10 folgende Mittelwerte für die Markenpositionierung:

Tabelle 6.10 Mittelwerte der Markenpositionierung der Marken der Vorstudien

	Vorstudie 1			Vorstudie 2			Vorstudie 3		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>n</i>
Elmex	2.01	1.13	176	2.23	1.15	79	2.10	1.10	100
Rivella	3.55	1.13	180	3.38	1.21	80	3.90	1.26	98
Freitag	3.40	1.18	175	3.69	1.17	74	3.74	1.12	97
Odlo	2.36	1.08	146	2.37	1.15	73	2.77	1.15	86
Lindt	3.87	1.10	178	3.72	1.12	79			
Rolex	3.71	1.36	177	3.61	1.36	77	4.13	1.19	100
Zweifel	3.62	1.17	178	3.44	1.22	80			
V-Zug							1.63	0.82	97
Flyer							1.87	0.99	84
Swatch							3.41	1.20	101

Anmerkungen. Die Mittelwerte wurden mit SPSS 25 berechnet. Leere Zeilen bedeuten, dass der entsprechende Wert in der jeweiligen Vorstudie nicht erhoben wurde. Die Skala reichte von 1 (funktionale Positionierung) bis 5 (emotionale Positionierung).

Das *Produktinvolvement* der Marken wurde als latente Variable in je einer Vorstudie erhoben. Das Konstrukt wurde mit drei Indikatoren gemessen (siehe Abschnitt 6.8.7) und ist damit gerade identifiziert. Es können keine Güterwerte ausgegeben werden (Kline, 2016, S. 201). Um die Mittelwerte einer latenten Variable über verschiedene Gruppen und Stichproben zu vergleichen, muss

Skalar-Invarianz vorliegen (Cieciuch & Davidov, 2016). Bei fehlender Skalar-Invarianz kann auf partielle Skalar-Invarianz zurückgegriffen werden. Zur Berechnung der partiellen Invarianz wurden die Faktorladung und der Achsenabschnitt einer Variable über die Gruppen hinweg freigegeben. Allerdings sind auch für partielle Skalar-Invarianz die Voraussetzungen nicht gegeben (siehe elektronisches Zusatzmaterial Tabelle A-3). Deshalb müssen die Mittelwerte mit Vorsicht interpretiert werden, können höchstens als Indikator für das ungefähre Produktinvolvement dienen und sollen nicht auf signifikante Unterschiede geprüft werden (Steinmetz et al., 2009, S. 608–610). Die Berechnung erfolgte mit Mplus über eine Multigruppenanalyse (siehe Tabelle 6.11). Der Mittelwert von Elmex wird per Definition auf 0 gesetzt. Alle anderen Mittelwerte sind als Differenz zu diesem Referenzwert zu verstehen.

Tabelle 6.11 Mittelwerte des Produktinvolvements der Marken der Vorstudien

Marke	Vorstudie 1	Vorstudie 2	Vorstudie 3
Elmex ($n = 92$)	0		
Rivella ($n = 90$)	−0.142		
Odlo ($n = 87$)	0.625		
Freitag ($n = 88$)	0.481		
Lindt ($n = 82$)		−0.132	
Rolex ($n = 81$)		0.605	
Zweifel ($n = 82$)		−0.253	
Swatch ($n = 101$)			0.890
V-Zug ($n = 98$)			1.058
Flyer ($n = 99$)			0.899

Anmerkungen. Mittelwerte für Produktinvolvement (PI) auf Basis partieller Skalar-Invarianz. Die Skala reichte von 1 bis 5. Höhere Werte zeugen von höherem Produktinvolvement.

Als funktionale Low-Involvement-Marke wurde *Elmex* gewählt. Lindt, Zweifel und Rivella haben eine vergleichsweise hohe emotionale Markenpositionierung und ein vergleichsweise tiefes Involvement. Weil Lindt und Zweifel in der Vorstudie 2 von einem Online-Sample beurteilt wurden, verfügt Rivella über die höhere Vergleichbarkeit mit weiteren Marken. *Rivella* wurde deshalb als emotionale Low-Involvement-Marke für das Hauptexperiment ausgewählt.

Für die High-Involvement-Marken blieben zur Auswahl: Freitag, Odlo, Rolex, V-Zug, Flyer und Swatch. *V-Zug* verfügt über eine funktionale Markenpositionierung und ein hohes Produktinvolvement und wurde als funktionale High-Involvement-Marke gewählt.

Rolex verfügt über ein hohes Produktinvolvement und eine emotionale Markenpositionierung. Allerdings besteht die Gefahr, dass Rolex als Luxusmarke weniger gut mit den bereits gewählten Marken vergleichbar ist. Da *Freitag* über die emotionalere Positionierung als Odlo verfügt, wurde Freitag als emotionale High-Involvement-Marke gewählt.

6.5.3 Brand Content

Zur Entwicklung der *Brand-Content-Botschaften (Posts)* wurden zuerst reale Posts von Schweizer Marken und ausländischen Marken gesammelt. Durch diese Anlehnung an reale Beispiele sollte sichergestellt werden, dass die fiktiven Posts realitätsnah umgesetzt wurden. Die eigentliche Umsetzung orientierte sich an der in Abschnitt 3.10 entwickelten Definition von informativem und unterhaltendem Brand Content, wonach Brand Content eine Unternehmens- und eine Userperspektive umfasst.

Im Zentrum der *Unternehmensperspektive* steht der zu kommunizierende funktionale (beim informativen Brand Content) oder emotionale (beim unterhaltenden Brand Content) Markennutzen. Der *funktionale und der emotionale Nutzen* wurden unter Berücksichtigung realer Markendokumentationen (insbesondere Webseite der Marke und soziale Medien) unter Mithilfe von Dozierenden im Bereich Markenführung und Marketingkommunikation und von Fachleuten aus der Praxis formuliert. Durch die Anlehnung an reale Information wurde auch sichergestellt, dass sich funktionaler und emotionaler Markennutzen aus derselben Kernfunktion des Produkts ergeben. Der emotionale Markennutzen kann entsprechend als eine Abstraktion des funktionalen Markennutzens angesehen werden (Heun, 2014a, S. 39; Puto & Wells, 1984). Die Unternehmensperspektive wurde weder in der Vorstudie noch im Hauptexperiment abgefragt. Dieses Vorgehen scheint plausibel, weil auch in der Praxis Brand Manager im Sinne einer Soll-Positionierung über die zu kommunizierenden Nutzen entscheiden.

Information als User-Gratifikation wurde anhand von «Tipps & Tricks» umgesetzt. Diese verfügen über hohes Informationspotenzial, erweitern das Wissen des Users und können als Hilfestellung bei Entscheidungen helfen (Bosshart, 2007; Mast, 2019). «Tipps & Tricks» sind entsprechend ein probates Mittel für Brand Content oder Content Marketing (Kreutzer & Land, 2017, S. 268). Als Vorlage

und Inspiration dienten hierzu unter anderem ein Blogeintrag des Skiherstellers Nordica (o. J.), in dem Tipps für die Auswahl des richtigen Skischuhs gegeben wurden, und ein Post von Ricola (16. April 2018) mit Tipps zur Pflege der Stimmbänder.

Für die Unterhaltung als User-Gratifikation wurden für die Pretestphase zwei Varianten entwickelt. Eine erste Version basiert auf «Memes» und lehnt sich in ihrer Konzeptualisierung an einen Post der Swisscom (16. August 2017) an. Nach Osterroth (2015) sind Memes «Sprache-Bild-Texte, deren Bedeutungsentfaltung durch kollektive (oft hyperbolisierte) Semiose stattfindet» (S. 33). Semiose versteht Osterroth in Anlehnung an Herwig (2010) als einen Prozess, bei dem ein Ursprungs-Meme («meme-fähiges Artefakt», S. 36) durch andere User aufgenommen und weiterverbreitet wird, wobei auch neue Variationen entstehen und ebenfalls weiterverbreitet werden (Osterroth, 2015; Shifman, 2013). Bilder von Memes werden «der Popkultur, der Politik oder dem Alltag» (S. 28) entnommen und in einen neuen Kontext gestellt (Osterroth, 2015). Das Unterhaltungspotenzial von Memes entsteht unter anderem durch die Verwendung von Humor und Bildern (Horváth & Mitev, 2016).

Eine zweite, unterhaltende Version basierte auf einfachen Rätseln oder Aufgaben, die dem User gestellt wurden. Als Inspiration dienten hier unter anderen ein Post von Rivella (1. Januar 2019), bei dem User herausfinden mussten, welche von sieben Skiabfahrten auf einem Bild zum Rivella führt. Ein ähnliches Spiel setzte Ricola ein (22. Februar 2018). In einem Post von Gerber Cheese (24. April 2018) mussten die User eine Rechnungsaufgabe lösen, wobei verschiedene Gerber-Produkte für verschiedene Zahlen standen. Solche Aufgaben sind erfolgreiche Social-Media-Inhalte (Eismann, Wagner, Baccarella & Voigt, 2018). Als Beispiel einer solchen Aufgabe nennen Eismann und Kollegen einen Post von Volkswagen. Der Post zeigte ein Bild eines mit Bällen gefüllten Autos und liess die User die Anzahl der Bälle im Inneren des Autos schätzen. Solche Aufgaben benötigen kognitives Involvement, stimulieren aber auch die Neugierde (Eismann et al., 2018). Deterding, Dixon, Khaled und Nacke (2011) bezeichnen den Einsatz von «Spielementen im Non-Spiel-Kontext» als «Gamification». Gamification ist eng mit Unterhaltung verbunden (J. Zhang, 2010) bzw. kann zur Unterhaltung beitragen (Demarmels et al., 2018).

Für alle definierten Marken (Elmex, Rivella, Freitag und V-Zug) wurde jeweils ein informativer, ein memebasierter unterhaltender und ein rätselbasierter unterhaltender Post entwickelt. Ein befreundeter Designer unterstützte die grafische Darstellung; die für die rätselbasierte Version notwendigen Illustrationen kreierte ein professioneller Illustrator.

Alle zwölf Posts wurden in der Vorstudie 4 ($n = 142$) auf ihren Informations- und Unterhaltungswert sowie auf die Einstellung gegenüber dem Brand Content getestet (siehe Abschnitt 6.8.3 für die Indikatoren). Die Teilnehmenden wurden vier Gruppen zugeteilt, wobei jede Gruppe jeweils einen Post pro Version beurteilte (total drei Posts pro Gruppe). Die Samplegrößen der Gruppen betrugen 24, 34, 39 und 43 Teilnehmende. Durch die geringen Gruppengrößen (Hair et al., 2017) wurden die Itemmittelwerte klassisch in SPSS berechnet. Bei fehlenden Werten erfolgte eine listenweise Löschung. Die Mittelwerte für die einzelnen Varianten pro Marke finden sich in Tabelle 6.12.

