



Moderation in Marketing Applications of Partial Least Squares Structural Equation Modeling (PLS-SEM)

Moderation in Marketinganwendungen von Partial Least Squares Strukturgleichungsmodellierungen (PLS-SEM)

Jonas Spengler

Otto-von-Guericke-Universität Magdeburg

Abstract

This thesis investigates how current marketing research deals with simple moderation analysis in marketing applications of partial least squares structural equation modeling (PLS-SEM). In this context, different approaches to generate the interaction term as well as options of data treatment before analysis, among others, pose challenges to researchers. The goal of this thesis is therefore to highlight the most common sources of error through a review of 33 relevant empirical articles from the PLS-SEM context and to derive recommendations for research. Using a framework that shows how to preferably generate the interaction term, it is found that about 70% of the analyzed studies use a suboptimal approach. Furthermore, the product-indicator approach is the most frequently used but only in exceptional cases the optimal approach. Additionally, both the approach used as well as the effect sizes of moderation effects are generally rarely specified and finally uncertainties in the interpretation of effects in moderation analyses can be observed. This confirms the urgency of the need for educating researchers in terms of moderation analysis in the context of PLS-SEM to minimize common mistakes in the future that distort results of moderation analyses or make their meaning incorrect.

Zusammenfassung

Die vorliegende Arbeit untersucht den Umgang der aktuellen Marketing-Forschung mit einfachen Moderationsanalysen in Marketinganwendungen von Partial Least Squares Strukturgleichungsmodellierungen (PLS-SEM). Dabei stellen u.a. verschiedene Ansätze zur Generierung des Interaktionsterms und die Datenbehandlung vor der Analyse Forscher vor Herausforderungen. Das Ziel dieser Arbeit ist es deshalb, durch ein Review 33 einschlägiger empirischer Artikel aus dem PLS-SEM-Kontext die häufigsten Fehlerquellen zu beleuchten und Empfehlungen für die Forschung abzuleiten. Mithilfe eines Frameworks zur Generierung des Interaktionsterms konnte gezeigt werden, dass etwa 70% der analysierten Untersuchungen einen suboptimalen Ansatz verwenden. Weiterhin ist der Produkt-Indikator-Ansatz der am häufigsten genutzte, aber nur in Ausnahmefällen optimale Ansatz. Hinzu kommen die allgemein seltene Angabe des verwendeten Ansatzes, der Effektgrößen von Moderationseffekten und Unsicherheiten bei der Interpretation der Effekte bei Moderationsanalysen. Dies bestätigt die Dringlichkeit des Bedarfs nach einer Aufklärung von Forschern, damit häufige Fehler, welche die Ergebnisse von Moderationsanalysen verzerren oder inhaltlich falsch werden lassen, in Zukunft minimiert werden.

Keywords: Moderationsanalyse; Interaktionseffekt; Moderator; PLS-SEM; Partial Least Squares Strukturgleichungsmodellierung.

1. Einleitung

Das Testen komplexer Beziehungen zwischen einer Vielzahl von Variablen, die direkt (manifeste Variablen) oder nur unter Zuhilfenahme mehrerer messbarer Variablen be-

obachtbar sind (latente Variablen), erfolgt zunehmend in verschiedensten wissenschaftlichen Disziplinen mithilfe von Partial Least Squares Strukturgleichungsmodellierung (PLS-SEM) (Sarstedt, Ringle et al., 2020). Dabei nimmt besagte

Komplexität der Beziehungen stetig zu, was u.a. darin resultiert, dass sich Forscher neben oftmals bereits untersuchten direkten Effekten zwischen Variablen vermehrt für Moderationseffekte (ME) interessieren (Henseler & Chin, 2010; Henseler, Fassott, Dijkstra & Wilson, 2012).

Moderation wird als der Einfluss auf die Stärke und/oder Richtung der Beziehung zweier Variablen beschrieben, der durch die Manipulation einer dritten Variable, der sog. Moderatorvariable (MV), verursacht wird (Baron & Kenny, 1986, S. 1174). Hierbei kann bspw. untersucht werden, unter welchen Umständen der positive Zusammenhang zwischen Kundenzufriedenheit und -loyalität wesentlich stärker oder schwächer ausgeprägt ist (Henseler & Fassott, 2010). Der ME wird dabei im aktuellen Kontext stets durch die Generierung eines sog. Interaktionsterms modelliert. Diese Modellierung von ME in Strukturmodellen ist jedoch nicht trivial und kommt mit einigen Herausforderungen einher, wie Memon et al. (2019) aufzeigen. Die Autoren weisen neben Schwierigkeiten bei der Hypothesisierung und Interpretation von ME insb. auf verschiedene Ansätze hin, zwischen denen Forscher wählen können, um den Interaktionsterm bei MA in PLS-SEM zu generieren.

So haben Forscher in der Regel die Wahl zwischen dem Produkt-Indikator-Ansatz, Zwei-Stufen-Ansatz und Orthogonalisierungsansatz (Becker, Ringle & Sarstedt, 2018). Eine Herausforderung besteht dabei darin, den passenden Ansatz zu wählen, da diese unterschiedliche Leistungen in Abhängigkeit von den Zielen von Forschern hervorbringen (Henseler & Chin, 2010). Vielmehr sorgen aber unterschiedliche Messmodelle, die latente Variablen in PLS-SEM messen können, für eine schwerwiegende potenzielle Fehlerquelle. Lediglich der Zwei-Stufen-Ansatz ist bspw. dazu in der Lage, eine inhaltlich korrekte MA durchführen zu können, wenn neben reflektiv gemessenen auch formativ gemessene latente Variablen innerhalb der MA vorkommen. Darüber hinaus kann die Behandlung der Daten vor der Analyse einen Einfluss auf die Schätzung des ME haben (Becker et al., 2018). Die Daten können hierbei standardisiert, mittelwertzentriert oder unstandardisiert in die Analyse eingehen. Diese Vielzahl von Entscheidungsmöglichkeiten kann für Verwirrung bei Forschern sorgen, wenn MA in PLS-SEM durchgeführt werden sollen (Hair, Sarstedt & Ringle, 2019).

Henseler und Fassott (2010) haben deshalb in ihrem Artikel einen Leitfaden gegeben, wie MA in PLS-SEM durchzuführen sind. Dieser beinhaltet jedoch nicht die Erkenntnisse der Simulationsstudie von Henseler und Chin (2010) zu den unterschiedlichen Leistungen der Ansätze und die Erkenntnisse von Becker et al. (2018) bezüglich der empfohlenen Datenbehandlung. Hierbei fehlt ebenfalls eine Berücksichtigung des Orthogonalisierungsansatzes. Aus diesem Grund soll in der vorliegenden Arbeit zunächst ein Framework entwickelt werden, das auf der aktuellen Forschung zur Durchführung von MA in PLS-SEM aufbaut. Wie der Leitfaden von Henseler und Fassott (2010) soll es dazu dienen, Forschern eine Vorgabe zu geben, MA in PLS-SEM bestmöglich umzusetzen.

Weiterhin kann u.a. eine bis heute andauernde Kontroverse bezüglich der Interpretation verschiedener Effekte, die

bei MA auftreten, beobachtet werden. So vertreten bspw. Hair, Hult, Ringle, Sarstedt und Thiele (2017, S. 256-258) den Standpunkt, Forscher sollten ein Modell ohne den ME schätzen, um die Beziehung zweier Variablen unabhängig vom Moderator testen zu können. Hiergegen argumentieren Li, Sharp, Bergh und Vandenberg (2019), dass eine Interpretation dieser Beziehung nur innerhalb eines Modells möglich ist, das den ME ebenfalls beinhaltet, da sie nach den Autoren immer vom ME abhängig sein muss, insofern dieser signifikant ist.

Die potenzielle Verwirrung sowie die angesprochene Kontroverse von Forschern im Umgang mit MA in PLS-SEM führt zu der Frage, inwieweit den Empfehlungen der Literatur zur Analyse von ME mithilfe von Interaktionstermen in PLS-SEM gefolgt wird und wo häufige Fehler und Probleme auftreten. Zur Beantwortung dieser Frage werden nach der Aufstellung des Frameworks einschlägige Artikel durch eine systematische Literaturrecherche identifiziert und anschließend mithilfe eines Reviews untersucht. Dabei besteht das Ziel darin, aufzuzeigen, wie in aktuellen Forschungsartikeln mit den genannten Herausforderungen umgegangen wird und welche Empfehlungen dabei häufig vernachlässigt oder befolgt werden. Hieraus sollen schließlich Empfehlungen für die Forschung abgeleitet werden, um der Unsicherheit entgegenzuwirken und den angemessenen Umgang mit MA in PLS-SEM in der Zukunft zu motivieren.

Im Folgenden wird hierfür im zweiten Abschnitt der theoretische Hintergrund von PLS-SEM und MA dargelegt. Weiterhin werden bekannte Herausforderungen der aktuellen Forschung aufgezeigt und das Framework zur Vorgehensweise bei MA in PLS-SEM aufgestellt. Der dritte Abschnitt stellt die wissenschaftliche Methode hinter dem anschließenden Review dar. Darauf aufbauend umfasst der vierte Abschnitt das Review der identifizierten Artikel zur Bestimmung von Herausforderungen im Umgang mit MA in PLS-SEM. Abschließend werden Implikationen für die Forschung abgeleitet und Limitationen der vorliegenden Arbeit diskutiert, worauf die abschließend dargestellten Gelegenheiten für zukünftige Forschung basieren.

2. Moderationsanalysen in PLS-SEM

In dem nachfolgenden Abschnitt wird der theoretische Hintergrund der vorliegenden Arbeit erläutert, auf welchem das anschließende Review von Moderationen in Marketinganwendungen von PLS-SEM basiert. Hierbei wird zunächst die Funktionsweise und das Anwendungsgebiet von PLS-SEM dargelegt, wonach eine allgemeine Erläuterung zu Moderationen in Marketinganwendungen sowie eine Aufschlüsselung verschiedener Formen von MA folgt. Weiterhin werden die Vorteile von PLS-SEM bei der Analyse von ME aufgezeigt sowie näher auf die Berechnung der ME in PLS-SEM mithilfe der Generierung von Interaktionstermen eingegangen. Nachdem schließlich die Interpretation von ME und häufig auftretende Fehler behandelt werden, wird aus den aktuellen Erkenntnissen zu Analysen von ME in PLS-SEM mit Interaktionstermen ein Framework aufgestellt, welches als Anleitung

für die Durchführung solcher Analysen in zukünftiger Forschung dienen soll und das zum Teil die Grundlage für das Review im vierten Abschnitt dieser Arbeit bildet.

2.1. Partial Least Squares Strukturgleichungsmodellierung (PLS-SEM)

Liegt der Fokus von Forschern auf der Betrachtung komplexer Beziehungen zwischen zwei oder mehr Variablen, findet häufig eine Form der multivariaten Analyse Anwendung, die simultan die Beziehung mehrerer Variablen zueinander analysiert (Hair et al., 2017, S. 2). Basierend auf einem theoretischen Modell bilden die Variablen oftmals ein Netzwerk kausaler Beziehungen, welches durch statistische Methoden geschätzt wird, die unter Strukturgleichungsmodellierung (SEM) zusammengefasst werden (Esposito Vinzi, Trinchera & Amato, 2010). Insbesondere dient die auf der multivariaten Analyse basierende SEM dabei dazu, auch Netzwerke bestehend aus nicht direkt beobachtbaren Variablen, sog. latenten Variablen bzw. Konstrukten, welche indirekt durch mehrere beobachtbare Indikatorvariablen bzw. manifeste Variablen messbar sind, zu schätzen (Hair et al., 2017, S. 2-4). Konstrukte werden dazu genutzt, abstrakte Ideen und Konzepte zu messen, um menschliches Verhalten besser verstehen und voraussagen zu können (Sarstedt, Hair Jr, Nitzl, Ringle & Howard, 2020). Dabei stellen bspw. die Zufriedenheit mit einem Produkt oder Loyalität zu einem Unternehmen Konstrukte dar. SEM wird hierbei dazu verwendet, um die komplexen wechselseitigen Beziehungen zwischen latenten und manifesten Variablen simultan zu erfassen und dabei Messfehler zu berücksichtigen, die bei der Messung von Konstrukten auftreten (Sarstedt, Hair, Ringle, Thiele & Gudergan, 2016).

Netzwerke aus latenten Variablen werden nach Hair, Ringle und Sarstedt (2011) in der Regel mithilfe sog. Pfadmodelle visualisiert. Abbildung 1 zeigt eine beispielhafte Darstellung eines Pfadmodells mit insg. drei Konstrukten. Diese sind nach den Autoren von Y_1 bis Y_3 benannt, werden über einseitig gerichtete Pfeile (Pfade) miteinander verbunden und besitzen zum Teil eine verschiedene Anzahl von Indikatorvariablen, die sich von x_1 bis x_7 erstrecken. Weiterhin gehen die Autoren darauf ein, dass Y_1 und Y_2 bei der Betrachtung des im aktuellen Kontext sog. inneren Modells (Strukturmodells) exogene latente Variablen darstellen, die andere Konstrukte im Modell erklären, was bedeutet, dass lediglich Pfeile von diesen Konstrukten weggehen. Y_3 hingegen verkörpert nach den Autoren ein endogenes Konstrukt, das im Modell erklärt wird, d.h. es gehen sowohl Pfeile zu diesem Konstrukt hin, sie können jedoch darüber hinaus auch davon ausgehen. Auf Basis einer hinreichenden theoretischen Herleitung der Beziehung zwischen den Konstrukten können die Pfade auch als kausale Beziehungen zwischen den Konstrukten interpretiert werden (Hair et al., 2017, S. 11).