Tabelle 6.12 Mittelwerte bezüglich Brand Content der geprüften Stimuli der Vorstudien

	Informations- gratifikation (3 Items)			Unterhaltungs- gratifikation (3 Items)			Einstellung gegenüber Brand Content (3 Items)		
	<i>M</i>	<i>Var</i>	<i>n</i>	<i>M</i>	<i>Var</i>	<i>n</i>	<i>M</i>	<i>Var</i>	<i>N</i>
<i>Informativer Brand Content</i>									
Elmex	3.82	0.049	39	2.04	0.086	38	2.66	0.061	39
Rivella	4.00	0.00	34	3.24	0.027	34	3.79	0.018	34
Freitag	4.02	0.033	43	2.36	0.211	43	3.04	0.029	42
V-Zug	4.01	0.028	25	2.33	0.137	25	3.28	0.018	26
<i>Membasierter unterhaltender Brand Content</i>									
Elmex	2.01	0.028	25	3.63	0.583	25	3.33	0.028	25
Rivella	2.00	0.008	39	3.40	0.268	38	3.39	0.030	39
Freitag	2.51	0.030	34	2.86	0.043	34	2.85	0.016	34
V-Zug	2.37	0.002	43	3.34	0.277	42	3.27	0.047	42
<i>Rätselbasierter unterhaltender Brand Content</i>									
Elmex	2.36	0.052	42	2.96	0.060	43	3.02	0.031	42
Rivella	2.00	0.00	25	3.08	0.037	26	3.45	0.066	26
Freitag	2.60	0.069	38	3.49	0.060	39	3.70	0.021	39
V-Zug	2.26	0.002	34	3.15	0.096	34	3.07	0.014	34

Anmerkungen. Berechnung der Skalenmittelwerte mit SPSS 25, listenweise Löschung bei fehlenden Werten. Die Skala reichte von 1 bis 5. Höhere Werte zeugen von höherer wahrgenommener Information und Unterhaltung bzw. von einer höheren Einstellung gegenüber dem Post.

Die Auswertungen zeigen, dass die informativen Posts bei allen vier Marken betreffend die Informationsgratifikation deutlich höher eingeschätzt werden als die beiden unterhaltenden Versionen. Die memebasierten, unterhaltenden Posts werden betreffend die Unterhaltungsgratifikation am höchsten eingeschätzt. Allerdings ist die wahrgenommene Unterhaltung des memebasierten Posts bei Rivella nur geringfügig höher als beim informativen Post. Bei Freitag wird der rätselbasierte Post als am unterhaltsamsten beurteilt. Aus Gründen der Vergleichbarkeit sollte jedoch für alle Marken jeweils dieselbe Version der Posts verwendet werden. Auf Basis des Pretests wurde die informative Version – basierend auf «Tipps & Tricks» – für das Hauptexperiment übernommen und die memebasierte, unterhaltende Version für das Hauptexperiment ausgewählt.

Sowohl bei der finalen informativen als auch bei der unterhaltenden Botschaft wurde mindestens ein Bild integriert, damit die Botschaften möglichst vergleichbar sind. Bilder sorgen im werblichen Kontext für erhöhte Aufmerksamkeit und Zuwendung zur Botschaft. Sie können schnell erfasst werden und lösen Emotionen aus (Lobinger, 2012, S. 125). Tabelle 6.13 gibt eine schriftliche Beschreibung der finalen Brand-Content-Botschaften der vier Marken.

Tabelle 6.13 Beschreibung der finalen Brand-Content-Botschaften mit Unternehmens- und Userperspektive

	Informative Botschaften	Unterhaltende Botschaften
Elmex	<p><i>(Einleitung)</i> Wie pflegt man die Zähne richtig? Mit Wissenschaftlern haben wir die 3 wichtigsten Regeln zusammengefasst.</p> <p><i>(Funktionaler Nutzen)</i> Am besten natürlich mit #elmex, weil elmex verfügt über hochwirksames Aminfluorid und schützt die Zähne langfristig und effektiv vor Karies.</p> <p><i>(Information; Bild einer Zahnärztin und zweier Kunden, Produktbild)</i> 3 einfache Tipps, um die Zähne richtig zu putzen.</p> <ul style="list-style-type: none">– 2-mal täglich: Die Zähne mindestens morgens und abends putzen.– Von rot nach weiss: Vom Zahnfleisch her hin zur Zahnkrone reinigen.– Nicht sofort nach dem Essen: Nach dem Essen ca. 30 Minuten warten bis zum Reinigen.	<p><i>(Einleitung)</i> In jeder Situation eine Geheimwaffe: Dein schönstes und persönlichstes Lächeln.</p> <p><i>(Emotionaler Nutzen)</i> Dank dem Aminfluorid in der #elmex schützt und bewahrst du dein Lächeln für die wirklich wichtigen Momente.</p> <p><i>(Unterhaltung)</i> Immer schön lächeln</p> <ul style="list-style-type: none">– (Bild eines Sturzes vom Fahrrad und Tränen-lachendes Emoticon)– (Bild einer Frau mit Daumen hoch und Emoticon mit dunkler Sonnenbrille)– (Bild zweier lachender Frauen und Tränen-lachendes Emoticon)– (Bild eines Paares und Emoticon mit Sternenaugen)– (Bild einer Frau beim Frühstück und Emoticon, welches Zunge seitlich rausstreckt)– (Bild eines mit Farbe verschmutzten Kindes und grinsendes Emoticon)

(Fortsetzung)

Tabelle 6.13 (Fortsetzung)

	Informative Botschaften	Unterhaltende Botschaften
Rivella	<p><i>(Einleitung)</i> Beim Wandern die CH entdecken ist aufregend – und anstrengend. Wir haben 3 Tipps zusammengestellt, damit euch das Wandern nächstes Mal etwas leichter fällt.</p> <p><i>(Funktionaler Nutzen)</i> Am besten klappt's mit einem #rivellarefresh: Die weniger süsse und leichtere Variante des Originals.</p> <p><i>(Information; Bild dreier Wanderer von hinten, Produktbild)</i> 3 einfache Tipps für eine sichere und erfolgreiche Wanderung.</p> <ul style="list-style-type: none"> – Planung ist das A und O. Den Weg, die Dauer und das geographische Umfeld der Wanderung vorab anschauen. – Die richtige Ausrüstung. Bei jeder Wanderung ein Muss: geeignetes Schuhwerk, Regenjacke und Smartphone. – Vertrauen ist gut, Kontrolle ist besser. Unterwegs immer prüfen, ob man im Zeitplan liegt und wie sich das Wetter entwickelt. 	<p><i>(Emotionaler Nutzen)</i> #rivellarefresh ist die weniger süsse und leichtere Variante des Originals. Damit fällt dir im Leben alles etwas leichter.</p> <p><i>(Übergang)</i> Schnapp dir ein Rivella und mach es wie die Leute unten.</p> <p><i>(Unterhaltung)</i> Ich mit einem Rivella</p> <ul style="list-style-type: none"> – (Bild einer Frau am Schreibtisch mit riesigem Papierberg) easy, mache ich – (Bild eines Eiskletterers) easy, ich komme – (Bild einer Frau in komplizierter Yoga-Stellung) easy, kann ich – (Bild eines mit Karton-Schachteln überladenen Mopeds) easy, bringe es – (Bild eines Jungen vor dem Greifautomaten) easy, hole ich – (Bild eines mit Farbe verschmutzten Kindes) easy, putze ich
Freitag	<p><i>(Einleitung)</i> Wir haben bei der CH Materialprüfungsanstalt nachgefragt, wie man die Qualität guter Freizeittaschen erkennt.</p> <p><i>(Funktionaler Nutzen)</i> Resistente Lastwagenplanen, gebrauchter Sicherheitsgurt zum Tragen und Umhängen, Schweizer Qualität – da halten wir uns bei #freitag nicht so schlecht, oder?</p> <p><i>(Information; Bild einer Frau in Näherei, Produktbild)</i> 3 einfache Tipps, um die Qualität von Freizeittaschen zu beurteilen.</p> <ul style="list-style-type: none"> – Material prüfen: Wie fühlt sich das Material an? Lässt es sich biegen und formen? Ist es resistent und wasserfest? – Nähte prüfen: Sind die Nähte gleichmässig und fest vernäht? – Tragegriffe prüfen: Sind die Tragegriffe breit, bequem und weich? 	<p><i>(Emotionaler Nutzen)</i> Dank den resistenten Lastwagenplanen, die wir bei #freitag für unsere Produkte benutzen, sind unsere Taschen nicht nur funktional, sondern auch absolut einzigartig.</p> <p><i>(Übergang)</i> So stichst du nicht nur aus der Masse heraus, sondern machst auch immer eine gute Figur.</p> <p><i>(Unterhaltung)</i> einzigartig und funktional?</p> <ul style="list-style-type: none"> – (Bild einer verbogenen Gabel, markiert mit rotem Kreuzchen) – (Bild eines Beils mit Haupt und Blatt aus Holz, markiert mit rotem Kreuzchen) – (Bild einer Brille mit nur einem, mittigen Glas, markiert mit rotem Kreuzchen) – (Bild eines Fahrrads ohne Pedalen, markiert mit rotem Kreuzchen) – (Bild einer Giesskanne, deren Ausfluss in die Kanne zurückfliesst, markiert mit rotem Kreuzchen) – (Bild einer Freitag-Freizeittasche, markiert mit grünem Häkchen)

(Fortsetzung)

Tabelle 6.13 (Fortsetzung)

	Informative Botschaften	Unterhaltende Botschaften
V-Zug	<p><i>(Einleitung)</i> Wie werden Kleider schonend gewaschen und langfristig gepflegt? Wir haben drei Tipps zusammengestellt.</p> <p><i>(Funktionaler Nutzen)</i> Das Waschen klappt natürlich am besten mit einer #vzug. Dank Schontrommel mit 14'000 speziell geformten Öffnungen bietet V-Zug maximale Schonung für eure Wäsche.</p> <p><i>(Information; Bild einer Wäscheleine, Produktbild)</i> 3 einfache Tipps, um Kleider zu pflegen und zu schonen.</p> <ul style="list-style-type: none">– Wenig waschen: Manchmal reicht es, ein Kleidungsstück an die frische Luft zu hängen und erst beim nächsten Mal wieder zu waschen.– Flecken sofort reinigen: Wenn die Lieblingskleider verschmutzt sind, dann sofort reagieren. Am besten mit einem speziellen Fleckenstift.– Wäschesack verwenden: Empfindliche Textilien in einem Wäschesack oder leichten Kissenbezug waschen.	<p><i>(Emotionaler Nutzen)</i> #vzug bietet maximale Schonung für eure Wäsche. So sind die Lieblingskleider immer optimal gepflegt für den nächsten grossen Auftritt.</p> <p><i>(Übergang)</i> Da darf auch mal ein kleines Ungeschick passieren...</p> <p><i>(Unterhaltung)</i> Ich dank meiner V-Zug</p> <ul style="list-style-type: none">– (Bild eines Sturzes vom Fahrrad) easy– (Bild eines Sturzes vom Fahrrad) easy– (Bild eines mit Tomatensauce besudelten Mannes) easy– (Bild eines Kindes in einer Pfütze) easy– (Bild einer Frau, die in einen Bach fällt) easy– (Bild eines mit Farbe verschmutzten Kindes) easy, oder?

Anmerkungen. Notizen in den Klammern dienen dem Verständnis für die Leserin und den Leser und waren nicht Teil der finalen Botschaften.

6.6 Hauptexperiment

Als empirischer Hauptteil dieser Arbeit wurde im Frühling 2019 – basierend auf den Vorarbeiten in den Vorstudien – ein 2 x 2 x 2 x 2-between-subject-Experiment durchgeführt (siehe Tabelle 6.14).

Tabelle 6.14 Versuchsplan des Hauptexperiments

Unabhängige-Variablen	<ul style="list-style-type: none">– Manipulation der Marke in Elmex (funktionale Low-Involvement-Marke), Rivella (emotional, Low Involvement), Freitag (emotional, High Involvement) und V-Zug (funktional, High Involvement)– Manipulation des Kanals in Twitter (informativer Kanal durch primäre Informationserwartungen) und Instagram (unterhaltender Kanal durch primäre Unterhaltungserwartungen)– Manipulation des Brand Contents in informativen Brand Content und unterhaltenden Brand Content
-----------------------	--

(Fortsetzung)

Tabelle 6.14 (Fortsetzung)

Experimental- gruppen	Gruppe 1 ($n = 89$) Marke: Elmex Kanal: Instagram BC: Informativ	Gruppe 2 ($n = 87$) Marke: Elmex Kanal: Instagram BC: Unterhaltend	Gruppe 3 ($n = 84$) Marke: Elmex Kanal: Twitter BC: Informativ	Gruppe 4 ($n = 85$) Marke: Elmex Kanal: Twitter BC: Unterhaltend
	Gruppe 5 ($n = 90$) Marke: Rivella Kanal: Instagram BC: Informativ	Gruppe 6 ($n = 89$) Marke: Rivella Kanal: Instagram BC: Unterhaltend	Gruppe 7 ($n = 83$) Marke: Rivella Kanal: Twitter BC: Informativ	Gruppe 8 ($n = 82$) Marke: Rivella Kanal: Twitter BC: Unterhaltend
	Gruppe 9 ($n = 87$) Marke: Freitag Kanal: Instagram BC: Informativ	Gruppe 10 ($n = 86$) Marke: Freitag Kanal: Instagram BC: Unterhaltend	Gruppe 11 ($n = 84$) Marke: Freitag Kanal: Twitter BC: Informativ	Gruppe 12 ($n = 85$) Marke: Freitag Kanal: Twitter BC: Unterhaltend
	Gruppe 13 ($n = 89$) Marke: V-Zug Kanal: Instagram BC: Informativ	Gruppe 14 ($n = 82$) Marke: V-Zug Kanal: Instagram BC: Unterhaltend	Gruppe 15 ($n = 91$) Marke: V-Zug Kanal: Twitter BC: Informativ	Gruppe 16 ($n = 81$) Marke: V-Zug Kanal: Twitter BC: Unterhaltend
Abhängige Variablen bzw. Mediatoren*	<ul style="list-style-type: none"> – *Wahrgenommene Information (BC_IG) und Unterhaltung (BC_UG) – Einstellung gegenüber dem Brand Content (BC_ES) – Funktionales (M_FI) und emotionales Markenimage (M_EI) – Globalimage der Marke (M_GI) – Markenbezogenes Verhalten (M_MV) 			

Anmerkungen. BC = Brand Content. n bezieht sich auf die bereinigte Stichprobe.

6.6.1 Beschreibung der Stichprobe

Für das Hauptexperiment wurden 1'600 Personen aus dem Online-Panel der deutschsprachigen Schweiz des Marktforschungsinstituts Bilendi rekrutiert. Die Befragung fand vom 28. März 2019 bis am 11. April 2019 statt. Die eingeladene Stichprobe war ausgeglichen bezüglich Geschlecht (50 % Frauen und 50 % Männer) und Bildung (50 % mit tertiärer Bildung und 50 % ohne tertiäre Bildung), tendenziell aber etwas jünger als der Schweizer Durchschnitt (60 % zwischen 18 und 39 Jahren, 40 % zwischen 40 und 65 Jahren), weil jüngere Personen soziale Medien stärker nutzen. Die Teilnehmenden erhielten für die Beantwortung des experimentellen Fragebogens über das Marktforschungsinstitut eine Entschädigung. Das finale Sample nach der Datenbereinigung (siehe Abschnitt 6.7) betrug 1'374 Personen.

Abbildung 6.2 Internetnutzung und Nutzung sozialer Medien bei der Stichprobe. (Eigene Darstellung auf Basis der erhobenen Daten im Hauptexperiment ($n = 1'374$))

Abbildung 6.2 gibt einen Überblick über die Nutzung digitaler Kanäle der Stichprobe. Es zeigt sich, dass über 95 % aller Teilnehmerinnen und Teilnehmer mindestens einmal am Tag im Internet sind. Knapp 60 % nutzen zudem täglich soziale Medien. Bei den einzelnen Kanälen oder sozialen Medien zeigen sich hingegen grosse Unterschiede. Während die Hälfte der Befragten täglich Facebook nutzt, konsultieren nur gerade 6.2 % der Befragten täglich Twitter. Bei YouTube beläuft sich der Kreis der täglichen User auf 39.1 %, bei Instagram auf 34.9 %. Über 80 % der Befragten nutzen Twitter nie oder maximal einmal pro Monat, bei Instagram beläuft sich diese Zahl auf 49.6 %, bei Facebook auf 28.3 % und bei YouTube auf geringe 11.3 %.

171 der 1'374 befragten Personen (12.4 %) folgen Rivella auf sozialen Medien. Bei Elmex (4.7 %), Freitag (3.9 %) und V-Zug (5.7 %) fällt dieser Wert tiefer aus.

6.6.2 Bildung der Experimentalgruppen

Die Stimuli wurden zufällig den Teilnehmenden gezeigt (*Randomisierung*). Alle 16 Stimuli konnten dabei maximal 100 Mal gezeigt werden. Wurde ein Stimulus 100 Mal gezeigt, wurde er geblockt und nicht weiter angezeigt. Auf eine Parallelisierung wurde aufgrund der Gruppengrösse ($n = 100$) verzichtet (Koch et al., 2019, S. 102).

6.6.3 Aufbau des Fragebogens

Der Fragebogen (siehe elektronisches Zusatzmaterial) startete mit den Fragen zu Geschlecht, Alter und Bildung zur Kontrolle der Quotenvorgaben. Danach wurden den Teilnehmenden in Block 1 vier allgemeine Fragen zu den Marken gestellt: Kaufverhalten der Produkte der Marken, Markenpositionierung der Marken (Manipulationscheck), Ex-ante-Einstellung gegenüber den Marken (ein Item) und ob den Marken auf sozialen Medien gefolgt wird. In einem zweiten Block wurde die Nutzung des Internets und von sozialen Medien sowie die Social-Media-Affinität abgefragt. Im dritten Block erfolgte der Manipulationscheck des Experimentalkanals. Dazu wurden die Informationserwartungen, die Unterhaltungserwartungen und die Einstellung gegenüber dem Kanal erhoben. Es folgten im vierten Block Fragen zum Produktinvolvement der Marke. In Block 5 wurde einer der 16 Experimentalstimuli für eine Minute eingeblendet. Der Stimulus wurde zufällig ausgewählt, wobei jede Botschaft maximal 100 Mal gezeigt wurde.

Obwohl alle Marken über eine hohe Bekanntheit in der Schweiz verfügen, wurden die *Marken im Fragebogen kurz vorgestellt* (siehe Mendini et al., 2018, S. 199). Die Informationen wurden den offiziellen Kanälen (Webseite, Instagram, Twitter, LinkedIn etc.) entnommen und wo nötig für den Rahmen dieser Arbeit umformuliert.

Elmex: Elmex entwickelt Zahnpflegeprodukte. Die Marke wurde 1963 von der Basler GABA AG lanciert, nachdem das Unternehmen Aminfluoride zum Schutz der Zähne entdeckt hatte. Elmex hilft, Zahnschmelz und Karies zu verhindern, und sorgt «ein Leben lang für gesunde Zähne».

Rivella: Rivella ist ein Schweizer Erfrischungsgetränk auf der Basis von Milchserum. Das Unternehmen Rivella wurde 1952 in Rothrist im Kanton Aargau gegründet. Mit seinem einzigartigen Geschmack steht Rivella für erfrischende Momente und das unverwechselbare Lebensgefühl der aktiven Schweiz.

Freitag: Freitag ist ein Schweizer Modeunternehmen, das Freizeittaschen aus alten Lastwagenplanen produziert. Das Unternehmen wurde 1993 von den Grafikdesignern Markus und Daniel Freitag in Zürich gegründet. Mit den farbigen, rezyklierten und individuellen Produkten wurde Freitag zum inoffiziellen Ausrüster aller urbanen, velofahrenden Individualistinnen und Individualisten.

V-Zug: V-Zug ist der Schweizer Marktführer für Haushaltsgeräte. Das traditionelle Industrieunternehmen wurde 1913 in Zug gegründet und hat bis heute dort seinen Hauptsitz. Mit den beiden Hauptproduktgruppen Küchengeräte und Waschmaschinen steht V-Zug für «Schweizer Perfektion für zuhause».

Es folgten die Fragen zur wahrgenommenen Information, zur wahrgenommenen Unterhaltung und zur Einstellung gegenüber dem Brand Content. Block 6 beinhaltete die Items zum funktionalen und emotionalen Markenimage, zum Globalimage der Marke sowie zum markenbezogenen Verhalten. Die restlichen demografischen Angaben bildeten in Block 7 den Abschluss des Fragebogens.

Während die Fragen in den Blöcken 1 und 2 allgemein gehalten waren, bezogen sich die Fragen in den Blöcken 3 (Kanal), 4 (Produktinvolvement), 5 (Brand Content) und 6 (Markenimage) jeweils nur auf die im gezeigten Stimulus verwendeten Kanäle, Produkte, Brand-Content-Botschaften und Marken.

6.7 Datenbereinigung und Datenprüfung

In diesem Kapitel wird auf die Datenbereinigung eingegangen.

6.7.1 Durchklicker und Straightlining

Der ursprüngliche Datensatz ($n = 1'600$) wurde in einem mehrstufigen Verfahren bereinigt (Eilers, 2014, S. 92; Weiber & Mühlhaus, 2014, S. 174). In einem ersten Schritt wurde der Datensatz auf Basis der *Beantwortungszeit* bereinigt. Die Variable *Beantwortungszeit* hatte einen unbereinigten Mittelwert von 14 Minuten und 30 Sekunden und einen Median von 7 Minuten und 48 Sekunden bei diversen Ausreissern nach oben. In einem ersten Schritt wurden die obersten 1 % und die untersten 1 % der Fälle gemäss der Variable *Beantwortungszeit* aus dem Datensatz entfernt (total 32 Fälle). Es blieben 1'568 Fälle übrig.

Der Mittelwert der Beantwortungszeit betrug nach dieser ersten Bereinigung 10 Minuten und 38 Sekunden. In einem zweiten Schritt wurden Durchklicker aus dem Datensatz entfernt. Bei Durchklickern handelt es sich um Personen, bei denen aufgrund einer vergleichsmässig tiefen Beantwortungszeit davon ausgegangen werden muss, dass sie die Fragen nicht tatsächlich beantwortet haben, sondern sich zufällig und möglichst rasch durch den Fragebogen geklickt haben. Als Schwellenwert der minimalen Beantwortungszeit wurden 5 Minuten festgelegt – etwas weniger als die Hälfte der durchschnittlichen Beantwortungszeit (Eilers, 2014, S. 93). 121 Personen, die weniger als 5 Minuten zur Beantwortung benötigten, wurden ausgeschlossen.

In einem zweiten Schritt wurde das Antwortverhalten über die zentralen Variablen «Kanal», «Produktinvolvement», «Brand Content», «Markenimage» und «Markenverhalten» analysiert. 26 Personen hatten über diese 38 Variablen hinweg eine Standardabweichung von 0. Die Analyse dieser Fälle zeigte, dass die Personen entweder fehlende Werte aufweisen (z. B. in Form von «ich weiss nicht») oder alle (restlichen) Variablen mit der gleichen Option auf der Likert-Skala beantworteten. Das Verhalten, alle Fragen einer Skala mit demselben Wert zu beantworten, wird als *Straightlining* (Rossmann, 2017, S. 86) bezeichnet. In einem dritten Schritt wurde dasselbe Vorgehen für die 18 Variablen zum «Brand Content» und «Markenimage» wiederholt. Es wurden hier weitere 49 Fälle mit gleichbleibendem Antwortverhalten ausgeschlossen.