Das sog. äußere Modell (Messmodell) besteht in Abbildung 1 sowohl aus reflektiv als auch aus formativ gemessenen Konstrukten. Nach Hair et al. (2017, S. 13) werden Y_1 und Y_2 in der gezeigten Abbildung 1 reflektiv gemessen, was daran zu erkennen ist, dass die Pfeile ausschließlich von den

Konstrukten zu den zugehörigen Indikatoren (von Y_1 zu x_1 und x_2 ; von Y_2 zu x_3 , x_4 und x_5) verlaufen. Nach den Autoren wird Y_3 demnach formativ von den Indikatorvariablen x_6 und x_7 gemessen, da sich sämtliche Pfeile von den Indikatoren zu dem Konstrukt erstrecken.

Werden latente Variablen reflektiv gemessen, sorgen Änderungen im Wert des Konstruktes zu Veränderungen der respektiven Indikatorenwerte und vice versa bei formativ gemessenen Konstrukten (Hair et al., 2011). Reflektive Messungen von Konstrukten sind dabei dadurch gekennzeichnet, dass die Indikatoren die kausale Folge der latenten Variable darstellen und wegen desselben Ursprungs hoch miteinander korreliert sein sollten, weshalb sie bei nur geringen Auswirkungen auf das Konstrukt austauschbar und teilweise auch löschar sind (Henseler & Fassott, 2010). Werden latente Variablen formativ gemessen, müssen die Indikatoren jedoch nicht wie bei reflektiven Messungen derselben Konstrukt-Domäne entspringen, sondern sie formen bzw. verursachen vielmehr die latente Variable, weshalb diese als die Konsequenz der zugrundeliegenden Indikatorvariablen betrachtet werden kann (Hair et al., 2017, S. 249, Henseler & Chin, 2010). Da jeder Indikator einen spezifischen Aspekt der Domäne des formativ gemessenen Konstruktes abbildet, sind die Indikatoren nicht austauschbar oder ohne weiteres löschar, weil dies die Bedeutung des Konstruktes verändern würde (Hair et al., 2017, S. 47). Darüber hinaus können latente Variablen auch mithilfe nur eines einzelnen Indikators gemessen werden, wobei der Indikator in diesem Fall weder den Ursprung noch die Konsequenz der Variable darstellt und das Messmodell damit unerheblich ist (Hair et al., 2017, S. 108-109; Henseler & Fassott, 2010).

Hair et al. (2011) unterscheiden bei Methoden zur Schätzung von Strukturgleichungsmodellen zwischen Covariance-Based SEM (CB-SEM) und PLS-SEM. Die Autoren gehen darauf ein, dass PLS-SEM ein Ansatz zur kausalen Modellierung sei und das Ziel verfolge, die erklärte Varianz des betrachteten abhängigen Konstruktes zu maximieren. PLS-SEM wird deshalb vorrangig als explorative Methode betrachtet, bei der mit begrenztem Vorwissen zu den Beziehungen der zu inkludierenden Variablen nach möglichen Richtungen und Stärken der Beziehungen gesucht wird und bei der im Zuge dessen Theorien der Zusammenhänge zwischen den Variablen entwickelt werden (Hair et al., 2017, S. 3-4). Das Ziel von CB-SEM stellt hingegen die Überprüfung bestehender Theorien über strukturelle Beziehungen dar und die Verwendung erfordert die Erfüllung bestimmter Voraussetzungen, wie ein ausreichend großer Stichprobenumfang und eine Normalverteilung der Daten (Hair et al., 2011). Diese Voraussetzungen sind bei der Verwendung von PLS-SEM eher zu vernachlässigen. Deshalb sollten Forscher sich bei der Schätzung von Netzwerken aus latenten Variablen für die Verwendung von PLS-SEM entscheiden, wenn das Ziel die Voraussage von Beziehungen zwischen Konstrukten darstellt, ein kleiner Stichprobenumfang vorliegt, die erhobenen Daten nicht normalverteilt sind, formativ gemessene Konstrukte im Modell vorkommen und das Modell durch viele Konstrukte und Indikatoren sehr komplex ist

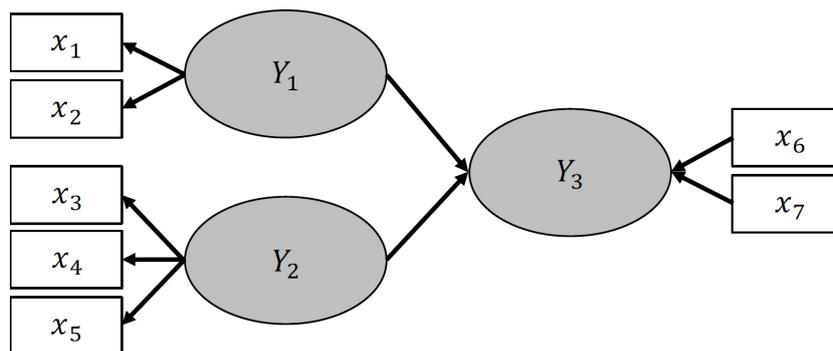


Abbildung 1: Beispielhafte Darstellung eines Pfadmodells (in Anlehnung an Hair et al., 2011)

(Hair et al., 2017, S. 23, Hair et al., 2011). In der vorliegenden Arbeit liegt der Fokus auf PLS-SEM, weshalb CB-SEM größtenteils von den nachfolgenden Betrachtungen ausgeschlossen wird.

Der varianzbasierte Algorithmus, welcher PLS-SEM zugrunde liegt, geht zurück auf Herman Wold (bspw. Wold, 1975) und basiert nach Hair et al. (2011) auf einem zweistufigen Ansatz mit dem Ziel der Maximierung der erklärten Varianz in den endogenen Konstrukten im Modell. Die erste Stufe umfasst nach den Autoren die Approximation der Werte der latenten Konstrukte (latent variable scores (LVS)) für jedes Konstrukt im Modell. Die zweite Stufe schätzt mithilfe der zuvor approximierten LVS, welche nun in je einer multiplen Ordinary Least Squares Regression pro endogenem Konstrukt als unabhängige und abhängige Variablen (je nach Position der betrachteten Konstrukte im Pfadmodell) dienen, die Pfadkoeffizienten des Strukturmodells sowie die finalen Gewichte und Ladungen der Indikatoren (Esposito Vinzi et al., 2010; Hair et al., 2011; Hair et al., 2017, S. 82-86). Gewichte und Ladungen umfassen nach Hair et al. (2011) dabei die Koeffizienten formativ bzw. reflektiv gemessener Beziehungen in PLS-SEM. Dabei wird nach den Autoren die erste Stufe so oft wiederholt, bis die Summe der Veränderungen der berechneten äußeren Gewichte bzw. Ladungen zwischen zwei Iterationen kleiner als ein zuvor festgelegtes Limit ist und die Gewichte bzw. Ladungen dementsprechend konvergieren. Die zweite Stufe bringt schließlich neben Pfadkoeffizienten und den finalen Gewichten und Ladungen auch die erklärte Varianz der endogenen latenten Variablen in Form des Bestimmtheitsmaßes R^2 hervor (Hair et al., 2017, S. 86).

Letztlich sollte nach Hair et al. (2017, S. 104-110) einer Interpretation der Ergebnisse von PLS-SEM zuvor die Evaluation der reflektiven und formativen Messmodelle sowie des gesamten Strukturmodells vorangehen. Dabei sollten nach den Autoren bei einem oder mehreren reflektiv gemessenen Konstrukten im Modell die interne Konsistenzreliabilität, konvergente Validität und diskriminante Validität evaluiert werden. Weiterhin weisen die Autoren darauf hin, dass formative Messmodelle im Modell ebenfalls im Hinblick auf konvergente Validität geprüft werden sollten, darüber hinaus aber auch auf Kollinearität zwischen den Indikatoren sowie auf die Signifikanz und Relevanz der äußeren Gewichte.

Das gesamte Strukturmodell ist nach den Autoren anschließend auf das Bestimmtheitsmaß R^2 , die prädiktive Relevanz Q^2 , Größe und Signifikanz der Pfadkoeffizienten sowie auf die Effektgrößen f^2 und q^2 zu analysieren. Dabei wird bei der Interpretation der Ergebnisse die Signifikanz der Pfadkoeffizienten aufgrund der nichtparametrischen Eigenschaften von PLS-SEM, wegen der keine spezifische Verteilung von Daten vorausgesetzt wird, durch die Herleitung einer eigenen Verteilung der zugrundeliegenden Daten mithilfe des sog. Bootstrapping-Verfahrens getestet (Hair et al., 2017, S. 87). Hierbei werden Standardfehler produziert, aus denen t-Werte und entsprechend auch p-Werte geschätzt werden, die wiederum zur Betrachtung der statistischen Signifikanz der Pfadkoeffizienten dienen (Hair et al., 2017, S. 185-186).

2.2. Moderation in Marketinganwendungen

2.2.1. Definition und Verwendung von Moderation

Moderation bzw. ME treten auf, wenn die Veränderung des Wertes einer dritten Variable (der MV) einen Einfluss auf die Stärke und/oder Richtung der Beziehung zweier anderer Variablen aufweist (Baron & Kenny, 1986, S. 1174). Die MV kann jedoch neben einem Einfluss auf eine spezifische Beziehung auch das gesamte Modell und demnach jede Beziehung in ihm beeinflussen, wobei die Zusammenhänge zwischen den Variablen hierbei nicht für die gesamte Stichprobe konstant sind, sondern in Abhängigkeit vom Moderator variieren (Hair et al., 2017, S. 243). In letzterem Fall kann ein ME auf alle Beziehungen im Modell mithilfe eines Vergleichs mehrerer Untergruppen der Stichprobe in Form einer Multigruppenanalyse getestet werden (Memon et al., 2019). Multigruppenanalysen werden aus den aktuellen Betrachtungen jedoch ausgeschlossen, da der Gegenstand dieser Arbeit insb. der Fall darstellt, in dem eine Moderatorvariable auf eine spezifische Beziehung im Modell wirkt. Diese sog. einfache MA wird in Abbildung 2 visualisiert, wobei Einkommen als MV auf den Zusammenhang zwischen der exogenen Variable Kundenzufriedenheit und der endogenen Variable Kundenloyalität wirkt (Hair et al., 2017, S. 243-244; Memon et al., 2019)¹.

¹Im weiteren Verlauf der Arbeit meint eine MA immer eine einfache MA, bei welcher der Einfluss einer MV auf die Beziehung einer exogenen und einer endogenen Variable geschätzt werden soll.

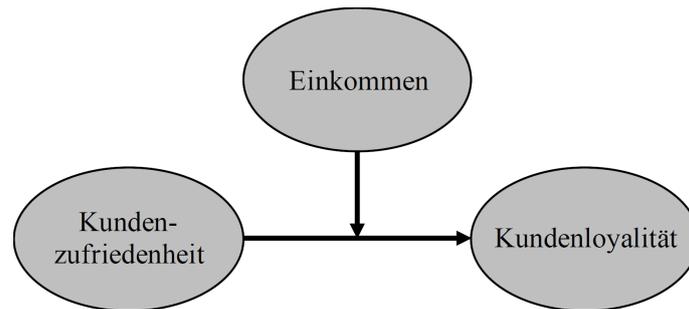


Abbildung 2: Beispielhafte Darstellung eines ME (in Anlehnung an Hair et al., 2017, S. 244)

Der hierbei auftretende ME wird auch als Interaktionseffekt bezeichnet, wobei Hall und Sammons (2014) Interaktionseffekte zunächst als den kombinierten Effekt zweier unabhängiger Variablen auf eine abhängige Variable beschreiben. Die Autoren weisen dabei darauf hin, dass ein Interaktionseffekt nur äquivalent zu einem ME sei, insofern die zugrundeliegende theoretische Grundlage dies gewährleistet. Ein signifikanter Interaktionseffekt kann demnach nur als ME interpretiert werden, insofern der ME a priori hypothetisiert wurde (Field, 2018, S. 486). Auch nach Hair et al. (2017, S. 243) sollte die moderierte Beziehung a priori hypothetisiert werden, wobei zu erkennen sein muss, ob die MV erwartungsgemäß einen Einfluss auf eine oder alle Beziehungen im Modell haben wird. Dabei sollte die Hypothese auf einer theoretischen Herleitung basieren, d.h. auf existierender Literatur und einer Begründung, wieso die Inklusion des Moderators zu einer besseren Erklärung einer Beobachtung führt und in welche Richtung der ME erwartungsgemäß wirken wird (Andersson, Cuervo-Cazurra & Nielsen, 2014; Frazier, Tix & Barron, 2004). Der theoretische Hintergrund kann nach Memon et al. (2019) dabei auf Grundlage von systematischen Literaturrecherchen oder Metaanalysen und den darin enthaltenen Vorschlägen zu zukünftiger Forschung und identifizierten potenziellen Moderatoren gebildet werden. Weiterhin dienen nach den Autoren zur Identifikation potenzieller ME auch Widersprüche in Erkenntnissen verschiedener Forscher, welche Beziehungen zwischen Variablen in unterschiedlichen Bedingungen getestet haben, oder Übertragungen von Erkenntnissen aus anderen wissenschaftlichen Kontexten in den aktuellen.

Unabhängig, ob MV, wie bspw. das Einkommen, Alter, Geschlecht oder die Nationalität, auf eine oder mehrere Beziehungen wirken, können sie dazu in der Lage sein, Heterogenität in den Daten zu erklären (Hair et al., 2017, S. 243). Forscher fragen sich demnach wann, unter welchen Umständen bzw. für wen eine oder mehrere zuvor definierte kausale Beziehungen zu beobachten sind, wenn sie Moderatorvariablen einbinden möchten (Hall & Sammons, 2014). Das übergeordnete Ziel besteht dabei meistens aus der Generierung neuer theoretischer Erkenntnisse und Theorien oder der Schließung von Forschungslücken (Andersson et al., 2014).

2.2.2. Arten von Moderatorvariablen

Die MV kann direkt oder nicht direkt beobachtbare Eigenschaften, wie bspw. die Einstellung zu einer Marke, umfassen, die entsprechend mit einem Indikator oder Skalen aus mehreren formativen oder reflektiven Indikatoren messbar sind (Hair et al., 2017, S. 244-245). Dabei wird nach Henseler und Fassott (2010) in erster Linie abhängig von der verwendeten Skalierung der MV zwischen kontinuierlichen und kategorischen MV unterschieden.