6.7.2 Finaler Datensatz und finale Stichprobe

Insgesamt verfügt der *finale Datensatz* über 1'374 Fälle. Eine Analyse der demografischen Variablen zeigte, dass sich der Datensatz durch die Löschung der genannten Fälle in seiner Gewichtung nur marginal verändert hat (siehe elektronisches Zusatzmaterial Tabelle A-4). Die genannten Quoten sind auch im bereinigten Datensatz klar zu erkennen. Auch bei den demografischen Variablen ohne Quotenvorgabe (Wohnort, Erwerbstätigkeit und Haushaltseinkommen) gab es keine nennenswerten Verschiebungen. Schliesslich hat sich auch die Verteilung der manipulierten Experimentalvariablen nicht grundlegend geändert.

6.7.3 Fehlende Werte

Bei der Auswertung eines Datensatzes müssen fehlende Werte eine entsprechende Beachtung finden. Es können drei Typen fehlender Werte nach deren Zustandekommen unterschieden werden (Lüdtke, Robitzsch, Trautwein & Köller, 2007):

- *Missing Completely At Random (MCAR)*: Fehlende Werte einer Variable X sind zufällig über die Stichprobe verteilt. Die fehlenden Werte können nicht durch andere Variablen erklärt werden.
- *Missing At Random (MAR)*: Fehlende Werte einer Variable X können durch Drittvariablen erklärt werden. Die Ausprägung der Variable X selbst hat keinen Einfluss auf das Zustandekommen der fehlenden Werte. Bei Lüdtke et al. (2007) findet sich das Beispiel von Einkommen und Familienstand. Angaben zum Einkommen sind MAR, wenn fehlende Werte auf den Familienstand zurückgeführt werden können, innerhalb unterschiedlicher Familienstandsgruppen fehlende Einkommenswerte aber wiederum MCAR sind.
- *Missing Not At Random (MNAR)*: Wenn fehlende Werte einer Variable X auch nach Berücksichtigung weiterer im Datensatz vorhandener Variablen von der Variable X selbst abhängig sind, dann spricht man von MNAR. Lüdtke et al. (2007) führen hier als Beispiel auf, wenn leistungsschwache Schüler an «Schulleistungsuntersuchungen» nicht teilnehmen, unabhängig von Drittvariablen wie dem sozialen Status der Eltern.

Kline (2016, S. 83) sieht es als unproblematisch an, wenn in einem Datensatz weniger als 5 % der Daten fehlen. Insbesondere ab 10 % fehlender Werte müssen diese detailliert betrachtet und allenfalls ersetzt werden (Hair et al., 2017). Ignoriert man automatisch erfasste Variablen wie «Beantwortungszeit» oder «verwendeter Browser» und sämtliche eigenständig kreierte Variablen (Umkodierungen, Transformationen), dann verfügt der Gesamtdatensatz der 1'374 analysierten Fälle über 4.4 % fehlende Werte. Werden nur die «zentralen» Variablen (Variablen des Gesamtmodells und Produktinvolvement) berücksichtigt, dann steigt der *Prozentsatz fehlender Werte auf 5.9 %*. Insofern liegt die Anzahl fehlender Werte bei den zentralen Variablen dieser Arbeit marginal über der Schwelle von 5 %, aber deutlich unter der Schwelle von 10 %. Im Folgenden werden diejenigen Variablen mit der höchsten Anzahl fehlender Werte näher untersucht.

Eine Analyse des Datensatzes zeigt, dass die Marken V-Zug und insbesondere Freitag nicht bei allen Testpersonen bekannt ist. Bei der Frage, ob die jeweiligen

Marken schon einmal gekauft worden sind oder ein Kauf in Erwägung gezogen worden ist, gaben 14.2 % (Freitag) bzw. 8.2 % (V-Zug) an, die Antwort nicht zu wissen oder die Marke nicht zu kennen. Diese Antwort wurde nicht als «fehlend» kodiert, der Umstand wirkt sich jedoch auf das Aufkommen fehlender Werte (i. S. v. «ich weiss nicht / ich kenne die Marke nicht») bei zwei Folgefragen aus. 23.6 % der Befragten konnten nicht einschätzen, ob es sich beim Kauf von Freitag-Produkten um eine rationale/kopforientierte oder eine emotionale/gefühlsorientierte Entscheidung handelt (12.7 % bei V-Zug). 26 % konnten ex ante nicht einschätzen, inwiefern Freitag eine gute Marke ist (12.6 % bei V-Zug).

Weiter zeigte eine Analyse des Datensatzes, dass die meisten fehlenden Werte bei den Gratifikationserwartungen gegenüber dem Kanal (insgesamt 6 Items), bei der Bewertung der Einstellung gegenüber dem Kanal (3 Items) und bei der Bewertung des funktionalen Markenimages (4 Items) vorkommen. Die Variable *K_ES_1* wies mit 13.9 % fehlenden Werten den höchsten Wert auf. Hier hängt die Verteilung der fehlenden Werte vom beurteilten Kanal ab. Musste Twitter beurteilt werden, dann sind vermehrt fehlende Werte zu finden als bei Instagram. Das ist darauf zurückzuführen, dass Twitter deutlich weniger genutzt wird als Instagram.

Zur Handhabung fehlender Werte gibt es unterschiedliche Lösungen (Lüdtke et al., 2007, S. 105–107). Lüdtke et al. unterscheiden zwischen klassischen, imputationsbasierten und modellbasierten Verfahren. Fall- und paarweiser Ausschluss sind Beispiele eines klassischen Verfahrens, die Ersetzung des fehlenden Wertes durch den Mittelwert ist ein Beispiel für ein imputationsbasiertes Verfahren. Bei Strukturgleichungsmodellen bieten sich modellbasierte Verfahren im *Rahmen der Maximum-Likelihood-Schätzung* an (Weiber & Mülhhaus, 2014, S. 176). Dieses Vorgehen wird auch bei *Mplus* für alle drei Varianten fehlender Werte (MCAR, MAR, MNAR) eingesetzt (Muthén & Muthén, 2017, S. 443) und kommt damit im Rahmen dieser Arbeit zur Anwendung.

6.8 Operationalisierung und Prüfung der Variablen

In den folgenden Kapiteln wird beschrieben, wie die latenten Variablen dieser Arbeit operationalisiert, also messbar gemacht werden. Im ersten Schritt werden dazu übergreifende Resultate einer *gemeinsamen konfirmatorischen Faktorenanalyse* (CFA) vorgestellt (Hair et al., 2017). Danach werden die latenten Variablen einzeln besprochen, wobei auf die Resultate der CFA eingegangen wird. Tabelle 6.15 gibt einen Überblick über die latenten Variablen dieser Arbeit.

Tabelle 6.15 Übersicht über die latenten Variablen

Latente Variable	Abkürzung	Indikatoren
Informationsgratifikation des Brand Contents	BC_IG	BC_IG_1 bis BC_IG_3
Unterhaltungsgratifikation des Brand Contents	BC_UG	BC_UG_1 bis BC_UG_4*
Einstellung gegenüber dem Brand Content	BC_ES	BC_ES_1 bis BC_ES_4
Informationserwartungen gegenüber Kanal	K_IE	K_IE_1 bis K_IE_3
Unterhaltungserwartungen gegenüber Kanal	K_UE	K_UE_1 bis K_UE_3
Einstellung gegenüber Kanal	K_ES	K_ES_1 bis K_ES_3
Funktionales Markenimage	M_FI	M_FI_1 bis M_FI_4
Emotionales Markenimage	M_EI	M_EI_1 bis M_EI_3
Globalimage der Marke	M_GI	M_GI_1 bis M_GI_3
Markenbezogenes Verhalten I	M_MV_I	M_MV_1 und M_MV_3
Markenbezogenes Verhalten II	M_MV_II	M_MV_2 und M_MV_4
Produktinvolvement	PI	PI_1 bis P_3
Social-Media-Affinität	SM	SM_1 bis SM_3

Anmerkungen. *BC_UG_2 wird auf Basis der konfirmatorischen Faktorenanalyse entfernt und nicht in die Auswertung einbezogen (siehe Abschnitt 6.8.3).

6.8.1 Konfirmatorische Faktorenanalyse

Im ersten Schritt wurden zusammengehörende latente Variablen (z. B. Variablen zu Kanalgratifikationen, Variablen zu Brand Content, Variablen zum Markenimage) in separaten, partiellen Modellen geprüft. Auf Basis dieser Resultate wurden wenige Anpassungen vorgenommen, die in den jeweiligen Kapiteln beschrieben sind. Danach wurden alle latenten Variablen einer *gemeinsamen CFA* unterzogen (Bruhn et al., 2012; Hair et al., 2017; Kleinke et al., 2017). Die CFA wurde einmal mit Maximum Likelihood (ML) und einmal mit Maximum Likelihood With Robust Standard Errors (MLR) geschätzt. MLR liefert robuste Resultate, wenn die Daten nicht normalverteilt sind bzw. wenn nicht zweifelsfrei von Normalverteilung ausgegangen werden kann (Muthén & Muthén, 2017).

Eine Prüfung der zentralen Variablen auf Schiefe und Wölbung ergibt keine kritischen Grössen gemäss den in Abschnitt 6.4.9 genannten Richtwerten. Die höchsten Werte für Schiefe (−1.4) und Wölbung (2.6) ergaben sich für die Variable «Die Produkte der Marke haben eine gute Qualität» (M_FI_1). Weil aufgrund fehlender Werte kein Mardias-Test auf multivariate Normalverteilung gemacht werden kann (Weiber & Mühlhaus, 2014, S. 181), bietet es sich an, eine

MLR-Schätzung zu prüfen. Die Gegenüberstellung der beiden Modellschätzungen in Tabelle 6.16 zeigt, dass beide Modelle eine hohe Güte aufweisen (Hair et al., 2014, S. 584), der Modell-Fit mit der MLR-Schätzung allerdings marginal besser ist. Deshalb muss davon ausgegangen werden, dass die Daten nicht perfekt normalverteilt sind. Im weiteren Verlauf wird deshalb die MLR-Schätzung verwendet.

Tabelle 6.16 Prüfung und Gegenüberstellung der Schätzmethoden ML und MLR

Modell-Nr.	Schätz-methode	Df	χ^2	χ^2/df	CFI	TLI	RMSEA	SRMR
1	ML	624	1912.015	3.06	0.967	0.961	0.039 ^a	0.034
2	MLR	624	1705.615	2.73	0.968	0.962	0.036 ^b	0.034

Anmerkungen. ML = Maximum-Likelihood-Schätzung mit konventionellen Standardfehlern, MLR = Maximum-Likelihood-Schätzung mit robusten Standardfehlern gegenüber Nicht-Normalverteilung (Muthén & Muthén, 2017, S. 667); CFI = Comparative Fit Index, TLI = Tucker Lewis Fit Index, RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square Residual.

^a Mit 90 % K.I. zwischen 0.037 und 0.041. ^b Mit 90 % K.I. zwischen 0.034 und 0.038.

Tabelle 6.16 zeigt, dass die Gesamt-CFA über eine hohe Modellgüte verfügt und die theoretisch hergeleiteten Operationalisierungen der reflektiven Konstrukte durch die Daten abgebildet werden. Für die Berechnung wurde zur Modellidentifikation die Varianz der latenten Variablen auf 1 festgelegt (Weiber & Mühlhaus, 2014, S. 153).

Die in den jeweiligen Kapiteln aufgeführten Resultate zu den einzelnen Indikatoren und Konstrukten beziehen sich auf diese gesamtheitliche CFA mit MLR-Schätzung. Zur Prüfung von Reliabilität und Validität wurden zudem die durch die CFA erhaltenen Werte zu Faktorladungen, Indikatorreliabilität, Faktorreliabilität und durchschnittlich extrahierter Varianz ausgewiesen bzw. berechnet (siehe Abschnitt 6.4.4). Die Berechnung erfolgte in Excel auf Basis des Mplus-Outputs und der Formeln durch Weiber und Mühlhaus (2014, S. 153–154). Zudem wird die Item-Korrelation anhand des klassischen Wertes Cronbach’s Alpha ausgewiesen (Bruhn et al., 2012). Zur Berechnung des Cronbach’s Alpha wurde bei fehlenden Werten eine listenweise Löschung angewandt. Die Berechnung erfolgte mit SPSS.

Bevor die einzelnen latenten Variablen besprochen werden, werden die Themen Multikollinearität und diskriminante Validität diskutiert, weil dazu die Resultate der CFA aller latenten Variablen gegenübergestellt werden müssen. Eine Auswertung der Korrelationskoeffizienten auf Basis der CFA zeigt gemäss

Tabelle 6.17, dass *keine problematischen Korrelationen* im Sinne einer Multikollinearität gemäss dem genannten *Schwellenwert von 0.9* vorliegen (Kühnel & Krebs, 2007, S. 545). Allerdings erreichen gewisse Korrelationen hohe Werte von über 0.8.

Tabelle 6.17 Korrelationen der latenten Variablen

LV	1	2	3	4	5	6	7	8	9	10	11	12	13
1 BC_IG	1.00												
2 BC_UG	0.42	1.00											
3 BC_ES	0.76	<i>0.83</i>	1.00										
4 K_IE	0.22	0.28	0.24	1.00									
5 K_UE	0.17	0.20	0.20	0.71	1.00								
6 K_ES	0.18	0.19	0.19	0.64	0.69	1.00							
7 M_FI	0.33	0.30	0.38	0.15	0.18	0.16	1.00						
8 M_EI	0.34	0.44	0.43	0.23	0.23	0.21	0.78	1.00					
9 M_GI	0.28	0.35	0.39	0.09	0.15	0.13	<i>0.86</i>	<i>0.87</i>	1.00				
10 M_MV_I	0.28	0.36	0.37	0.15	0.17	0.16	0.75	<i>0.82</i>	<i>0.89</i>	1.00			
11 M_MV_II	0.45	0.58	0.56	0.33	0.21	0.25	0.42	0.63	0.49	0.64	1.00		
12 PI	0.18	0.17	0.12	0.11	0.11	0.07	0.20	0.36	0.22	0.24	0.41	1.00	
13 SM	0.13	0.18	0.15	0.46	0.46	0.44	0.07	0.20	0.08	0.11	0.20	0.12	1.00

Anmerkungen. Korrelationen über 0.8 sind kursiv.

Zur *Prüfung diskriminanter Validität* wird die durchschnittlich extrahierte Varianz (DEV) pro latenter Variable (in der Diagonale in Tabelle 6.18) den quadrierten Korrelationen zwischen den latenten Variablen (die restlichen Einträge in Tabelle 6.18) gegenübergestellt (Weiber & Mühlhaus, 2014, S. 165). In den meisten Fällen ist diskriminante Validität gegeben. Beim funktionalen Markenimage (M_FI) ist die DEV kleiner als die gemeinsame Varianz (oder quadrierte Korrelation) mit dem emotionalen Markenimage (M_EI) und mit dem Globalimage (M_GI). Beim emotionalen Markenimage ist die DEV kleiner als die gemeinsame Varianz mit dem Globalimage. Beim markenbezogenen Verhalten (M_MV_I) ist die DEV kleiner als die gemeinsame Varianz mit dem Globalimage. In den vier Fällen sind die Werte der DEV nur marginal tiefer als die quadrierte Korrelation. Weil die latenten Variablen Teilaspekte des Markenimages bzw. des markenbezogenen Verhaltens messen, ist deren hohe Korrelation keine Überraschung. Zudem

ist das Fornell-Larcker-Kriterium eine strenge Prüfgrösse diskriminanter Validität (siehe Abschnitt 6.4.4). Eine weniger strenge Variante besteht darin, zu zeigen, dass die Korrelation der latenten Variablen kleiner als 1 ist. Ein Signifikanztest in Mplus zeigt, dass die Korrelationen zwischen M_FI und M_EI ($p < 0.001$), zwischen M_FI und M_GI ($p < 0.001$), zwischen M_EI und M_GI ($p < 0.001$) und zwischen M_GI und M_MV_I ($p < 0.001$) alle kleiner als 1 sind. Insofern kann davon ausgegangen werden, dass keine relevante Verletzung diskriminanter Validität vorliegt.

Tabelle 6.18 Test auf diskriminante Validität der latenten Variablen

LV	1	2	3	4	5	6	7	8	9	10	11	12	13
1 BC_IG	0.79												
2 BC_UG	0.17	0.77											
3 BC_ES	0.57	0.68	0.79										
4 K_IE	0.05	0.08	0.06	0.60									
5 K_UE	0.03	0.04	0.04	0.50	0.67								
6 K_ES	0.03	0.04	0.04	0.41	0.48	0.91							
7 M_FI	0.11	0.09	0.14	0.02	0.03	0.02	0.58						
8 M_EI	0.11	0.20	0.19	0.05	0.05	0.04	0.61	0.73					
9 M_GI	0.08	0.12	0.15	0.01	0.02	0.02	0.73	0.76	0.82				
10 M_MV_I	0.08	0.13	0.14	0.02	0.03	0.02	0.56	0.68	0.78	0.74			
11 M_MV_II	0.21	0.33	0.31	0.11	0.04	0.06	0.18	0.40	0.24	0.40	0.56		
12 PI	0.03	0.03	0.01	0.01	0.01	0.01	0.04	0.13	0.05	0.06	0.16	0.53	
13 SM	0.02	0.03	0.02	0.21	0.21	0.19	0.01	0.04	0.01	0.01	0.04	0.01	0.59

Anmerkungen. Diagonale: DEV, sonstige Zellen = quadrierte Korrelation.

6.8.2 Verwendete Skalen

Die Indikatoren aller latenten Variablen von «Brand Content», «Markenimage» und «Produktinvolvement» sowie diejenigen von «Einstellung gegenüber dem Kanal» wurden auf einer fünfstufigen Likert-Skala mit dem Anfangswert *trifft überhaupt nicht zu* (1) und dem Endwert *trifft voll und ganz zu* (5) gemessen (Horstmann, 2017, S. 145).

Tabelle 6.19 Beispielhafte Likert-Skala zur Erhebung der Zustimmung

	trifft überhaupt nicht zu	trifft eher nicht zu	teils, teils	trifft eher zu	trifft voll und ganz zu	ich weiss nicht
Indikator A	☐	☐	☐	☐	☐	☐
Indikator B	☐	☐	☐	☐	☐	☐
etc.	☐	☐	☐	☐	☐	☐

Anmerkungen. Nach Hoffmann (2017).

Die Indikatoren der latenten Variablen zu «Informationsgratifikation des Kanals», «Unterhaltungsgratifikation des Kanals» und «markenbezogenes Verhalten» wurden auf einer fünfstufigen Likert-Skala mit dem Anfangswert *sehr unwahrscheinlich* (1) und dem Endwert *sehr wahrscheinlich* (5) erfasst (Jers, 2012, A 38).

Tabelle 6.20 Beispielhafte Likert-Skala zur Erhebung von Wahrscheinlichkeiten

	sehr unwahr- scheinlich	unwahr- scheinlich	teils, teils	wahr- scheinlich	sehr wahr- scheinlich	ich weiss nicht
Indikator A	☐	☐	☐	☐	☐	☐
Indikator B	☐	☐	☐	☐	☐	☐
etc.	☐	☐	☐	☐	☐	☐

Anmerkungen. Nach Jers (2012).

Die Reihenfolge der Items wurde im Hauptexperiment für alle latenten Variablen randomisiert. Tabelle 6.19 und Tabelle 6.20 geben die Skalen beispielhaft wieder.

6.8.3 Variablen zum Brand Content

Die Brand-Content-Botschaften wurden auf drei Dimensionen beurteilt: Informationsgratifikation bzw. wahrgenommene Information (i. S. v. GO), Unterhaltungsgratifikation und Einstellung gegenüber dem Brand Content. Die *Informationsgratifikation* wurde anhand von drei Items (Indikatoren) operationalisiert (siehe

Tabelle 6.21). Zwei Items wurden von Jahn und Kunz (2012), welche die funktionalen und hedonistischen Inhaltsgratifikationen von Facebook-Brand-Fan-Pages untersuchten, übernommen und ins Deutsche übersetzt. Zwei weitere von ihnen verwendete Items, «functional» und «practical», wurden nicht übernommen. Erstens erwies sich die Übersetzung in ein verständliches und gebräuchliches, äquivalentes deutsches Wort als schwierig. Zweitens wurde angestrebt, die Konstrukte aus Platzgründen im Fragebogen anhand von drei Items zu messen. Auf Basis der erfolgreichen Pretests bei den Kanalgratifikationen wurde als drittes Item «informativ» verwendet.

Tabelle 6.21 Operationalisierung Informationsgratifikation des Brand Contents

	Indikator/Item	Quelle
BC_IG_1	Dieser Post ist hilfreich.	Jahn & Kunz, 2012, S. 360
BC_IG_2	Dieser Post ist nützlich.	Jahn & Kunz, 2012, S. 360
BC_IG_3	Dieser Post ist informativ.	i. A. a. Ducoffe, 1995, S. 5 und Bosshart, 2007

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Die *Unterhaltungsgratifikation* des Brand Contents wurde im Pretest mit drei Items operationalisiert. Die drei Items «unterhaltsam», «spannend» und «macht Spass» wurden Jahn und Kunz (2012) entlehnt und ebenfalls ins Deutsche übersetzt (siehe Tabelle 6.22). Auf das von Jahn und Kunz zusätzlich verwendete Item «pleasant» wurde verzichtet. Die Ergebnisse des Pretests zeigten, dass das Konstrukt zwar akzeptable Werte erzielt, insbesondere aber das Item «spannend» nicht perfekt mit den anderen korrelierte. Für das Hauptexperiment wurde deshalb in Anlehnung an Bosshart und Hellmüller (2009) ein zusätzliches Item formuliert.

Tabelle 6.22 Operationalisierung Unterhaltungsgratifikation des Brand Contents

	Indikator/Item	Quelle
BC_UG_1	Dieser Post ist unterhaltsam.	Jahn & Kunz, 2012, S. 360
BC_UG_2	Dieser Post ist spannend. (entfernt)	Jahn & Kunz, 2012, S. 360
BC_UG_3	Dieser Post macht Spass.	Jahn & Kunz, 2012, S. 360
BC_UG_4	Dieser Post bereitet mir Vergnügen.	i. A. a. Bosshart & Hellmüller, 2009, S. 3

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Die *Einstellung gegenüber dem Brand Content* wurde mit vier Items von Stenger (2012) gemessen (siehe Tabelle 6.23). Stenger übersetzte die Items von MacKenzie et al. (1986). Bei Stenger und MacKenzie et al. wurde die Einstellung gegenüber der Botschaft als semantisches Differential erhoben (*finde ich gut* versus *finde ich nicht gut*). Im Rahmen dieser Arbeit wurden daraus Aussagen in Item-Form gemacht, die ebenfalls auf einer fünfstufigen Likert-Skala eingeschätzt wurden.