Kontinuierliche MV treten in Form von metrischen Variablen auf und werden deshalb durch Likert-Skalen, unter der Annahme gleichgroßer Abstände zwischen den Ausprägungen der Skala, und somit in Form von quasi-metrischen, sowie mit intervall- und verhältnisskalierten Variablen gemessen (Hair et al., 2017, S. 8-9; Henseler & Fassott, 2010). Der zuvor in Abbildung 2 gezeigte Moderator stellt bspw. eine kontinuierliche MV dar, wobei die Höhe der MV die Beziehung der exogenen und endogenen Variable in ihrer Stärke bzw. Richtung beeinflusst, insofern ein signifikanter ME vorliegt (Hair et al., 2017, S. 246). Die Beziehung wäre demnach konstant, wenn kein signifikanter ME wirkt.

Kategorische MV werden nominal- oder ordinalskaliert gemessen und beinhalten nur einen Indikator pro Konstrukt, welcher die Kategorien in Form von mehreren Antwortmöglichkeiten abbildet (Henseler & Fassott, 2010). Sie haben ebenfalls einen Einfluss auf die Richtung bzw. Stärke der Beziehung der exogenen und endogenen Variable zwischen verschiedenen Kategorien, können darüber hinaus aber auch zeigen, dass in einer Kategorie ein Effekt zwischen exogener und endogener Variable auftritt, während in anderen kein Effekt beobachtbar ist (Field, 2018, S.484). So kann bspw. das Geschlecht als MV die Beziehung zweier Variablen beeinflussen bzw. aufzeigen, dass ein spezifischer Effekt einer exogenen auf eine endogene Variable lediglich bei Frauen und nicht bei Männern beobachtbar ist. In der Regel werden die Kategorien dummy-kodiert, wobei die Referenzkategorie, mit welcher eine oder mehrere andere Kategorien verglichen werden sollen, mit „0“ kodiert wird (Hair et al., 2017, S. 245). Beinhaltet die MV mehr als zwei Kategorien, werden diese nach Henseler und Fassott (2010) mithilfe mehrerer Dummy-Variablen gleichzeitig in das Modell eingeschlossen.

Es besteht bei kontinuierlichen MV nach den Autoren ebenfalls die Möglichkeit mithilfe von Dichotomisierung, also

der Teilung der Stichprobe zur Schaffung mindestens zweier künstlicher Gruppen, Vergleiche von Kategorien durchzuführen. Dies ist jedoch mit dem Verlust von Informationen in Form von Varianz der MV und reduzierter statistischer Power² verbunden oder führt sogar dazu, dass Effekte gefunden werden, wo in der Population eigentlich keine auftreten (Frazier et al., 2004). Deshalb empfehlen Henseler und Fassott (2010) sich stets gegen eine Dichotomisierung kontinuierlicher MV zu entscheiden und sie nicht mithilfe eines Gruppenvergleichs zu analysieren.

2.3. Moderation in PLS-SEM

2.3.1. Modellierung des Moderationseffekts

Soll eine einfache MA in PLS-SEM, wie in Abbildung 3 gezeigt, mithilfe der Modellierung eines spezifischen ME (p_3) auf die Beziehung (p_1) einer exogenen (Y_1) und endogenen (Y_2) latenten Variable durchgeführt werden, erfordert dies die Inklusion einer MV (M) in das Pfadmodell, die ebenfalls einen direkten Einfluss (p_2) auf die endogene Variable hat (Hair et al., 2017, S. 247). Der Pfad p_2 verhindert dabei eine Überschätzung des ME (p_3) durch die Kontrolle des direkten Effektes der latenten MV auf das endogene Konstrukt, wobei die Steigung der exogenen latenten Variable nicht konstant sein, sondern linear von der Höhe des Moderators abhängen soll (Henseler & Fassott, 2010).

Der Einschluss des Moderatorkonstrukts erfolgt dabei zunächst in die folgende lineare Gleichung, welche die Main Effects (p_1 und p_2) von Y_1 und M auf Y_2 abbildet, wobei die Konstante p_0 den Schnittpunkt mit der Ordinate darstellt, d.h. wenn alle Steigungen (Regressionsparameter) der Variablen Y_1 und M (d.h. p_1 und p_2) Null sind (Henseler & Fassott, 2010):

$$Y_2 = p_0 + p_1 \cdot Y_1 + p_2 \cdot M + \varepsilon \quad (1)$$

Dabei bildet der Fehlerterm ε die nicht erklärte Varianz ab, d.h. den Messfehler, der bei der Messung der latenten Variablen auftritt (Becker et al., 2018; Henseler & Chin, 2010).

Der in Abbildung 3 dargestellte ME kann anschließend durch die Inklusion des Pfades p_3 nach (Hair et al., 2017, S. 247) sowie Henseler und Fassott (2010) mit der folgenden Gleichung ausgedrückt werden, in der die Stärke der exogenen Variable Y_1 vom Niveau von M abhängt ($(p_1 + p_3 \cdot M) \cdot Y_1$), weshalb p_1 nicht mehr als Main Effect sondern nun als Simple Effect bezeichnet wird:

$$Y_2 = p_0 + (p_1 + p_3 \cdot M) \cdot Y_1 + p_2 \cdot M + \varepsilon \quad (2)$$

Diese Gleichung kann umgestellt werden, um zu verdeutlichen, wie der ME in PLS-SEM integriert wird, wobei das Produkt $Y_1 \cdot M$ den sog. Interaktionsterm darstellt und der ME p_3 im Fall standardisierter LVS ausdrückt, inwieweit sich

der Simple Effect (p_1) verändert, wenn M um eine Standardabweichung verringert oder erhöht wird (Becker et al., 2018; Hair et al., 2017, S. 247-248):

$$Y_2 = p_0 + p_1 \cdot Y_1 + p_2 \cdot M + p_3 \cdot (Y_1 \cdot M) + \varepsilon \quad (3)$$

Der als zusätzliche latente Variable in das PLS-SEM Pfadmodell eingebundene Interaktionsterm wird in Abbildung 4 visualisiert und demnach zusammen mit der exogenen latenten Variable Y_1 und dem Moderatorkonstrukt M in demselben Modell geschätzt (Field, 2018, S. 486; Hair et al., 2017, S. 248). Hierbei drückt der Simple Effect (p_1) die Stärke eines Effektes aus, wenn alle anderen Steigungen (p_2 und p_3) null sind, d.h. wenn die MV einen Wert von null hat, womit sich der Simple Effect in seiner Berechnung und Bedeutung vom Main Effect (Effekt, wenn nur $p_2 = 0$) unterscheidet und das Modell mit einem ME anders als das Modell ohne ME betrachtet werden muss (Hair et al., 2017, S. 257-258; Henseler & Fassott, 2010). Letztlich zeigt jedoch der Pfadkoeffizient des Interaktionsterms p_3 und damit der Interaktionseffekt (ME) zusammen mit einem hinreichenden zugrundeliegenden theoretischen Fundament auf, ob eine signifikante Moderation der Beziehung der exogenen und endogenen Variable vorliegt oder nicht (Field, 2018, S. 485-486).

Soll mehr als ein latenter Moderator auf dieselbe Beziehung wirken, was bspw. in Form einer „Cascaded MA“, d.h. einer moderierten Moderation, erfolgen kann, muss dieser nach Henseler und Fassott (2010) innerhalb einer Three-Way Interaction modelliert werden. So ist nach Hair et al. (2017, S. 248) bspw. der ME von Einkommen auf die Beziehung zweier Konstrukte nicht konstant, sondern schwankt ebenfalls durch die Veränderung der MV Alter. Dabei werden im Fall von zwei MV (M_1 und M_2) und jeweils einer exogenen (Y_1) und endogenen (Y_2) Variable nach den Autoren insgesamt acht latente Variablen modelliert, da neben der „Three-Way Interaktion“ ($p_7 \cdot (Y_1 \cdot M_1 \cdot M_2)$) auch die Simple Effects ($p_{(1)}$ bis p_3) und die zuvor diskutierten „Two-Way Interactions“ ($p_4 \cdot (Y_1 \cdot M_1) + p_5 \cdot (Y_1 \cdot M_2) + p_6 \cdot (M_1 \cdot M_2)$) in der zugrundeliegenden linearen Gleichung beinhaltet sein müssen:

$$Y_2 = p_0 + p_1 \cdot Y_1 + p_2 \cdot M_1 + p_3 \cdot M_2 + p_4 \cdot (Y_1 \cdot M_1) + p_5 \cdot (Y_1 \cdot M_2) + p_6 \cdot (M_1 \cdot M_2) + p_7 \cdot (Y_1 \cdot M_1 \cdot M_2) + \varepsilon \quad (4)$$

2.3.2. Der Nutzen von PLS-SEM bei der Analyse von Moderationseffekten

Sollen ME innerhalb multivariater Analysen mit latenten Variablen analysiert werden, haben Forscher nach Sarstedt, Ringle et al. (2020) die Wahl zwischen CB-SEM und PLS-SEM, aber auch Regressionsanalysen mithilfe des PROCESS-Makros von Hayes (2013) finden hierbei Anwendung. Dabei unterliegen Regressionsanalysen nach den Autoren allerdings dem Nachteil, dass sie keine Messfehler bei der Berechnung von LVS in die Betrachtung mit einbeziehen. Hair et al. (2017) zeigen, dass eine Kombination aus Über- und Unterschätzung der Berechnung des gesamten Pfadmodells

²Nach Cohen (1988, S. 4) ist die statistische Power eines Tests die Wahrscheinlichkeit dafür, einen signifikanten Effekt zu finden, wenn dieser tatsächlich in der Population existiert.

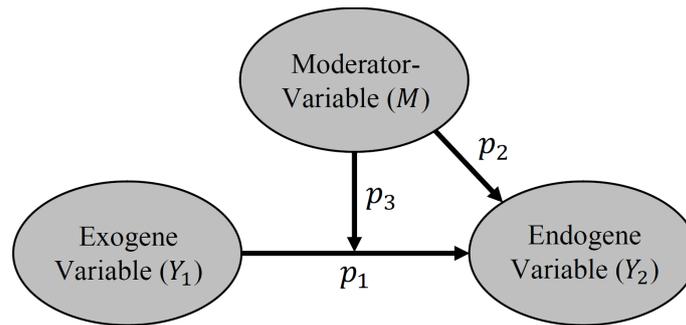


Abbildung 3: Moderationseffekt im Pfadmodell (in Anlehnung an Hair et al., 2017, S. 247)

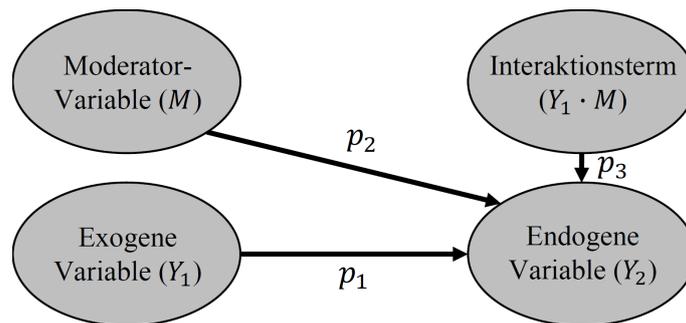


Abbildung 4: Einbindung eines Interaktionsterms (in Anlehnung an Hair et al., 2017, S. 248)

resultiert, wenn Messfehler nicht korrigiert werden. Die mit der Berechnung der LVS einhergehenden Messfehler können allerdings mithilfe von SEM berücksichtigt werden, was die Verzerrung des Pfadmodells verhindert (Sarstedt et al., 2016). Insbesondere bei der Analyse von MA ist dies von Bedeutung, da bei der Generierung des Interaktionsterms der verzerrende Effekt von Messfehlern durch die Multiplikation zweier fehlerbehafteter Konstrukte verstärkt wird (Li et al., 2019). Hinzukommt, dass ME zumeist klein ausfallen, was auch zu einer geringen statistischen Power beiträgt, weshalb bereits eine marginale negative Verzerrung des Effektes schwer ins Gewicht fallen und den Effekt nicht auffindbar machen kann (Aguinis, Beaty, Boik & Pierce, 2005; Sarstedt, Hair Jr et al., 2020). Somit ist eine Art der SEM einer Regressionsanalyse vorzuziehen, wenn ME in komplexen Modellen mit latenten Variablen geschätzt werden sollen.

Darüber hinaus ist nach Henseler und Fassott (2010) aufgrund des Ablaufes der zweiten Stufe des PLS-SEM-Algorithmus eine Schätzung von ME in PLS-SEM wie in multiplen Regressionsanalysen möglich. Nach den Autoren und nach Sarstedt, Ringle et al. (2020) ermöglicht die lineare Kombination der Indikatoren der exogenen Variable und der MV zu LVS eine analoge Vorgehensweise wie bei Regressionen, weshalb Forscher den Interaktionseffekt leicht in PLS-SEM abbilden und interpretieren können. CB-SEM hingegen liefert keine LVS, weshalb Interaktionen schwieriger abzubilden sind (Hair et al., 2017, S. 15). Hinzukommt, dass PLS-SEM mehrere Ansätze zur Modellierung des Interaktionsterms liefert, auf welche nachfolgend im Detail eingegangen wird und die es ermöglichen, sowohl reflektiv

als auch formativ gemessene Konstrukte als MV in PLS-SEM einzubinden (Memon et al., 2019). Auch komplexe Modelle mit einer Vielzahl von Konstrukten und Indikatoren, sowie komplexe Beziehungen (d.h. Moderationen und Mediationen von Beziehungen) lassen sich ohne Probleme in PLS-SEM abbilden (Sarstedt, Hair Jr et al., 2020). Anhand der dargelegten Argumente kann PLS-SEM als der zu favorisierende SEM-Ansatz zur Durchführung von MA in Modellen mit latenten Variablen identifiziert werden.