Tabelle 6.23 Operationalisierung der Einstellung gegenüber Brand Content

	Indikator/Item	Quelle
BC_ES_1	Ich finde diesen Post gut.	i. A. a. Stenger, 2012, S. 181
BC_ES_2	Ich mag diesen Post.	i. A. a. Stenger, 2012, S. 181
BC_ES_3	Dieser Post ist mir sympathisch.	i. A. a. Stenger, 2012, S. 181
BC_ES_4	Ich finde diesen Post ansprechend.	i. A. a. Stenger, 2012, S. 181

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Die drei latenten Variablen und die damit verbundenen Indikatoren wurden einer konfirmatorischen Faktorenanalyse⁴ unterzogen. In einem ersten Schritt wurden alle aufgeführten Indikatoren und die drei latenten Variablen getestet. In diesem Modell hatte der Indikator B_UG_2 eine vergleichsweise tiefere Faktorladung von $\lambda = 0.74$. Deshalb wurde dieser Indikator in einem zweiten Modell ausgeschlossen. Die Güte des Modells konnte dadurch verbessert werden. Es wird allerdings eine hohe Korrelation zwischen Unterhaltung und Einstellung gegenüber dem Brand Content von 0.83 ausgewiesen. Aus diesem Grund wurde ein drittes Modell mit zwei Faktoren gerechnet, wobei die drei Indikatoren zu Unterhaltung und die vier Indikatoren zur Einstellung gegenüber dem Brand Content zu einer gemeinsamen latenten Variable zusammengefasst wurden. Dieses Zwei-Faktoren-Modell verfügt allerdings über eine schlechtere Modellgüte (siehe elektronisches Zusatzmaterial Tabelle A-5). Auf Basis dieser Berechnungen wurde das *dreifaktorielle Modell angenommen* und der Indikator B_UG_2 nicht berücksichtigt. Dieses Modell bildet die Daten gut ab ($df = 31$, $\chi^2 = 83.158$, $\chi^2/df = 2.68$, CFI = 0.994, TLI = 0.991, RMSEA = 0.035 mit 90 % K.I. zwischen 0.026 und 0.044, SRMR = 0.020).

⁴ Bevor die gesamthafte, alle latenten Variablen umfassende CFA gerechnet wurde, wurden jeweils separate CFA für zusammengehörende latente Variablen gerechnet.

Die Daten in der folgenden Tabelle beziehen sich auf die Gesamt-CFA mit allen latenten Variablen (Hair et al., 2017). Die Faktorladungen sind durchgehend signifikant und hoch (siehe Tabelle 6.24). Die mit Excel auf Basis des Mplus-Outputs berechneten Werte zu Indikatorreliabilität, Faktorreliabilität und zur durchschnittlich extrahierten Varianz pro Faktor bleiben über den geforderten Schwellenwerten. Die latenten Variablen in Bezug auf Brand Content verfügen insgesamt über eine hohe Reliabilität und Validität.

Tabelle 6.24 Teststatistiken zu den Indikatoren und latenten Variablen bezüglich Brand Content

Ind. ⁵		LV	Std.	Unstd.	SE	C. R.	<i>p</i>	IR	C _a	FR	DEV
BC_IG_1	←	BC_IG	0.850	1.078	0.023	46.336	***	0.723	0.918	0.95	0.79
BC_IG_2	←	BC_IG	0.907	1.155	0.020	56.849	***	0.823			
BC_IG_3	←	BC_IG	0.902	1.155	0.021	54.841	***	0.814			
BC_UG_1	←	BC_UG	0.853	0.996	0.022	45.681	***	0.728	0.909	0.91	0.77
BC_UG_3	←	BC_UG	0.891	1.062	0.020	52.040	***	0.794			
BC_UG_4	←	BC_UG	0.884	1.021	0.020	50.891	***	0.781			
BC_ES_1	←	BC_ES	0.891	1.051	0.021	49.298	***	0.794	0.936	0.94	0.79
BC_ES_2	←	BC_ES	0.907	1.068	0.020	53.607	***	0.823			
BC_ES_3	←	BC_ES	0.869	1.007	0.022	45.057	***	0.755			
BC_ES_4	←	BC_ES	0.877	1.053	0.021	50.835	***	0.769			

Anmerkungen. Ind. = Indikator (beobachtete Variable); LV = latente Variable; Std. = standardisierte Faktorladung; Unstd. = unstandardisierte Faktorladung; SE = Standardfehler; C. R. = Critical Ratio; *p* = Wahrscheinlichkeit; IR = Indikatorreliabilität; C_a = Cronbach's Alpha; FR = Faktorreliabilität; DEV = durchschnittlich extrahierte Varianz.

*** signifikant auf Level $p < .001$

In einem unabhängigen Item wurde zudem erhoben, inwiefern die Brand-Content-Botschaft zur Marke passt. Die Aussage dazu lautete: «Dieser Post passt zur Marke [Marke]» (siehe Mendini et al., 2018, S. 198). Die Abkürzung dieses Items lautet B_PASST.

⁵ Der Einfachheit und besseren Übersicht halber werden die Abkürzungen in diesem Kapitel nur in dieser Tabelle wiedergegeben, nicht aber in den darauffolgenden Tabellen, die über dieselben Abkürzungen verfügen.

6.8.4 Variablen zum Kommunikationskanal

Das soziale Medium, das als Kommunikationskanal des Brand Contents dient (Twitter oder Instagram, siehe Abschnitt 6.5.1), wurde auf drei Dimensionen bewertet: Informationserwartungen, Unterhaltungserwartungen (LaRose & Eastin, 2004) und die Einstellung gegenüber dem Kanal. Um die *Informationserwartungen* abzufragen, wurde auf den Vorschlag von Eastin et al. (2014) zurückgegriffen (siehe Tabelle 6.25). Die Items wurden ins Deutsche übersetzt. Die Indikatoren zu Information und Unterhaltung wurden eingeleitet mit: «Wenn Sie Twitter [Instagram] nutzen / nutzen würden, wie wahrscheinlich ist es, dass Folgendes eintritt? Es ist [Auswahl auf Skala von 1 = sehr unwahrscheinlich bis 5 = sehr wahrscheinlich]... dass Sie [Indikator].»

Tabelle 6.25 Operationalisierung der Informationserwartungen gegenüber Kanal

	Indikator/Item	Quelle
K_IE_1	Informiert bleiben.	Eastin, Cicchirillo, Cunningham und Liang, 2014, S. 12
K_IE_2	Informationen finden, die es sonst nirgendwo gibt.	Eastin et al., 2014, S. 12
K_IE_3	Etwas Neues erleben.	Eastin et al., 2014, S. 12

Anmerkungen. Fünfstufige Likert-Skala von *sehr unwahrscheinlich* (1) bis *sehr wahrscheinlich* (5).

Um die *Unterhaltungserwartungen* abzufragen, wurde auf Items von Jers (2012) bzw. Ruehl und Ingenhoff (2017) zurückgegriffen (siehe Tabelle 6.26).

Tabelle 6.26 Operationalisierung der Unterhaltungserwartungen gegenüber Kanal

	Indikator/Item	Quelle
K_UE_1	Spass haben.	Jers, 2012, S. 249; Ruehl & Ingenhoff, 2017, S. 176
K_UE_2	Sich unterhalten fühlen.	Jers, 2012, S. 249; Ruehl & Ingenhoff, 2017, S. 176
K_UE_3	Sich die Zeit vertreiben.	Jers, 2012, S. 249; Ruehl & Ingenhoff, 2017, S. 176

Anmerkungen. Fünfstufige Likert-Skala von *sehr unwahrscheinlich* (1) bis *sehr wahrscheinlich* (5).

Die *Einstellung gegenüber dem Kommunikationskanal* wurde analog zum Globalimage der Marke mit der von Warwitz (2016) vorgeschlagenen Skala erhoben (siehe Tabelle 6.27).

Tabelle 6.27 Operationalisierung der Einstellung gegenüber Kanal

	Indikator/Item	Quelle
K_ES_1	Ich finde [Kanal] gut.	Warwitz, 2016, S. 152
K_ES_2	Ich mag [Kanal].	Warwitz, 2016, S. 152
K_ES_3	[Kanal] ist mir sympathisch.	Warwitz, 2016, S. 152

Anmerkungen. Fünfstufige Likert-Skala von *sehr unwahrscheinlich* (1) bis *sehr wahrscheinlich* (5).

In einem ersten Schritt wurden die drei kanalbezogenen latenten Variablen einer gemeinsamen CFA unterzogen. Das Resultat zeigt einen guten Fit und lässt auf keine nennenswerten Fehlspezifikationen schließen ($df = 24$, $\chi^2 = 83.086$, $\chi^2/df = 3.46$, CFI = 0.991, TLI = 0.986, RMSEA = 0.044 mit 90 % K.I. zwischen 0.034 und 0.054, SRMR = 0.022). Im Folgenden (siehe Tabelle 6.28) werden die Resultate der Gesamt-CFA präsentiert. Auch bei den kanalbezogenen Items bleiben Faktorladungen, Indikatorreliabilität, Faktorreliabilität und DEV über den geforderten Werten.

Tabelle 6.28 Teststatistiken zu den kanalbezogenen Indikatoren und latenten Variablen

Ind.		LV	Std.	Unstd.	SE	C. R.	p	IR	C _a	FR	DEV
K_IE_1	←	K_IE	0.811	1.032	0.028	36.531	***	0.658	0.822	0.82	0.60
K_IE_2	←	K_IE	0.742	0.953	0.028	33.787	***	0.551			
K_IE_3	←	K_IE	0.777	0.975	0.029	34.043	***	0.604			
K_UE_1	←	K_UE	0.867	1.101	0.026	41.948	***	0.752	0.872	0.86	0.67
K_UE_2	←	K_UE	0.883	1.148	0.027	42.579	***	0.694			
K_UE_3	←	K_UE	0.756	1.035	0.033	31.450	***	0.572			
K_ES_1	←	K_ES	0.954	1.227	0.020	61.487	***	0.910	0.965	0.97	0.91
K_ES_2	←	K_ES	0.954	1.277	0.019	67.568	***	0.910			
K_ES_3	←	K_ES	0.941	1.215	0.020	59.651	***	0.885			

Anmerkungen. *** signifikant auf Level $p < .001$

6.8.5 Variablen zum Markenimage

Das Markenimage wurde über ein funktionales Partialimage, ein emotionales Partialimage und über ein nachgelagertes Globalimage erfasst. Die Aufteilung in vorgelagerte Partialimages und ein nachgelagertes Globalimage hat sich im Themengebiet der Markenführung bewährt (Bruhn et al., 2012; Eilers, 2014). Die beiden Dimensionen «funktionales» und «emotionales» Image gelten als die beiden wesentlichsten Image-Dimensionen aus einer Kundensicht (Voss et al., 2003) und finden sich in ähnlicher, wenn auch breiterer Konzeptualisierung auch in der Reputationsforschung (Ingenhoff, 2007, 2018).

Die Operationalisierungen *des funktionalen und des emotionalen Markenimages* sind Horstmann (2017) entnommen (siehe Tabelle 6.29 und Tabelle 6.30). In ähnlicher Form wurden diese bereits in früheren Arbeiten anderer Forschenden der Schule der identitätsorientierten Markenführung verwendet (Becker, 2012; Burmann & Stolle, 2007; Eilers, 2014). Die Operationalisierungen basieren auf diversen Vorarbeiten (D. A. Aaker, 1991; Chaudhuri & Holbrook, 2001; Sweeney & Soutar, 2001). Auch in der Reputationsforschung finden sich ähnliche Formulierungen (Ingenhoff & Buhmann, 2016).

Tabelle 6.29 Operationalisierung des funktionalen Markenimages

	Indikator/Item	Quelle
M_FI_1	Die Produkte der Marke [Marke] haben eine gute Qualität.	Horstmann, 2016, S. 151
M_FI_2	Die Produkte der Marke [Marke] haben exzellente Eigenschaften.	Horstmann, 2016, S. 151
M_FI_3	Die Produkte der Marke [Marke] sind zuverlässig.	Horstmann, 2016, S. 151
M_FI_4	Die Produkte der Marke [Marke] haben ein gutes Preis-Leistungs-Verhältnis.	Horstmann, 2016, S. 151

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Die Formulierung der Items des *emotionalen Markenimages* wurde im Vergleich zu Horstmann (2017) um den Konjunktiv ergänzt (Diehl, 2009, S. 242), weil nicht davon ausgegangen werden kann, dass alle Befragten die Marke tatsächlich verwenden.

Tabelle 6.30 Operationalisierung des emotionalen Markenimages

	Indikator/Item	Quelle
M_EI_1	Die Marke [Marke] macht bzw. würde mich glücklich machen, wenn ich sie verwende.	Horstmann, 2016, S. 151
M_EI_2	Die Marke [Marke] gibt bzw. würde mir ein gutes Gefühl geben, wenn ich sie verwende.	Horstmann, 2016, S. 151
M_EI_3	Die Marke [Marke] macht bzw. würde mir Spaß machen, wenn ich sie verwende.	Horstmann, 2016, S. 151

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Die drei Items zum Messen des *Globalimages* stammen von Warwitz (2016) und gehen auf A. A. Mitchell (1986, S. 15) zurück (siehe Tabelle 6.31). Horstmann (2017, S. 149) nutzte ähnliche, in der Formulierung aber etwas längere Items.

Tabelle 6.31 Operationalisierung des Globalimages der Marke

	Indikator/Item	Quelle
M_GI_1	Ich finde [Marke] gut.	Warwitz, 2016, S. 152
M_GI_2	Ich mag [Marke].	Warwitz, 2016, S. 152
M_GI_3	[Marke] ist mir sympathisch.	Warwitz, 2016, S. 152

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

In einem ersten Schritt der Modellprüfung wurden alle Indikatoren und die damit verbundenen latenten Variablen des Markenimages einer konfirmatorischen Faktorenanalyse unterzogen. Es zeigte sich, dass insbesondere die Variable M_FI_4 wenig mit den anderen Items des funktionalen Images korrespondierte und eine entsprechend tiefe Faktorladung von $\lambda = 0.59$ hat. Deshalb wurde ein zweites Modell ohne diesen Indikator gerechnet. Weil sich die Modellgüte dadurch nicht signifikant steigerte, blieb die Variable erhalten. Weil die beiden Partialimages eine hohe Korrelation von 0.86 (Funktionales Image – Globalimage) bzw. 0.88 (Emotionales Image – Globalimage) mit dem Globalimage ausweisen, wurden zwei Zwei-Faktoren-Modelle gerechnet, wobei jeweils eines der Partialimages mit dem Globalimage zusammengeführt wurde, und ein Ein-Faktoren-Modell getestet. Es zeigte sich, dass weder die Reduktion von Faktoren

noch das Entfernen zur Verbesserung des Modells beitrugen (siehe elektronisches Zusatzmaterial Tabelle A-6). Deshalb wurde das Ursprungsmodell mit drei latenten Variablen angenommen, das insgesamt über einen guten Fit verfügt ($df = 32$, $\chi^2 = 316.037$, $\chi^2/df = 9.88$, CFI = 0.958, TLI = 0.941, RMSEA = 0.081 mit 90 % K.I. zwischen 0.073 und 0.089, SRMR = 0.041).

Die Angaben in der untenstehenden Tabelle 6.32 beziehen sich auf die Gesamt-CFA. Alle Indikatoren bleiben über den geforderten Schwellenwerten, ausser die Indikatorreliabilität des Faktors M_FI_4. Diese liegt unter dem Grenzwert von 0.4. Weil allerdings, wie oben beschrieben, das Modell nicht signifikant verbessert werden konnte durch Entfernen des Indikators, blieb er erhalten.

Tabelle 6.32 Teststatistiken zu den markenbezogenen Indikatoren und latenten Variablen

Ind.		LV	Std.	Unstd.	SE	C. R.	<i>p</i>	IR	C _a	FR	DEV
M_FI_1	←	M_FI	0.835	0.699	0.029	23.799	***	0.697	0.830	0.86	0.58
M_FI_2	←	M_FI	0.825	0.773	0.026	29.615	***	0.681			
M_FI_3	←	M_FI	0.831	0.709	0.028	25.200	***	0.691			
M_FI_4	←	M_FI	0.589	0.593	0.031	19.201	***	0.347			
M_EI_1	←	M_EI	0.861	1.034	0.024	43.854	***	0.741	0.891	0.89	0.73
M_EI_2	←	M_EI	0.854	0.972	0.026	37.397	***	0.729			
M_EI_3	←	M_EI	0.849	1.005	0.024	41.659	***	0.721			
M_GI_1	←	M_GI	0.923	0.966	0.025	38.808	***	0.852	0.930	0.93	0.82
M_GI_2	←	M_GI	0.913	1.017	0.024	42.927	***	0.834			
M_GI_3	←	M_GI	0.881	0.910	0.026	35.056	***	0.776			

Anmerkungen. *** signifikant auf Level $p < .001$

6.8.6 Variablen zum markenbezogenen Verhalten

Um das *markenbezogene Verhalten* zu untersuchen, wurden drei Items von Stenger (2012) übernommen. Stenger befasst sich mit dem Einfluss viraler Markenkommunikation auf die Markenwahrnehmung. Die drei Items wurden um ein viertes ergänzt, das Stenger mit «Weiterleitungsabsicht» umschrieb (siehe Tabelle 6.33).

Tabelle 6.33 Operationalisierung des markenbezogenen Verhaltens

	Indikator/Item	Quelle
M_MV_1	Wie wahrscheinlich ist es, dass Sie Produkte der Marke [Marke] kaufen?	Stenger, 2012, S. 182
M_MV_2	Wie wahrscheinlich ist es, dass Sie die Webseite der Marke [Marke] besuchen?	Stenger, 2012, S. 182
M_MV_3	Wie wahrscheinlich ist es, dass Sie die Marke [Marke] weiterempfehlen?	Stenger, 2012, S. 182
M_MV_4	Wie wahrscheinlich ist es, dass Sie den gesehenen Post online teilen / weiterleiten würden?	i. A. a. Stenger, 2012, S. 181

Anmerkungen. Fünfstufige Likert-Skala von *sehr unwahrscheinlich* (1) bis *sehr wahrscheinlich* (5).

Für das markenbezogene Verhalten wurde ebenfalls eine konfirmatorische Faktorenanalyse gerechnet. Vier Indikatoren bei einer latenten Variable sind die minimalen Voraussetzungen, um ein sinnvolles CFA-Ergebnis zu erhalten (Hair et al., 2017, S. 168). Der erhaltene Fit des Modells ist allerdings ungenügend (siehe elektronisches Zusatzmaterial Tabelle A-7). Die kleinste Faktorladung wurde für das Item M_MV_4 ausgemacht ($\lambda = 0.44$). Weil die Berechnung mit nur drei Indikatoren nicht möglich ist, wurde eine umfassende konfirmatorische Faktorenanalyse zusammen mit dem Markenimage gemacht. Dabei wurde einerseits das Markenverhalten mit vier Indikatoren und mit drei Indikatoren (ohne M_MV_4) getestet. Auf Basis der Modifikationsindizes (Muthén & Muthén, 2017, S. 804) wurde zudem ersichtlich, dass sowohl M_MV_1 und M_MV_3 als auch M_MV_2 und M_MV_4 jeweils stark korrelieren. Deshalb wurde das Markenverhalten in einem weiteren Modell in zwei Faktoren aufgeteilt und ebenfalls zusammen mit dem Markenimage getestet. Die gemeinsame Prüfung zusammen mit dem Markenimage ist aus zwei Gründen plausibel. Einerseits beziehen sich sowohl das Markenimage als auch das markenbezogene Verhalten auf jeweils dieselben Marken. Andererseits kann das markenbezogene Verhalten im Sinne der Drei-Komponenten-Theorie als der konativ orientierte Bestandteil des Markenimages betrachtet werden (siehe Abschnitt 2.3.3).

Es zeigte sich, dass eine CFA zusammen mit dem Markenimage und zwei markenbezogenen Faktoren eine akzeptable Modellgüte aufweist ($df = 67$, $\chi^2 = 511.064$, $\chi^2/df = 7.63$, CFI = 0.958, TLI = 0.943, RMSEA = 0.070 mit 90 % K.I. zwischen 0.064 und 0.075, SRMR = 0.043). Deshalb wurden die vier markenbezogenen Indikatoren behalten, jedoch in zwei latente Variablen aufgesplittet. Die Werte zu Faktorladungen, Indikatorreliabilität, Faktorreliabilität und DEV erreichen allesamt zufriedenstellende Werte (siehe Tabelle 6.34).

Tabelle 6.34 Teststatistiken zu den Indikatoren und latenten Variablen bezüglich Markenverhalten

Ind.		LV	Std.	Unstd.	SE	C. R.	<i>p</i>	IR	C _a	FR	DEV
M_MV_1	←	MV_I	0.851	1.105	0.026	41.693	***	0.724	0.851	0.85	0.74
M_MV_3	←	MV_I	0.870	1.113	0.025	44.001	***	0.757			
M_MV_2	←	MV_II	0.775	0.992	0.030	32.649	***	0.601	0.720	0.72	0.56
M_MV_4	←	MV_II	0.726	0.856	0.034	25.095	***	0.527			

Anmerkungen. *** signifikant auf Level $p < .001$

6.8.7 Variablen zum Produktinvolvement

Das *Produktinvolvement* wurde mit drei Items von Meeder (2007) erhoben (siehe Tabelle 6.35), die sich ihrerseits an Ratchford (1987) orientiert. Die Formulierungen wurden im Vergleich zu Meeder um den Konjunktiv ergänzt, weil davon ausgegangen werden muss, dass die befragten Personen nicht direkt in eine Kaufentscheidung bezüglich der im Experiment verwendeten Produkte involviert sein müssen.