2.3.3. Ansätze zur Modellierung des Interaktionsterms

Der aus latenten Variablen generierte Interaktionsterm zur Abbildung eines ME in PLS-SEM kann in MA mithilfe verschiedener Ansätze generiert werden: Produkt-Indikator-Ansatz, Orthogonalisierungsansatz, Zwei-Stufen-Ansatz und hybrider Ansatz (Henseler & Chin, 2010). Diese sind in Tabelle 1 zusammengefasst, die einen Vergleich der vier Ansätze anhand der Funktionsweisen, der Anwendbarkeit in Abhängigkeit von den Messmodellen der zugehörigen MV und der exogenen Variablen sowie der Vor- und Nachteile der Ansätze ermöglicht.

Der in Regressionsanalysen als Standardansatz zur Generierung des Interaktionsterms geltende Produkt-Indikator-Ansatz wurde erstmals von Chin, Marcolin und Newsted (2003) in den PLS-SEM-Kontext übertragen, um den zuvor diskutierten Nachteilen von Regressionsanalysen bei der Schätzung von ME im Zusammenhang mit latenten Variablen auszuweichen (Hair et al., 2017, S. 249; Henseler & Fassott, 2010). Henseler und Chin (2010) gehen darauf ein, dass die hierbei erfolgende Bildung des Interaktionsterms durch

die paarweise Multiplikation jedes Indikators der latenten exogenen Variable mit denen der MV relativ einfach in PLS-SEM möglich sei. Die so entstehenden Produkt-Indikatoren dienen nach den Autoren schließlich im Strukturmodell als die Indikatoren des Interaktionsterms. Abbildung 5 verdeutlicht die Funktionsweise des Produkt-Indikatoren Ansatzes im Fall von drei Indikatoren der exogenen Variable und zwei Indikatoren der MV (Hair et al., 2017, S. 250).

Es ist hierbei zu erkennen, dass sich aus den Produkten aller möglichen Kombinationen der Indikatoren der MV und exogenen Variable insg. sechs Produkt-Indikatoren ergeben, die den Interaktionsterm reflektieren.

Die multiplizierten Indikatoren müssen bei der Verwendung des Produkt-Indikator-Ansatzes von reflektiv gemessenen Konstrukten stammen, da nur so angenommen werden kann, dass diese auch tatsächlich jeweils zu genau einer entsprechenden inhaltlichen Domäne der MV bzw. der exogenen latenten Variable gehören und die Multiplikation dieser Indikatoren auch tatsächlich eine Interaktion der Variablen und ihrer Domänen widerspiegelt (Hair et al., 2017, S. 249). Die Indikatoren müssen dafür austauschbar sein und recht hoch miteinander korrelieren (Henseler & Chin, 2010). Formativ gemessene Indikatoren können jedoch unabhängig voneinander sein und müssen nicht zwingend derselben Domäne eines Konstruktes entspringen, das sie formen, weshalb formativ gemessenen Konstrukten keine Domäne eindeutig zugeordnet werden kann (Hair et al., 2017, S. 249; Henseler & Fassott, 2010). Deshalb lassen Produkt-Indikatoren, die vollständig oder zum Teil aus einer Menge formativer Indikatoren bestehen, bei der Multiplikation der Indikatoren nicht zwingend auf denselben zugrunde liegenden Interaktionseffekt schließen (Chin et al., 2003). Dadurch wird die konzeptuelle Domäne des Interaktionsterms gestört und er verliert seine inhaltliche Bedeutung, was dazu führt, dass der Produkt-Indikator-Ansatz bei formativ gemessenen Konstrukten konzeptuell nicht abbildbar ist (Hair et al., 2017, S. 249).

Letztlich wird durch die Wiederverwendung der Indikatoren des exogenen Konstrukts und des Moderatoronstrukts im Messmodell des Interaktionsterms Multikollinearität, also eine hohe Korrelation der Konstrukte, im Pfadmodell hervorgerufen (Henseler & Chin, 2010). Dies kann nach Hair et al. (2017, S. 249-251) zu verzerrten Schätzungen der Pfadkoeffizienten oder überschätzten Standardfehlern führen, da die Indikatoren des Interaktionsterms einen Teil der Varianz der Indikatoren der exogenen Variable und der MV teilen. Zur Reduktion der Multikollinearität und Erleichterung der Interpretation des ME werden deshalb nach den Autoren die Indikatoren der latenten MV vor der Generierung des Interaktionsterms häufig standardisiert (Umwandlung zu Variablen mit Mittelwert von null und Standardabweichung von eins).

Die Autoren weisen weiterhin darauf hin, dass die Standardisierung der Indikatoren vor der Generierung des Interaktionsterms jedoch die resultierende Multikollinearität nicht vollständig eliminiert und den Vergleich von Main Effect und Simple Effect bei der Interpretation des ME erschwert. Um die Korrelation des Interaktionsterms und der

exogenen Variablen und der MV zu beseitigen, haben Little, Bovaird und Widaman (2006) nach den Autoren den Orthogonalisierungsansatz als Erweiterung des Produkt-Indikator-Ansatzes entwickelt, der ebenfalls die Bildung von Produkt-Indikatoren des Interaktionsterms auf dieselbe Art und Weise beinhaltet. Diese werden jedoch nicht direkt als Indikatoren des Interaktionsterms genutzt, sondern es wird eine Regression pro Produkt-Indikator durchgeführt, wobei diese als abhängige und die Indikatoren der latenten exogenen Variable und MV als unabhängige Variablen in der Regression dienen (Henseler & Chin, 2010). Im Fall des in Abbildung 6 gezeigten Modells würden aufgrund von drei Indikatoren des exogenen Konstrukts und zwei des Moderatoronstrukts insgesamt sechs Regressionsmodelle folgen, da bei der paarweisen Multiplikation aller Indikatoren sechs Produkt-Indikatoren entstehen (Hair et al., 2017, S. 250-251):

$$x_1 \cdot m_1 = b_{1,11} \cdot x_1 + b_{2,11} \cdot x_2 + b_{3,11} \cdot x_3 + b_{4,11} \cdot m_1 + b_{5,11} \cdot m_2 + e_{11}$$

$$x_1 \cdot m_2 = b_{1,12} \cdot x_1 + b_{2,12} \cdot x_2 + b_{3,12} \cdot x_3 + b_{4,12} \cdot m_1 + b_{5,12} \cdot m_2 + e_{12}$$

$$x_2 \cdot m_1 = b_{1,21} \cdot x_1 + b_{2,21} \cdot x_2 + b_{3,21} \cdot x_3 + b_{4,21} \cdot m_1 + b_{5,21} \cdot m_2 + e_{21}$$

$$x_2 \cdot m_2 = b_{1,22} \cdot x_1 + b_{2,22} \cdot x_2 + b_{3,22} \cdot x_3 + b_{4,22} \cdot m_1 + b_{5,22} \cdot m_2 + e_{22}$$

$$x_3 \cdot m_1 = b_{1,31} \cdot x_1 + b_{2,31} \cdot x_2 + b_{3,31} \cdot x_3 + b_{4,31} \cdot m_1 + b_{5,31} \cdot m_2 + e_{31}$$

$$x_3 \cdot m_2 = b_{1,32} \cdot x_1 + b_{2,32} \cdot x_2 + b_{3,32} \cdot x_3 + b_{4,32} \cdot m_1 + b_{5,32} \cdot m_2 + e_{32}$$

Die sog. standardisierten Residuen (e_{11} bis e_{32}) werden anschließend als Indikatoren des Interaktionsterms genutzt, was auch als residuale Zentrierung bezeichnet wird und ebenfalls in Abbildung 6 dargestellt ist (Henseler & Chin, 2010). Dieses zweistufige Prozedere sorgt dafür, dass die Varianz des neuen orthogonalen Interaktionsterms von der Varianz der exogenen Variable und der MV unabhängig ist und nur den Interaktionsterm repräsentiert (Little et al., 2006). Dadurch ist der Interaktionsterm unkorreliert mit den zugrundeliegenden Konstrukten (Becker et al., 2018).

Hair et al. (2017, S. 251) verweisen hierbei darauf, dass die Parameterschätzungen des Main Effects und Simple Effects in diesem Fall sehr gleich und bei der Interpretation des ME äquivalent zu betrachten sind, was diese vereinfacht. Dieses Fehlen eines Simple Effects kann nach Echambadi, Aroniz, Reinartz und Lee (2006) und Becker et al. (2018) allerdings auch zu Problemen in der Interpretation von Effekten führen, da die Möglichkeit, den Simple Effect zu interpretieren ausbleibt und der Main Effect fälschlicherweise als Simple Effect interpretiert wird. Letztlich kann auch dieser Ansatz nur bei reflektiv gemessenem exogenen und moderierenden Konstrukt angewendet werden, da er auf der Bildung von Produktindikatoren beruht, für welche dieselben Voraussetzungen wie im Produkt-Indikator-Ansatz gelten (Hair et al., 2017, S. 251).

Damit auch formativ gemessene latente Variablen in MA untersucht werden können, haben Chin et al. (2003) den Zwei-Stufen-Ansatz vorgeschlagen, welcher durch die von PLS-SEM ermöglichte Schätzung der LVS von Konstrukten die MA in zwei Stufen durchführt (Henseler & Fassott, 2010). Wie in Abbildung 7 zu erkennen ist, wird in der ersten Stufe zunächst das Modell der Main Effects geschätzt, in welchem

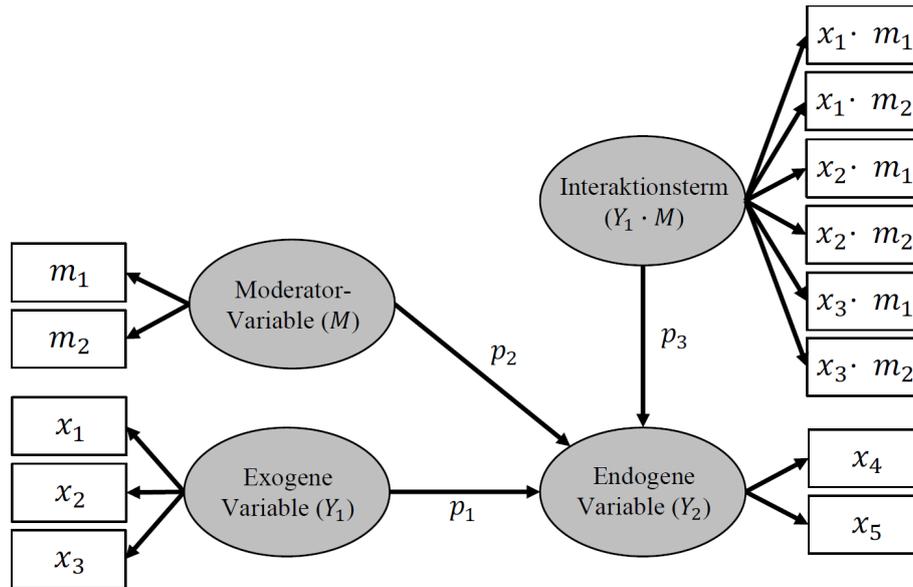


Abbildung 5: Produkt-Indikator-Ansatz (in Anlehnung an Hair et al., 2017, S. 250)

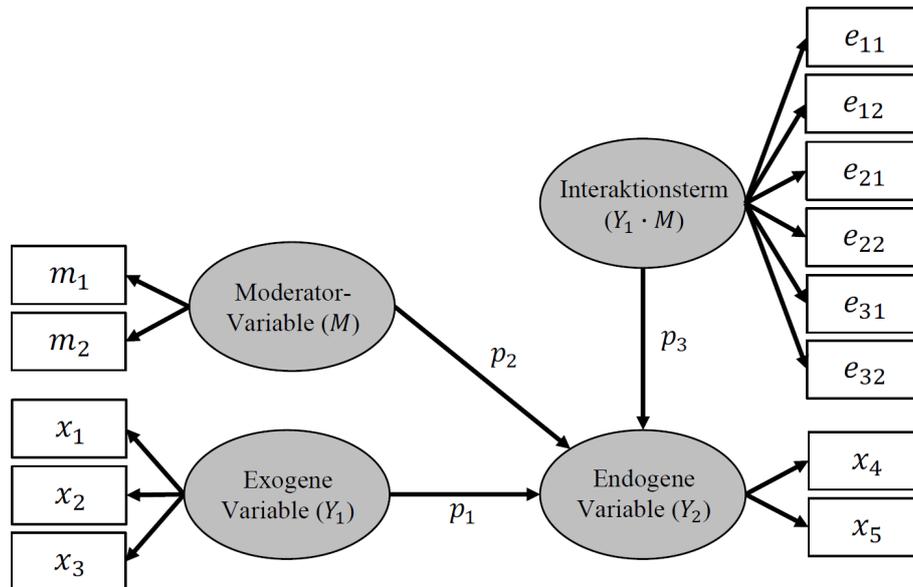


Abbildung 6: Orthogonalisierungsansatz (in Anlehnung an Hair et al., 2017, S. 252)

der Interaktionsterm noch nicht enthalten ist, und es entstehen die LVS der latenten exogenen Variable und der MV, die für die zweite Stufe gespeichert werden, in der die Multiplikation dieser LVS zu einem einzigen Indikator erfolgt, mit dem schließlich der Interaktionsterm gemessen wird (Hair et al., 2017, S. 251-252). Anschließend werden in der zweiten Stufe neben dem Interaktionsterm alle anderen latenten Variablen mit ihrem zuvor bestimmten LVS als unabhängige Variablen in einer multiplen linearen Regression auf die endogene Variable Y_2 genutzt (Henseler & Fassott, 2010).

Hierbei werden nach Henseler und Chin (2010) die LVS durch den PLS-SEM-Algorithmus stets standardisiert, der Interaktionsterm jedoch nicht, was ohnehin Schwierigkeiten in

der Interpretation des Interaktionseffektes hervorrufen würde. Durch die Standardisierung werden auch die bei diesem Ansatz auftretenden Probleme reduziert, die durch Multikollinearität im Strukturmodell hervorgerufen werden (Becker et al., 2018; Fassott, Henseler & Coelho, 2016). Letztlich weisen Henseler und Chin (2010) darauf hin, dass der Zweistufen-Ansatz auch bei ausschließlich reflektiv gemessenen MV sowie exogenen Variablen angewendet werden kann und nicht auf formative gemessene Konstrukte beschränkt ist.

Den finalen in der Literatur häufig besprochenen Ansatz stellt der hybride Ansatz dar, dessen Funktionsweise Wold (1982) geprägt hat und der von Henseler und Chin (2010) auf MA in PLS-SEM übertragen wurde.

Tabelle 1: Übersicht der gängigen Ansätze zur Modellierung des Interaktionsterms in PLS-SEM, deren Funktionsweise, die jeweils unterstützten Messmodelle und Vor- sowie Nachteile (in Anlehnung an Becker et al., 2018; Hair et al., 2017, S. 249-252; Henseler & Chin, 2010; Henseler & Fassott, 2010; Henseler et al., 2012)

Ansatz	Funktionsweise	Unterstützte Messmodelle	Vorteile	Nachteile
Produkt-Indikator-Ansatz (Chin et al., 2003)	Multiplikation der Indikatoren der exogenen Variable und MV zur Generierung von Produktindikatoren des Interaktionsterms	nur reflektiv gemessene exogene und Moderatorstrukture	einfache Anwendung in PLS-SEM	Multikollinearität durch Wiederverwendung der Indikatoren
Orthogonalisierungsansatz (Henseler & Chin, 2010)	Regression der Indikatoren der MV und des exogenen Konstrukts auf die Produktindikatoren des Interaktionsterms und Verwendung der Residuen als Indikatoren des Interaktionsterms	nur reflektiv gemessene exogene und Moderatorstrukture	keine Multikollinearität durch zu anderen Konstrukten orthogonalem Interaktionsterm	(Interpretation des Simple Effects nicht möglich)
Zwei-Stufen-Ansatz (Chin et al., 2003)	1. Stufe: Schätzung des Main Effects Modells zur Generierung der LVS der MV und des exogenen Konstrukts 2. Stufe: Multiplikation beider LVS zur Generierung des einzigen Indikators des Interaktionsterms	reflektiv oder formativ gemessene exogene oder Moderatorstrukture	Anwendung bei formativ gemessener MV und/oder exogener Variable	Multikollinearität durch Wiederverwendung der Indikatoren in Form von LVS
Hybrider Ansatz (Henseler & Chin, 2010; Wold, 1982)	wie Zwei-Stufen-Ansatz, wobei Approximation des Interaktionsterms bei der Schätzung der LVS mit einbezogen wird	reflektiv oder formativ gemessene exogene und Moderatorstrukture	Anwendung bei formativ gemessener MV und/oder exogener Variable	in keiner gängigen Software implementiert

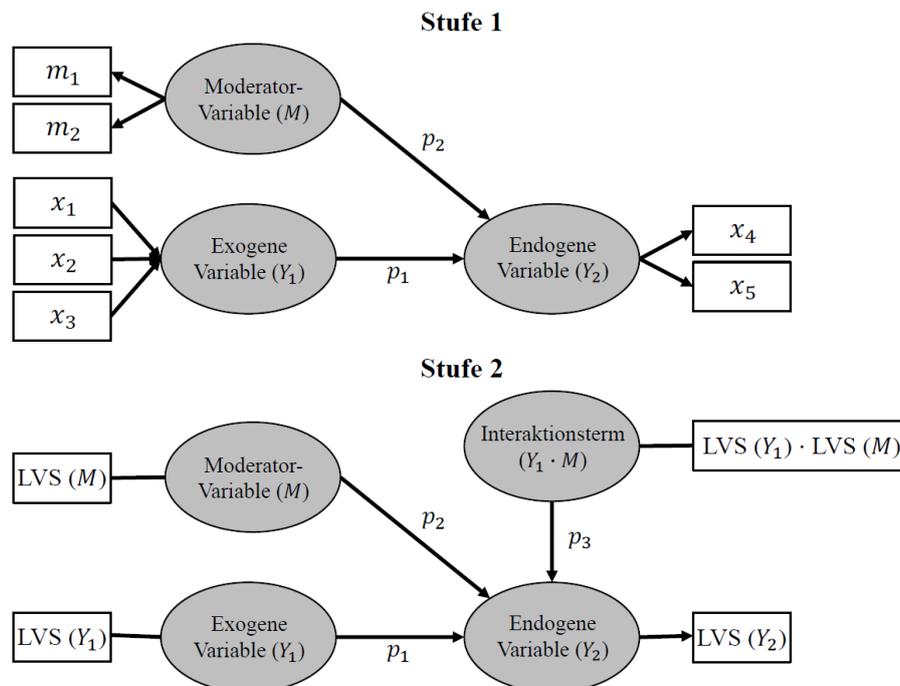


Abbildung 7: Zwei-Stufen-Ansatz (in Anlehnung an Hair et al., 2017, S. 253)

Nach Henseler und Chin (2010) kombiniert dieser Ansatz die Vorteile des Zwei-Stufen-Ansatzes und des Produkt-Indikatoren Ansatzes. Insbesondere gehen die Autoren darauf ein, dass der Interaktionsterm dabei wie bei dem Zwei-Stufen-Ansatz mithilfe des Produkts der LVS der MV und des exogenen Konstruktes gemessen wird. Darüber hinaus findet nach den Autoren wie bei dem Produkt-Indikator-Ansatz allerdings auch die Einbeziehung des Interaktionsterms selbst bei der Generierung der LVS statt, während der PLS-SEM-Algorithmus die LVS bestimmt. Aus dem Grund, dass hierbei LVS genutzt werden, können wie bei dem Zwei-Stufen-Ansatz auch formativ gemessene MV und exogene Konstrukte innerhalb der MA mit dem hybriden Ansatz genutzt werden (Henseler & Chin, 2010). Allerdings wird dieser Ansatz aus den weiteren Betrachtungen von MA in PLS-SEM ausgeschlossen, da er nach der Simulationsstudie von Henseler und Chin (2010) nicht besser als die anderen Ansätze bei der Schätzung von ME in ausschließlich reflektiv gemessenen Modellen funktioniert, nicht in gängiger Software zur Schätzung von PLS-SEM implementiert ist und aus diesem Grund bisher auch seltener von Forschern genutzt wurde (Becker et al., 2018).

Schließlich ist die Wahl des verwendeten Ansatzes unerheblich, wenn die exogene Variable und die MV mit nur einem Indikator gemessen werden, da das Messmodell hierbei keinen Einfluss auf die Messung der „latenten“ Variable hat (Henseler & Fassott, 2010).

2.3.4. Auswertung der Ergebnisse von Moderationsanalysen in PLS-SEM

Modellevaluation

Die zuvor besprochene Evaluation der Messmodelle der latenten Variablen und des gesamten Strukturmodells sollten nach Hair et al. (2017, S. 255) ebenfalls in Modellen mit MV durchgeführt werden. Auch hier muss nach den Autoren im Rahmen der Messmodellevaluation nach reflektiven und formativen Moderatorstrukturen unterschieden werden. Sie weisen jedoch auch darauf hin, dass der Interaktionsterm nicht zu evaluieren ist. Werden bspw. der Orthogonalisierungsansatz oder der Produkt-Indikator-Ansatz bei der MA genutzt, liegen dem Interaktionsterm nach den Autoren durch die Multiplikation der Indikatoren beider Konstrukte anstatt einer spezifischen zwei konzeptionelle Domänen zugrunde. Vielmehr bringen die Autoren zum Ausdruck, dass das Messmodell des Interaktionsterms lediglich ein Hilfsmittel darstellt, um die wechselseitige Beziehung zwischen der latenten MV und dem exogenen Konstrukt im Pfadmodell zu integrieren. Insofern existiert der Interaktionsterm nur im Modell, drückt aber kein wirklich existentes Konstrukt aus, das tatsächlich gemessen werden kann, weshalb die Evaluation hinfällig ist.

Entgegen dieser auf den Prinzipien hinter der Messung latenter Variablen basierenden Sichtweise, bringen Aguinis, Edwards und Bradley (2017) zum Ausdruck, dass die Evaluation des Interaktionsterms durchaus eine große Bedeutung haben sollte. Sie beziehen sich wie auch Li et al. (2019) dabei vorrangig auf die Evaluation der Reliabilität des Interak-

tionsterms. Diese soll nach den Autoren insb. bei der Multiplikation von unkorrelierten Konstrukten aus dem Produkt der einzelnen Reliabilitätsmessungen bestehen und demnach wesentlich geringer ausfallen. Folgt man dieser Argumentation, spricht auch das Argument von (Hair et al., 2017, S. 255) für eine sorgfältige Evaluation des Interaktionsterms, da nach den Autoren die Wiederverwendung von Indikatoren der exogenen und moderierenden Variable bei der Bildung des Interaktionsterms für eine hohe Korrelation der Konstrukte sorgt und damit gegen die diskriminante Validität verstößt. Sie argumentieren dabei allerdings nicht für die Evaluation des Interaktionsterms, sondern dagegen, da die designbedingte Verletzung der diskriminanten Validität für die Autoren einen weiteren Grund darstellt, keine Evaluation durchzuführen.

Unabhängig von diesen unterschiedlichen Sichtweisen weisen die Autoren darauf hin, dass bei der Verwendung des Zwei-Stufen-Ansatzes keine Evaluation des Messmodells des Interaktionsterms erfolgen muss, da dieser lediglich mit einem Indikator aus dem Produkt der LVS der exogenen Variable und der MV gemessen wird. Im Anschluss an die Messmodellevaluation muss nach den Autoren das gesamte Strukturmodell ebenfalls im Hinblick auf die entsprechenden zuvor dargelegten Kriterien evaluiert werden, wobei die Effektgröße f^2 besondere Beachtung erhalten sollte.

Ergebnisinterpretation

Die Interpretation der Ergebnisse von MA umfasst insg. drei Aspekte: Betrachtung der Signifikanz des ME sowie des Simple Effects, Berechnung und Berichten der Effektgröße f^2 , Erfassung der Stärke und Richtung des ME mithilfe eines Simple Slope Plots (Memon et al., 2019). Hierbei ist die inhaltliche Bedeutung dieser Aspekte in den Vordergrund zu stellen und mit dem zuvor dargelegten theoretischen Fundament zusammen zu interpretieren (Andersson et al., 2014). In der Regel sollte ein signifikanter ME und eine vom Forscher als substantiell eingeschätzte Effektgröße f^2 vorliegen, damit sichergestellt werden kann, dass der ME auch tatsächlich einen hinreichend großen Einfluss auf die Beziehung der Variablen ausübt und so wirklich zu den anschließend diskutierten Implikationen der MA für Forschung und Management führt (Henseler et al., 2012).

Die Untersuchung, ob die statistische Signifikanz des Interaktionsterms vorliegt, sollte demnach zunächst der Fokus der Interpretation der Ergebnisse einer MA sein (Hair et al., 2017, S. 256). Dies kann mithilfe des Effekts des Interaktionsterms auf das endogene Konstrukt betrachtet werden, wobei die Hypothese des ME unterstützt werden kann, wenn der entsprechende Pfadkoeffizient (p_3) signifikant ist (Henseler & Fassott, 2010). Durch das Bootstrapping-Prozedere wird dabei festgestellt, ob die MV einen signifikanten ME auf die Beziehung der exogenen und endogenen Variable ausübt, indem der Pfadkoeffizient p_3 darauf geprüft wird, ob sich der ME signifikant von null unterscheidet und damit tatsächlich in der Population existiert (Hair et al., 2017, S. 256; Henseler & Fassott, 2010).

Bei der Interpretation der ME ist nach

Hair et al. (2017, S. 256) ebenfalls eine Betrachtung des Simple Effects und ggf. des Main Effects zu erfolgen. Dabei sollte nach den Autoren der Wert null einen sinnvollen Bestandteil der Skala der MV darstellen, da der Simple Effect p_1 die Beziehung zwischen exogenem und endogenem Konstrukt darstellt, wenn die MV einen Wert von null annimmt. Deshalb wird nach Henseler und Fassott (2010) Mittelwertzentrierung der Indikatoren der MV und des exogenen Konstruktes vorgeschlagen, wobei der Referenzpunkt zum Mittelwert der Indikatoren verschoben wird. Hierfür müssen jedoch die Standardabweichungen der verschiedenen Variablen nach den Autoren in ihrer Interpretation untereinander vergleichbar sein. Hair et al. (2017, S. 257) empfehlen hierfür die Standardisierung der Indikatoren der MV, wodurch die Umwandlung zu Variablen mit dem Mittelwert bei null und der Standardabweichung von eins erfolgt. Somit werden sie in ihrer Bedeutung vergleichbar gemacht und die null stellt einen sinnvollen Bestandteil der Skalen der Variablen dar.

Weiterhin schlagen die Autoren die getrennte Betrachtung des Main Effects und Simple Effects vor, wenn der Produkt-Indikator-Ansatz oder der Zwei-Stufen-Ansatz verwendet werden, da sich der Simple Effect in seiner Bedeutung und dem geschätzten Pfadkoeffizienten vom Main Effect unterscheidet. Nach den Autoren sowie nach Chin et al. (2003) sollten Forscher, die sich für einen Test einer Hypothese interessieren, welche explizit den Main Effect beinhaltet und nicht den von der MV abhängigen Simple Effect, zunächst ein Modell ohne die MV heranziehen, um Verwechslungen der Effekte zu verhindern. Wird der Simple Effect für den Main Effect gehalten, d.h. aus dem Modell entnommen, das den ME enthält, kann dies nach Henseler und Fassott (2010) zu irreführenden oder inkorrekten Schlussfolgerungen führen. Hierbei sollte angemerkt werden, dass Aguinis et al. (2017) und Li et al. (2019) eine andere Sichtweise zum Ausdruck bringen. Die Autoren weisen darauf hin, dass eine Interpretation des Main Effects nicht sinnvoll sei, da das Vorliegen eines signifikanten ME bedeutet, dass kein konstanter Effekt der exogenen auf die endogenen Variable besteht, sondern dieser immer mit der Ausprägung der MV schwankt. Deshalb sollte nach den Autoren stets der Simple Effect auch im Rahmen von Hypothesen herangezogen werden, die explizit nur den Einfluss der exogenen auf die endogene Variable testen wollen.