Tabelle 6.35 Operationalisierung des Produktinvolvements

	Indikator/Item	Quelle
PI_1	Die Wahl dieser Produkte ist (wäre) eine wichtige Entscheidung für mich.	Meeder, 2007, S. 191
PI_2	Ich beschäftige mich (bzw. würde mich) relativ ausführlich mit der Kaufentscheidung bzgl. dieser Produkte (beschäftigen).	Meeder, 2007, S. 191
PI_3	Wenn ich in dieser Produktkategorie einen Fehlkauf tätige (tätigen würde), wäre das ein verhältnismässig grosser finanzieller Verlust für mich.	Meeder, 2007, S. 191

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Das Produktinvolvement verfügt über drei Indikatoren und ist deshalb ein «gerade identifiziertes» Modell mit 0 Freiheitsgraden. Solche Modelle verfügen im mathematischen Sinne über einen perfekten Fit, ermöglichen jedoch keine sinnvolle Durchführung einer CFA (Hair et al., 2017, S. 168). Deshalb konnte die Modellgüte des Produktinvolvements nicht selbstständig geprüft werden. In analoger Vorgehensweise wie beim markenbezogenen Verhalten wurde deshalb auch

das Produktinvolvement zusammen mit den drei latenten Variablen des Markenimages in einer CFA überprüft. Weil sich auch in diesem Fall alle latenten Variablen auf jeweils dieselbe Marke beziehen, scheint dieses Vorgehen ebenfalls vertretbar. Das gemeinsame Modell (Markenimage & Produktinvolvement) verfügt über eine akzeptable Güte ($df = 59$, $\chi^2 = 468.850$, $\chi^2/df = 7.95$, CFI = 0.952, TLI = 0.936, RMSEA = 0.071 mit 90 % K.I. zwischen 0.065 und 0.077, SRMR = 0.046). Entsprechend wurde das Produktinvolvement mit den drei vorgesehenen Indikatoren operationalisiert. Die Angaben der CFA zeigen allerdings, dass die Indikatorreliabilität des Faktors PI_3 unter dem Schwellenwert von 0.4 liegt (siehe Tabelle 6.36). Das Item wurde beibehalten, weil die Faktorladung genug hoch und signifikant ist. Zudem wird bei Gruppenvergleichen im Rahmen von Strukturgleichungsmodellen empfohlen, latente Variablen mit mindestens drei Indikatoren zu messen (Steinmetz et al., 2009, S. 614).

Tabelle 6.36 Teststatistiken zu Indikatoren und latenter Variable bezüglich Produktinvolvement

Ind.		LV	Std.	Unstd.	SE	C. R.	<i>p</i>	IR	C _a	CR	DEV
PI_1	←	PI	0.748	0.958	0.034	28.442	***	0.560	0.753	0.77	0.53
PI_2	←	PI	0.835	1.096	0.034	31.814	***	0.697			
PI_3	←	PI	0.582	0.832	0.041	20.125	***	0.339			

Anmerkungen. *** signifikant auf Level $p < .001$

6.8.8 Markenpositionierung und weitere markenbezogene Variablen

Der Manipulationscheck für die *Markenpositionierung* erfolgte über ein Item (Kempf, 1999, S. 43; K. Klein & Melnyk, 2016, S. 135; Lu, Liu & Fang, 2016, S. 334). Die Personen mussten auf einer fünfstufigen Likert-Skala angeben, ob es aus ihrer Sicht *eine komplett rationale/kopforientierte Entscheidung* (1) oder *eine komplett emotionale/gefühlsorientierte Entscheidung* (5) ist, die Produkte der vier Marken zu kaufen.

Neben diesem Manipulationscheck wurden drei weitere markenbezogene Fragen gestellt, die einerseits den Zugang in die Umfrage erleichtern sollten und zusätzlich als Kontrollvariablen dienen. Zu Beginn des Fragebogens wurde erhoben, ob der Teilnehmer oder die Teilnehmerin bereits Produkte der verwendeten Marken gekauft hat. Mögliche Antwortoptionen waren: (1) habe ich nicht gekauft und Kauf auch nicht in Erwägung gezogen, (2) habe ich nicht gekauft, aber einen

Kauf in Erwägung gezogen, (3) habe ich schon gekauft, werde es aber nicht mehr tun, (4) habe ich schon gekauft und werde es wieder tun und (5) Ich weiss nicht / kenne die Marke nicht. Ebenfalls wurde die Ex-ante-Einstellung gegenüber den vier Marken erhoben: [Marke] ist eine gute Marke. Die Aussage musste auch in einer fünfstufigen Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5) eingeordnet werden. Letztlich wurde gefragt, ob den Marken auf den sozialen Medien gefolgt wird. Ja, Nein und Ich weiss nicht dienten als Antwortoptionen.

6.8.9 Mediennutzung und Affinität mit sozialen Medien

Die Mediennutzung in folgenden Kategorien wurde zur Beschreibung der Stichprobe erhoben (siehe Abschnitt 6.6.1): Internet allgemein, soziale Medien allgemein, Facebook, Twitter, Instagram, YouTube. Als Antwortoptionen dienten nie, einmal pro Monat oder seltener, einmal pro Woche oder mehrmals pro Monat, mehrmals pro Woche, täglich oder mehrmals täglich. Zudem wurde Ich kenne den Kanal nicht als Antwortoption zur Verfügung gestellt.

Die *Social-Media-Affinität* (SMA) diene als Kontrollvariable. Um die Affinität mit sozialen Medien zu erheben, wurden drei Items von Wozniak, Schaffner, Stanoevska-Slabeva und Lenz-Kesekamp (2017) übernommen (siehe Tabelle 6.37). Die in Bezug auf Smartphones entwickelten Items wurden auf das Thema «soziale Medien» umformuliert und vom Englischen ins Deutsche übersetzt. Beim zweiten Item wurde eine sprachliche Präzisierung vorgenommen. Anstelle der ursprünglichen Formulierung «Ich benutze Smartphones 24/7» wurde «Ich benutze soziale Medien rund um die Uhr» verwendet. Der Begriff «soziale Medien» wurde im Fragebogen anhand von Beispielen (z. B. Facebook, Instagram, Twitter, YouTube, Snapchat) definiert.

Tabelle 6.37 Operationalisierung der Social-Media-Affinität

	Indikator/Item	Quelle
SM_1	Ich kann mir ein Leben ohne soziale Medien nicht vorstellen.	Wozniak et al., 2017, S. 325
SM_2	Ich benutze soziale Medien rund um die Uhr.	Wozniak et al., 2017, S. 325
SM_3	Ich bin «süchtig» nach sozialen Medien.	Wozniak et al., 2017, S. 325

Anmerkungen. Fünfstufige Likert-Skala von *trifft überhaupt nicht zu* (1) bis *trifft voll und ganz zu* (5).

Die latente Variable Social-Media-Affinität wurde anhand dreier Indikatoren gemessen. Damit ist das Modell analog dem Produktinvolvement gerade identifiziert. Zur Überprüfung wurde die latente Variable zusammen mit den drei kanalbezogenen latenten Variablen einer gemeinsamen CFA unterzogen. Weil dieses Modell einen guten Fit aufweist ($df = 48$, $\chi^2 = 139.865$, $\chi^2/df = 2.91$, CFI = 0.991, TLI = 0.989, RMSEA = 0.037 mit 90 % K.I. zwischen 0.030 und 0.045, SRMR = 0.021), wurde es gemäss obigem Vorschlag operationalisiert. Die folgenden Angaben beziehen sich auf die Gesamt-CFA mit allen latenten Variablen. Alle Werte befinden sich auf einem akzeptablen Niveau (siehe Tabelle 6.38).

Tabelle 6.38 Teststatistiken zu Indikatoren und latenter Variable bezüglich Social-Media-Affinität

Ind.		LV	Std.	Unstd.	SE	C. R.	P	IR	C _a	CR	DEV
SM_1	←	SM	0.720	0.947	0.031	30.941	***	0.518	0.805	0.81	0.59
SM_2	←	SM	0.767	1.012	0.030	33.652	***	0.588			
SM_3	←	SM	0.811	0.926	0.029	32.287	***	0.658			

Anmerkungen. *** signifikant auf Level $p < .001$

6.8.10 Weitere demografische Variablen

Die Teilnehmenden wurden gebeten, ihr *Alter*, ihr *Geschlecht* und ihre höchste abgeschlossene *Ausbildung* anzugeben. Diese drei Fragen wurden zu Beginn des Fragebogens gestellt, weil diesbezüglich Quoten erfüllt werden sollten.

Als höchste abgeschlossene Bildung wurden neun Antwortoptionen vorgegeben, die sich in die beiden Kategorien «mit tertiärem Bildungsabschluss» und «ohne tertiären Bildungsabschluss» unterteilen liessen (siehe Tabelle 6.39). Antwortoptionen und Zuteilung erfolgten nach den Vorgaben der Schweizer Berufsberatung (2019).

Tabelle 6.39 Kategorien des Bildungsabschlusses

Abschluss	Bildung
Obligatorische Schule	nicht-tertiär
Berufliche Grundbildung	nicht-tertiär
Matura / Gymnasium (gymnasiale Matura, Fachmatura, Berufsmatura)	nicht-tertiär
Höhere Berufsbildung	Tertiär
Bachelor (Uni, FH)	Tertiär
Master / Lizenziat (Uni, FH)	Tertiär
Doktorat	Tertiär
Andere Ausbildung	fehlender Wert
Keine Angabe	fehlender Wert

Anmerkungen. Zuordnung der Abschlüsse auf Kategorie Bildung gemäss Schweizer Berufsberatung.

Nach den ersten demografischen Merkmalen wurden die Teilnehmenden dazu befragt, ob sie die Marken bereits gekauft haben (siehe Tabelle 6.40). Alle Teilnehmenden haben diese Frage für alle vier Marken beantwortet. Die Kontrollvariable «Wiederkauf» wurde derart umkodiert, dass sich der Wert auf die jeweils gezeigte Marke bezog. Weil jedoch der Anteil derjenigen Testpersonen, welche die Marke bereits gekauft haben und die Marke wieder kaufen würden, bei Elmex (78.6 %) und Rivella (83.5 %) deutlich höher ist als bei Freitag (12 %) und bei V-Zug (35.7 %), ist die Variable «Wiederkauf» Richtung Elmex und Rivella verzerrt. Das heisst, dass von den 740 Personen, welche die ihnen gezeigte Marke wiederkaufen würden (vs. 551 Personen, welche das nicht tun würden bzw. die Marke gar noch nicht gekauft haben), 77.4 % entweder eine Botschaft von Elmex oder Rivella gesehen haben.

Tabelle 6.40 Kaufverhalten und Wiederkauf der Marken

Antwortkategorie	Wiederkauf
Habe ich nicht gekauft und Kauf auch nicht in Erwägung gezogen.	Hat Marke noch nicht gekauft oder würde sie nicht mehr kaufen.
Habe ich nicht gekauft, aber einen Kauf in Erwägung gezogen.	(wie oben)
Habe ich schon gekauft, werde es aber nicht mehr tun.	(wie oben)
Habe ich schon gekauft und werde es wieder tun.	Hat Marke schon gekauft und wird es wieder tun.

Anmerkungen. Eigene Zuordnung der Antwortkategorien auf die Variable Wiederkauf.

Am Ende des Fragebogens wurden das monatliche Haushaltseinkommen (brutto), die Wohnsituation und die Erwerbstätigkeit erfragt. Beim Haushaltseinkommen standen sechs Optionen zur Auswahl, die in drei Kategorien eingeteilt wurden (siehe Tabelle 6.41):

Tabelle 6.41 Kategorien des Haushaltseinkommens

Haushaltseinkommen (brutto)	Einkommen
unter CHF 4'000	tiefes Einkommen
CHF 4'000 bis 5'999	tiefes Einkommen
CHF 6'000 bis 7'999	mittleres Einkommen
CHF 8'000 bis 9'999	hohes Einkommen
über CHF 10'000	hohes Einkommen
keine Angabe	(fehlender Wert)

Anmerkungen. Eigene Zuordnung der Antwortkategorien auf die Variable Einkommen

Bei der Wohnsituation konnte zwischen «Stadt», «Agglomeration» und «Land» sowie «keine Angabe» gewählt werden. Die Optionen bei der Erwerbstätigkeit unterteilten sich in «Vollzeit, 90 % bis 100 %», «Teilzeit, 10 bis 89 %», «keine Erwerbstätigkeit» und «keine Angabe». Wohnsituation und Erwerbstätigkeit dienten einzig der Beschreibung der Stichprobe.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.