Aus diesem Grund kann das von Echambadi et al. (2006) kritisierte Prinzip der residualen Zentrierung nach Becker et al. (2018) auch zu Problemen der Interpretation führen, wenn der Orthogonalisierungsansatz verwendet wird. Hierbei entsteht nach den Autoren durch die Orthogonalität des Interaktionsterms kein Simple Effect. Hair et al. (2017, S. 258) sind hierbei wie bereits erwähnt der Ansicht, dass dies zu keinen Problemen in der Interpretation führt, da der Fokus nach den Autoren nicht auf dem Simple Effect sondern auf dem Main Effect liegen sollte, wenn Forscher den konstanten Effekt der exogenen auf die endogene Variable betrachten möchten. Die Autoren bringen demnach zum Ausdruck, dass der ME und Main Effect aus demselben Modell entnommen

werden sollen, während Aguinis et al. (2017) und Becker et al. (2018) hierbei das Problem sehen, dass die Betrachtung des Main Effects, als wäre er ein Simple Effect, zu inkorrekten Interpretationen führen kann.

Der zweite Aspekt der Ergebnisinterpretation umfasst die Berechnung und das Berichten der Effektgröße f^2 , die aufzeigt, wie sehr die MV zur Erklärung des endogenen Konstruktes beiträgt (Hair et al., 2017, S. 256). Sie ist nach Carte und Russell (2003) einer Einschätzung der Größe des ME anhand des Pfadkoeffizienten p_3 vorzuziehen. Dabei berechnet sich f^2 nach Henseler und Fassott (2010) aus der Änderung der erklärten Varianz R^2 im endogenen Konstrukt, wenn der Interaktionsterm im Modell enthalten ist (R^2_{included}), entgegengesetzt dem Fall, dass er nicht enthalten ist (R^2_{excluded}). Diesbezüglich schlagen die Autoren folgende Formel zur Berechnung der Effektgröße f^2 vor:

$$f^2 = \frac{R^2_{\text{included}} - R^2_{\text{excluded}}}{1 - R^2_{\text{included}}} \quad (5)$$

Hierbei wird f^2 in der Regel nach Cohen (1988, S. 412-414) bei $f^2=0,02$ als schwach, bei $f^2=0,15$ als moderat und bei $f^2=0,35$ als hoch eingeschätzt (Henseler & Fassott, 2010). Diese Einschätzung von Effektgrößen ist nach Aguinis et al. (2005) jedoch kontextabhängig und sollte von jedem Forscher selbst eingestuft werden. Aus diesem Grund hat Kenny (2018) auf Basis eines Reviews von Effektgrößen und statistischer Power in MA kategorischer Variablen im Kontext von multipler Regression, das von Aguinis et al. (2005) stammt und in dem eine mittlere Effektgröße bei ME von nur 0,009 gefunden wurde, die folgende realistischere Einschätzung von f^2 -Werten vorgeschlagen: $f^2=0,005$ sei schwach, $f^2=0,01$ sei moderat und $f^2=0,025$ sei hoch (Hair et al., 2017, S. 256). Dabei kann auch ein schwacher Effekt bspw. unter der Bedingung stark durch den ME beeinflusster Pfadkoeffizienten eine große Bedeutung haben und ist infolgedessen nicht direkt aus der Betrachtung auszuschließen (Chin et al., 2003; Aguinis et al., 2005). Aus diesem Grund sollte die Diskussion der Implikationen der MA für Manager und Forscher auch bereits auf Basis eines schwachen signifikanten Effektes nach Kenny (2018) erfolgen und nicht bei einer Unterschreitung eines schwachen Effektes nach Cohen (1988, S. 412-414) unterlassen werden, was bspw. von Henseler et al. (2012) empfohlen wird.

Um bei der Interpretation die Richtung und Stärke des ME korrekt einschätzen zu können, sollten die Resultate der MA in Form eines Simple Slope Plots dargestellt werden (Memon et al., 2019). Dabei werden die Steigungen der Kurven (Slopes) grafisch miteinander verglichen, welche die Beziehung der endogenen und exogenen Variable für verschiedene Niveaus oder Kategorien der MV darstellen (Field, 2018, S. 484-485). Der Pfadkoeffizient p_3 , also die Beziehung zwischen Interaktionsterm und endogenem Konstrukt, gibt hier außerdem Auskunft über die Richtung des ME. Dabei zieht ein positiver Pfadkoeffizient einen positiven ME nach sich und vice versa.

Abbildung 8 zeigt ein Beispiel eines positiven ME. Hierbei

stellt man nach Hair et al. (2017, S. 258) die Beziehung des exogenen und endogenen Konstruktes für verschiedene Niveaus der MV dar, wobei der Moderator im Fall des hohen Niveaus meistens auf einer Standardabweichung über bzw. im Fall des niedrigen Niveaus unter seinem Mittelwert betrachtet wird. Ist der Moderator auf seinem Mittelwert, also auf einem mittleren Niveau, entspricht die Steigung der abgebildeten Kurve dem Simple Effect p_1 , was in Abbildung 8 mithilfe der durchgezogenen Linie dargestellt ist. Es ist weiterhin zu erkennen, dass die gestrichelte Kurve, welche die Beziehung der exogenen und endogenen latenten Variable bei einem hohen Niveau der MV darstellt, steiler als die durchgezogene Kurve verläuft. Demnach ist die Beziehung auf dem hohen Niveau der MV stärker. Dies ist damit zu begründen, dass die Steigung der Kurve nun p_1+p_3 entspricht (Hair et al., 2017, S. 269). Gegenteiliges kann bei einem niedrigen Niveau (gepunktete Linie) der MV beobachtet werden, bei der die Beziehung zwischen der exogenen und endogenen Variable weniger stark ausgeprägt ist. In diesem Fall liegt wie bereits erwähnt ein positiver ME vor. Falls sich die Steigungen der Kurven bei verschiedenen Niveaus oder Kategorien der MV kaum bzw. gar nicht unterscheiden und die Kurven (eher) parallel verlaufen, deutet dies auf die Abwesenheit eines signifikanten ME hin (Field, 2018, S. 485). Letztlich ist bei einem negativen ME die Kurve steiler, welche die Beziehung der exogenen und endogenen Variable auf dem niedrigen Niveau der MV darstellt, als die Kurve auf dem Mittelwert der MV (Hair et al., 2017, S. 269). Wird der Moderator hierbei kleiner (größer), resultiert ein stärkerer (schwächerer) Zusammenhang der exogenen und endogenen Variable.

Wie eingangs erwähnt, sollte die Interpretation der Ergebnisse der MA und die Schlussfolgerung von Implikationen für die Forschung und/oder das Management stets auf die eingangs aufgestellte theoretische Grundlage zurückkommen, d.h. die Bedeutung der Ergebnisse im Kontext des theoretischen Hintergrunds in den Vordergrund stellen und analysieren, anstatt sich kontextlos auf die Signifikanz des ME zu fokussieren (Memon et al., 2019).

2.3.5. Herausforderungen bei Moderationsanalysen in PLS-SEM

Potenzielle Fehlerquellen

Bis hierhin sind bereits einige Fehlerquellen bei der Durchführung und Auswertung von MA herausgestellt worden, die nun zusammengefasst und um weitere Herausforderungen bei MA ergänzt werden sollen, was ebenfalls in Tabelle 2 dargestellt ist. Diese beinhaltet insg. neun potenzielle Quellen für Fehler, die kurz beschrieben werden und für die ggf. eine mögliche Lösung bzw. eine empfohlene Vorgehensweise zur Vermeidung der Fehler dargestellt wird.

Zunächst kann die Wahl des Ansatzes zur Generierung des Interaktionsterms dazu führen, dass inhaltliche Fehler entstehen, wie etwa durch die Nutzung des Produkt-Indikator-Ansatzes oder Orthogonalisierungsansatzes bei formativen Moderatorstrukturen bzw. exogenen Konstrukten. Der resultierende ME und die Interpretation sowie die aus der MA geschlossenen Implikationen basieren nicht auf

einer inhaltlich korrekten Bildung des Interaktionsterms, da den zugrundeliegenden formativ gemessenen Konstrukten keine eindeutige inhaltliche Domäne zugeordnet werden kann. (Hair et al., 2017, S. 249). Die Domäne des Interaktionsterms aus dem Produkt aus mindestens einem formativ gemessenem Konstrukt wäre folglich ebenfalls nicht bestimmbar und die inhaltliche Bedeutung des Interaktionsterms ist unklar. Deshalb darf bei einer formativ gemessenen latenten MV oder exogenen Variable lediglich der Zwei-Stufen-Ansatz gewählt werden.

Aber auch bei reflektiv gemessenen Konstrukten sind die Ansätze zur Bildung des Interaktionsterms nach einer Simulationsstudie von Henseler und Chin (2010) nicht gleichwertig zu behandeln, da sie abhängig vom Ziel der MA eine unterschiedliche Leistung erzielen. Wie eingangs erwähnt sorgt die Auswahlmöglichkeit aus den verschiedenen Ansätzen für Verwirrung bei Forschern (Hair et al., 2019). Welcher Ansatz unter welchen Bedingungen gewählt werden sollte, wird im nächsten Abschnitt der vorliegenden Arbeit näher beschrieben.

Der zu wählende Ansatz hängt in diesem Fall vordergründig von dem Ziel der Forscher ab, das sie bei der Durchführung von MA mit PLS-SEM verfolgen. Ein hierbei entstehendes Problem könnte die Unsicherheit von Forschern über ihr Ziel darstellen. In erster Linie werden im PLS-SEM-Kontext Effekte auf ihre Signifikanz getestet, was demnach das häufigste Ziel von Forschern darstellt und weshalb sich in den meisten Fällen für die Verwendung des Ansatzes mit der größten statistischen Power entschieden werden sollte. Oftmals wird jedoch die Nutzung von PLS-SEM damit begründet, dass das Hauptziel der Analyse die Vorhersage von Effekten sein soll. Daraus folgt, dass Forscher Probleme damit haben könnten, ihr Ziel korrekt einzuschätzen und infolgedessen einen nachteiligen Ansatz wählen.

Eine weitere Fehlerquelle stellt nach Becker et al. (2018) die Möglichkeit dar, bei der Anwendung des Orthogonalisierungsansatzes und Produkt-Indikator-Ansatzes über die Datenbehandlung der Indikatordaten vor der Generierung des Interaktionsterms zu entscheiden. Dabei besteht die Wahl zwischen Standardisierung, Mittelwertzentrierung oder keiner Behandlung der Daten (d.h. Verwendung unstandardisierter Daten). In ihrer Simulationsstudie haben die Autoren herausgefunden, dass alle Ansätze zur Generierung des Interaktionsterms bei der Genauigkeit der Schätzung der ME am besten mit der Verwendung von standardisierten Indikatordaten funktionieren. Dies erleichtert wie bereits dargelegt auch die Interpretation des Simple Effects (Hair et al., 2017, S. 257). Aus diesem Grund sollte von Mittelwertzentrierung oder gar keiner Behandlung der Daten vor der MA abgesehen und die Daten stattdessen unabhängig vom verwendeten Ansatz standardisiert werden, um Verzerrungen der Schätzung des Interaktionseffektes zu vermeiden (Becker et al., 2018).

Nach der Schätzung des Pfadmodells in PLS-SEM folgt die Evaluation des Messmodells. Dabei ist wie bereits dargelegt die MV genauso wie die exogene und endogene Variable zu evaluieren, wogegen der Interaktionsterm nach Hair et al. (2017, S. 255) kein Gegenstand der Evaluation des Mess-

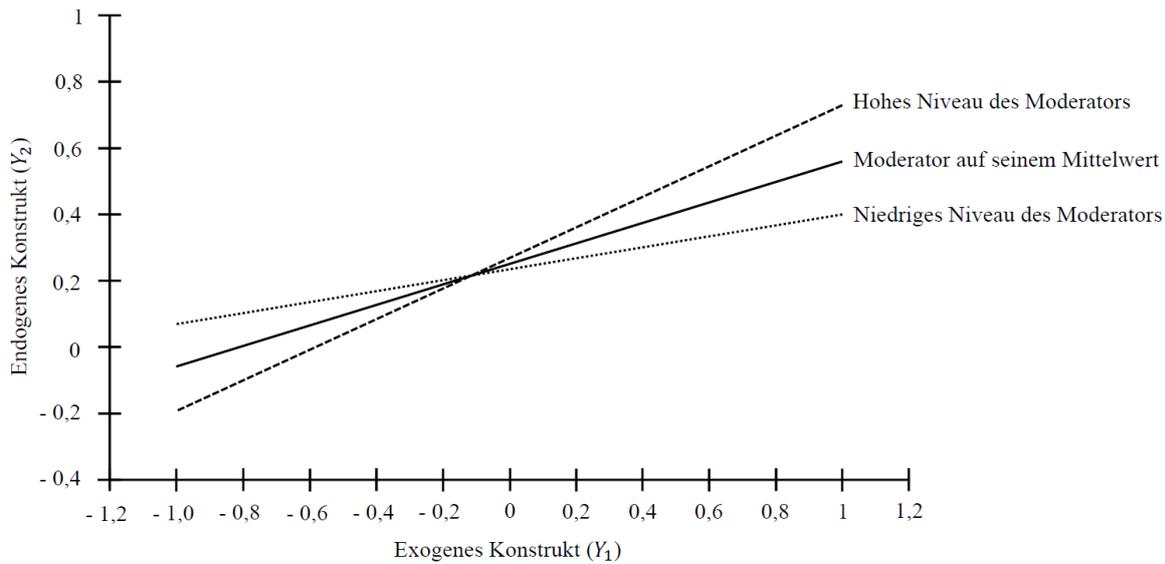


Abbildung 8: Beispielhafte Darstellung eines Simple Slope Plots (in Anlehnung an Hair et al., 2017, S. 259)

Tabelle 2: Potenzielle Fehlerquellen bei MA mit PLS-SEM mit Empfehlungen zur Problemvermeidung (in Anlehnung an Aguinis et al., 2017; Becker et al., 2018; Chin et al., 2003; Frazier et al., 2004; Hair et al., 2017, S. 249-258; Henseler & Chin, 2010; Henseler & Fassott, 2010; Li et al., 2019)

Fehlerquelle	Beschreibung der Fehlerquelle	Empfehlungen
Wahl des Ansatzes zur Generierung des Interaktionsterms (1)	nur Zwei-Stufen-Ansatz möglich bei formativ gemessenen Konstrukten	Verwendung des Zwei-Stufen-Ansatzes bei formativ gemessenen Konstrukten
Wahl des Ansatzes zur Generierung des Interaktionsterms (2)	Ansätze weisen verschiedene Leistung in Abhängigkeit vom Ziel der MA auf	siehe Framework in Abschnitt 2.4 für Umgang mit reflektiv gemessenen Konstrukten
Wahl des Ansatzes zur Generierung des Interaktionsterms (3)	Ziel der MA wird falsch von Forschern eingeschätzt	Ziel der MA ist im Zweifel das Herausstellen der statistischen Signifikanz
Datenbehandlung vor der Generierung des Interaktionsterms	Wahl zwischen Standardisierung, Mittelwertzentrierung und Verwendung unstandardisierter Daten	Standardisierung der Daten vor der Generierung des Interaktionsterms
Evaluation des Messmodells des Interaktionsterms	unterschiedliche Ansichten dazu, ob Evaluation (insb. der Reliabilität) erfolgen sollte	zumindest Angabe der individuellen Reliabilitäten des exogenen und moderierenden Konstrukts
Interpretation des Main Effects und Simple Effects	unterschiedliche Ansichten dazu, ob Main Effect und Simple Effect getrennt werden sollten	-
Signifikanztest des ME	Testen der MV auf Signifikanz	Testen des Interaktionsterms auf Signifikanz
Interpretation der Effektgröße f^2	keine Bestimmung von f^2 oder Interpretation nach Cohen (1988, S. 412-414)	Bestimmung von f^2 und Interpretation nach Kenny (2018)
Künstliche Dichotomisierung kontinuierlicher MV	künstliche Dichotomisierung sorgt für Verlust von Varianz in MV und geringerer Power	kontinuierliche MV nicht künstlich dichotomisieren

modells sein sollte. Insofern die Betrachtung von Messfehlern eine übergeordnete Rolle für Forscher spielt, sollten diese jedoch (zumindest) im Rahmen der Betrachtung der internen Konsistenzreliabilität nach Aguinis et al. (2017) auch für den Interaktionsterm berechnet werden. Forscher müs-

sen sich deshalb entscheiden, ob sie dem inhaltlich korrekten Weg nach Hair et al. (2017, S. 255) folgen und Messfehler des Interaktionsterms ignorieren oder der empirisch begründeten Vorgehensweise nach Aguinis et al. (2017) nachgehen und Messfehler mit in ihre Betrachtungen einbeziehen. Nach

allen genannten Autoren sollten aber in jedem Fall die individuellen Messfehler der exogenen, endogenen und moderierenden Variable angegeben werden.

In ähnlicher Art und Weise können auch die verschiedenen Sichtweisen von Hair et al. (2017, S. 257-258), Chin et al. (2003) und Henseler und Fassott (2010) auf der einen Seite und Aguinis et al. (2017) und Li et al. (2019) auf der anderen Seite dazu führen, dass Unsicherheiten im Umgang mit den Main Effects und Simple Effects entstehen. Erstgenannte Autoren befürworten eine hierarchische Vorgehensweise bei der Analyse der Effekte, wobei jeweils ein Modell mit MV zur Berechnung des Main Effects und ein Modell ohne MV zur Berechnung des Simple Effects geschätzt werden soll. Somit werden die Effekte strikt voneinander getrennt und eine Interpretation des Main Effects bspw. zum Testen einer Hypothese des linearen Effektes der exogenen auf die endogene Variable ist unabhängig vom ME möglich. Dies ist nach den zuletzt genannten Autoren allerdings inhaltlich nicht sinnvoll, da ein signifikanter ME bedeutet, dass der Main Effect nicht existiert und nur der Simple Effect eine Bedeutung hat. Erneut könnten die verschiedenen Sichtweisen hierbei für Unsicherheit bei Forschern sorgen.

Darüber hinaus besteht die Möglichkeit, dass Forscher nicht den Interaktionseffekt als Grundlage für die Interpretation eines signifikanten ME heranziehen, sondern den Pfadkoeffizienten (p_2) von der MV zu der endogenen Variable im Modell auf Signifikanz prüfen. Dies würde dazu führen, dass der Effekt der MV auf das endogene Konstrukt getestet wird und die MV dadurch wie eine exogene Variable behandelt werden würde. Weiterhin besteht hierbei auch die Möglichkeit der Verwechslung von Moderation und Mediation, wobei fälschlicherweise der direkte Effekt der MV auf das exogene Konstrukt (oder vice versa) betrachtet und somit auch kein ME untersucht werden würde. Demzufolge wird hierbei nicht die Moderation der Beziehung einer exogenen und endogenen latenten Variable analysiert, die folglich nur durch den zusätzlichen Interaktionsterm im Modell und seinen Effekt auf die endogene Variable modelliert werden kann. Daher ist stets der Interaktionsterm auf Signifikanz zu prüfen, um auf einen signifikanten ME schließen zu können (Henseler & Fassott, 2010).

Wurde die Signifikanz eines ME festgestellt, sollte dieser auf seine Effektgröße f^2 geprüft werden. Eine erste potenzielle Fehlerquelle besteht hierbei darin, die Effektgröße nicht zu bestimmen. Dies stellt eine wahrscheinliche Quelle für Fehler dar, da die Berechnung bisher nur in wenigen Softwarepaketen implementiert und standardmäßig bei MA durchgeführt wird (Memon et al., 2019). Ein zweiter potenzieller Fehler wäre hierbei der verfrühte Ausschluss von zu schwachen signifikanten Effekten nach Cohen (1988, S. 412-414) aus der MA und der anschließenden Diskussion der Implikationen des ME. Henseler et al. (2012) empfehlen diese Vorgehensweise, sie sollte jedoch dem Vorschlag von Kenny (2018) angepasst werden. Demnach wird empfohlen, die Untergrenze für ME von $f^2 = 0,005$ heranzuziehen und nur bei einer Unterschreitung anzunehmen, dass der ME zu schwach ist, um Implikationen nach sich zu ziehen (Hair et al., 2017,

S. 256).

Schließlich weisen Aguinis et al. (2017) neben den bisher besprochenen Fehlerquellen bei MA u.a. darauf hin, dass eine künstliche Dichotomisierung kontinuierlicher Variablen zu einer Reduktion der statistischen Power führt und die Größe des ME weiter verkleinert, da hierbei die Varianz der MV reduziert wird. Dies sollte demnach auch im PLS-SEM-Kontext bei MA vermieden werden und kontinuierliche Variablen sollten unverändert in die MA einbezogen werden (Henseler & Fassott, 2010).

Moderierte Mediation und Hierarchical Component Models

Im Kontext von MA in PLS-SEM stellt letztlich die Einbindung eines Moderators in eine Mediation eine besondere Herausforderung bei der Berechnung der Effekte und Interpretation dieser dar. Im Fall einer Mediation wirkt nach Hair et al. (2017, S. 247) eine dritte Variable in ähnlicher Art und Weise wie ein Moderator auf die Stärke der Beziehung zwischen zwei Variablen. Hierbei wird jedoch untersucht, inwiefern eine Änderung im exogenen Konstrukt zu einer Änderung im Mediatorkonstrukt führt, was nach den Autoren durch einen direkten Effekt zwischen diesen beiden Variablen modelliert wird. Moderierte Mediationsmodelle (Conditional Process Models) stellen die Kombination von Moderation und Mediation dar, wobei ein Moderator auf mindestens eine der direkten Beziehungen im Mediationsmodell wirkt und dieses damit abhängig von der Ausprägung des Moderators wird (Sarstedt, Hair Jr et al., 2020). Abbildung 9 stellt ein Beispiel eines Conditional Process Models dar, wobei die MV (M) einen Einfluss auf die Beziehung der exogenen Variable (Y_1) und der Mediatorvariable (Y_2) hat.

Die Berechnung und Interpretation des ME insb. im Zusammenhang mit der Mediation hängt hierbei allerdings von Faktoren, wie der Anzahl der Pfade, auf welche die MV wirkt, oder der Position des Pfades im Mediationsmodell, den die MV beeinflusst, ab (Hair et al., 2017, S. 261). Aus diesem Grund übersteigt die detaillierte Betrachtung von Conditional Process Models den Umfang der vorliegenden Arbeit.

Dies gilt ebenfalls für MA im Zusammenhang mit Hierarchical Component Models (HCM), welche Konstrukte umfassen, die nach Hair et al. (2017, S. 281-283) in mehreren Hierarchiestufen gemessen werden. Hierbei bilden nach den Autoren konkretere Bestandteile eines Konstruktes die Ebene erster Ordnung und das abstrakte Konstrukt die Ebene zweiter Ordnung. Weiterhin weisen die Autoren darauf hin, dass je nach Messmodell der Konstrukte erster oder zweiter Ordnung insg. vier verschiedene Arten von HCM entstehen können (d.h. reflektiv-reflektiv, reflektiv-formativ, formativ-reflektiv und formativ-formativ) und diese auf zwei verschiedene Weisen analysiert werden (d.h. mithilfe des Repeated Indicators Approach oder dem Zwei-Stufen-Ansatz). Becker, Klein und Wetzels (2012) gehen in diesem Zusammenhang auf die Schwierigkeiten bei der Analyse von ME im Zusammenhang mit HCM ein. Auch in diesem Fall wird von der detaillierten Betrachtung von HCM bei MA abgesehen, da eine Diskussion der zahlreichen Möglichkeiten bei der Analyse und Interpretation von HCM bei MA den Umfang dieser Ar-

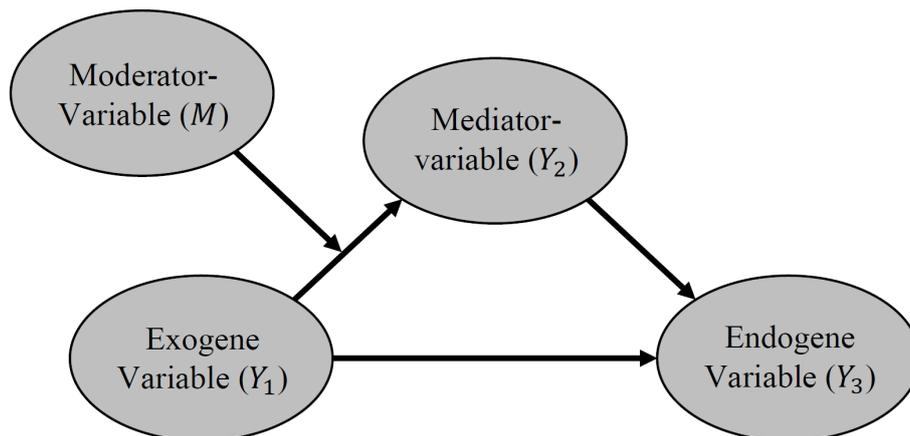


Abbildung 9: Beispielhafte Darstellung einer moderierten Mediation (in Anlehnung an Hair et al., 2017, S. 260)

beit übersteigt.

2.3.6. Framework zur Durchführung von Moderationsanalysen in PLS-SEM

Um die Unsicherheiten von Forschern bei der Berechnung und Interpretation von ME in PLS-SEM zu überwinden, wird nachfolgend ein Framework aufgestellt, das die Datenbehandlung vor der Generierung des Interaktionsterms, die Wahl des besten Ansatzes zur Generierung des Interaktionsterms und einen Teil der Vorgehensweise bei der Interpretation der ME und der Effektgröße f^2 zusammenfasst.

Damit das in Abbildung 10 dargestellte Framework zur Generierung des Interaktionsterms bei der Durchführung von MA verwenden werden kann, müssen nachfolgende Voraussetzungen erfüllt sein. Zunächst muss PLS-SEM zur Schätzung des Pfadmodells verwendet werden. Weiterhin sollte die Durchführung einer einfachen MA mit der Bildung eines Interaktionsterms zur Modellierung des ME erfolgen, wodurch Multigruppenanalysen zur Schätzung von ME ausgeschlossen werden. Außerdem kann die Verwendung des Frameworks, wie im vorigen Abschnitt dargelegt, nur mit Konstrukten erfolgen, die nicht Bestandteil eines HCM sind. Letztlich ist die Verwendung des Frameworks auf kontinuierliche latente bzw. manifeste MV (quasi-metrische, intervall- und verhältnisskalierte) limitiert, da die Variablen in den dem Framework zugrundeliegenden Simulationsstudien ebenfalls in jedem Fall kontinuierlich waren³.

Wie in Abbildung 10 zu erkennen ist, sollten die Indikatoren der latenten MV (M) sowie des exogenen (Y_1) und endogenen Konstruktes (Y_2) vor der Generierung des Interaktionsterms standardisiert werden. Hierbei wird nach Henseler und Fassott (2010) sichergestellt, dass die Variablen untereinander in ihrer Bedeutung vergleichbar sind, was die anschließende inhaltliche Interpretation vereinfacht. Außerdem wird die bei dem Zwei-Stufen-Ansatz und Produkt-Indikator-Ansatz hervorgerufene Multikollinearität im Pfadmodell durch die Standardisierung verringert. Schließlich

deuten auch die Erkenntnisse aus der Simulationsstudie von Becker et al. (2018), darauf hin, dass vor der Generierung des Interaktionsterms standardisierte Daten, im Gegensatz zu Daten, welche mittelwertzentriert oder unstandardisiert belassen wurden, die genauesten Schätzungen der Pfadkoeffizienten des Moderatormodells hervorbringen. Die Autoren verweisen aber ebenfalls darauf, dass der Zwei-Stufen-Ansatz die LVS, aus denen der Interaktionsterm gebildet wird, standardmäßig standardisiert, weshalb bei der Verwendung dieses Ansatzes nicht zwingend auf eine Standardisierung vor der Analyse geachtet werden muss.

Um im Anschluss an die Standardisierung der Indikatorindaten den passenden Ansatz für das vorliegende Modell identifizieren zu können, sollte nach Henseler und Fassott (2010) zunächst zwischen den Messmodellen der exogenen bzw. moderierenden Variable unterschieden werden. Deshalb folgt auch im Framework im Anschluss an die Standardisierung der Indikatoren aller Konstrukte in der MA die Definition des Messmodells der exogenen Variable und der MV. Kann die Aussage, dass M und Y_1 beide reflektiv gemessen seien, bestätigt werden, sollte dem linken Ast gefolgt werden. Dies wird durch das Häkchen im Framework visualisiert. Henseler und Chin (2010) haben in ihrer Simulationsstudie mit ausschließlich reflektiv gemessenen Konstrukten alle gängigen Ansätze zur Generierung des Interaktionsterms untersucht, um den besten Ansatz für verschiedene Ziele von Forschern bei der Durchführung von MA herausstellen zu können⁴. Infolgedessen schließt sich an die Definition des Messmodells die Bestimmung des Ziels der MA an, wenn lediglich reflektiv gemessene exogene latente Variablen und MV identifiziert worden sind. Hierbei unterscheiden Henseler und Chin (2010) zwischen den folgenden Zielen, nach denen folglich auch im Framework differenziert wird: Enthüllung der statistischen Signifikanz (Maximierung der statistischen Power), Minimierung des Schätzungsfehlers des ME (Maximierung der Punktgenauigkeit) und Maximierung der

³Für eine Verwendung kategorischer Variablen sei auf die Ausführungen von Henseler und Fassott (2010) verwiesen.

⁴Wie Becker et al. (2018) haben auch Henseler und Chin (2010) in ihrer Studie standardisierte Indikatoren-Daten verwendet.

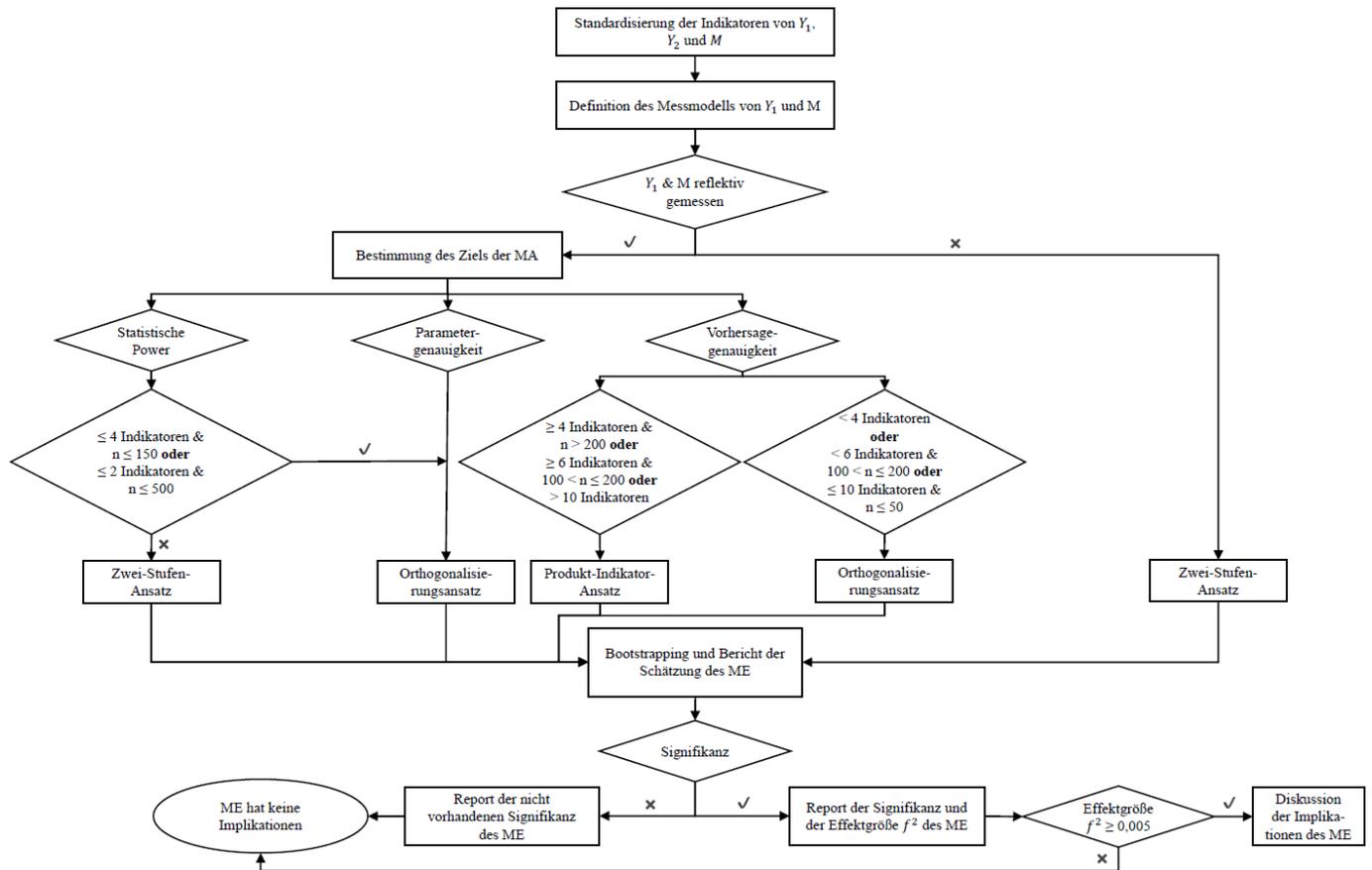


Abbildung 10: Framework zur Generierung des Interaktionsterms und der Interpretation von ME bei der Durchführung von MA in PLS-SEM (in Anlehnung an Becker et al. (2018); Henseler und Chin (2010); Henseler und Fassott (2010); Henseler et al. (2012); Kenny (2018))

Genauigkeit der Voraussage (Maximierung der Vorhersagegenauigkeit).

Bei der Betrachtung jedes Ziels wurde im Framework genauer als bspw. von Hair et al. (2017, S. 252-255) zwischen den Ansätzen differenziert, da die Ansätze innerhalb der Ziele teilweise unterschiedliche Leistungen in Abhängigkeit von Modellspezifikationen hervorbringen. Insbesondere weisen der Stichprobenumfang (n) und die jeweilige Anzahl der Indikatoren in den Konstrukten einen Einfluss auf die Leistung der Ansätze auf. So sollte nach Henseler und Chin (2010) zur Maximierung der statistischen Power bei einer gegebenen Effektgröße in der Population (= 0,136) und konstantem Signifikanzkriterium (= 0,05) bei sehr wenigen Indikatoren innerhalb der Konstrukte der Orthogonalisierungsansatz verwendet werden. Dabei dominiert dieser nach den Autoren bei sehr kleinen Stichprobenumfängen (n ≤ 150) und weniger als fünf Indikatoren in den Konstrukten die anderen Ansätze. Sollten sogar weniger als zwei Indikatoren in allen Konstrukten auftreten und der Stichprobenumfang übersteigt nicht 500, ist der Orthogonalisierungsansatz noch immer vorzuziehen. In den meisten Fällen wird nach den Autoren jedoch durch den Einsatz des Zwei-Stufen-Ansatzes die größte Power erzielt. Insofern die beschriebenen Gren-

zen nicht eingehalten werden und demnach eine mittlere bis große Anzahl von Indikatoren vorliegt oder der Stichprobenumfang groß genug ist, dominiert der Zwei-Stufen-Ansatz nach den Autoren demnach die anderen Ansätze.

Liegt der Fokus der Forscher jedoch auf der möglichst genauen Schätzung des ME und soll demnach die Punktgenauigkeit der Parameterschätzungen maximiert werden, bringt der Orthogonalisierungsansatz nach den Autoren stets die genauesten Schätzungen hervor und sollte in diesem Fall genutzt werden.

Letztlich ist für eine möglichst präzise Vorhersage des ME nach den Autoren entweder der Produkt-Indikator-Ansatz oder der Orthogonalisierungsansatz zur Generierung des Interaktionsterms zu verwenden. Hierbei kann erneut in Abhängigkeit vom Stichprobenumfang und von den Indikatoren der Konstrukte unterschieden werden. Wie die Autoren zeigen, dominiert der Produkt-Indikator-Ansatz im Fall vieler Indikatoren in den Konstrukten (> 10) bzw. auch bei weniger Indikatoren und einem dafür größeren Stichprobenumfang (bspw. ≥ 4 Indikatoren und n > 200). Der Orthogonalisierungsansatz hingegen erzielt nach den Autoren genauere Vorhersagen des ME bei wenigen Indikatoren der Konstrukte (< 4) oder bei einem kleinen bis mittelgroßen Stichproben-

umfang und einer mittleren Anzahl von Indikatoren (bspw. ≤ 10 Indikatoren und $n \leq 50$).

Wird mindestens eins der Konstrukte formativ gemessen und kann die Aussage, dass M und Y_1 reflektiv gemessen seien, demnach nicht bestätigt werden, sollte dem rechten Pfad nach der Definition der Messmodelle im Framework gefolgt werden. Dies ist anhand des Kreuzes oberhalb des rechten Pfades in Abbildung 10 erkennbar. Hieraus kann geschlussfolgert werden, dass bei der formativen Messung der latenten MV oder der latenten exogenen Variable stets der Zwei-Stufen-Ansatz verwendet werden sollte. Nur mit diesem Ansatz ist die inhaltliche Korrektheit des Interaktionsterms bei der Multiplikation formativer Konstrukte, denen keine eindeutige inhaltliche Domäne zugrunde liegt, gewährleistet und der Interaktionsterm konzeptionell abbildbar (Henseler & Fassott, 2010). Darüber hinaus haben auch Henseler et al. (2012) in ihrer Simulationsstudie die Überlegenheit des Zwei-Stufen-Ansatzes über dem Produkt-Indikator-Ansatz und dem Orthogonalisierungsansatz bei MA mit formativ gemessenen Konstrukten im Hinblick auf die wie bereits von Henseler und Chin (2010) getesteten Aspekte festgestellt⁵.

In jedem Fall sollte nach Henseler et al. (2012) die Prüfung auf Signifikanz des ME mithilfe des Bootstrapping-Prozederes erfolgen, wobei neben der Signifikanz auch die Schätzung der Höhe des Pfadkoeffizienten p_3 (d.h. Parameterschätzung des Interaktionseffektes) angegeben werden sollte. Im Anschluss an die Signifikanzprüfung sollte nach den Autoren die Größe des ME (f^2) bestimmt und berichtet werden, insofern der ME signifikant ist. Im Anschluss ist die Effektgröße nach den Autoren darauf zu kontrollieren, dass diese mindestens als klein angesehen werden kann. Hierfür wird im Framework die Untergrenze von Kenny (2018) herangezogen, ab welcher kleine Effekte, die nach Chin et al. (2003) ebenfalls bedeutsam sein können und Implikationen nach sich ziehen, mindestens 0,005 betragen müssen. Bei der Diskussion der Implikationen ist die theoretische Grundlage der Hypothese des ME erneut aufzugreifen (Andersson et al., 2014). Liegt kein signifikanter ME vor, sollte dieser dennoch berichtet werden (Henseler et al., 2012). Wie auch in dem Fall, dass der Effekt zwar signifikant, aber nicht mindestens als klein in seiner Größe einzustufen ist, zieht der ME dabei keine Implikationen nach sich.

3. Methode

Um die gestellte Forschungsfrage im Rahmen eines Reviews relevanter Untersuchungen in Forschungsartikeln beantworten zu können, thematisiert dieser Abschnitt die

⁵Die Autoren weisen darauf hin, dass der ebenfalls bei formativen Konstrukten anwendbare hybride Ansatz den Zwei-Stufen-Ansatz in allen Kategorien leistungsmäßig übertrifft, dieser aufgrund fehlender Software-Implementierung aber nur wenigen Forschern zur Verfügung steht. Sollte er jedoch verfügbar sein, empfehlen die Autoren, den hybriden Ansatz zur Generierung des Interaktionsterms bei formativen Konstrukten zu verwenden.

Durchführung der systematischen Literaturrecherche. Zunächst wird die Festlegung relevanter Inklusionskriterien beschrieben. Nachfolgend wird der Ablauf des Rechercheprozesses, dessen Ergebnisse und die Vorgehensweise bei der Kodierung der gefundenen Artikel dargelegt.

3.1. Inklusionskriterien

Vor der Literaturrecherche wurden Inklusionskriterien festgelegt, die jede Untersuchung erfüllen musste, um ein Bestandteil des Reviews werden zu können. Ein wichtiges Einschlusskriterium bei der Auswahl zu verwendender Untersuchungen war deren Analyse mindestens eines ME. Insbesondere muss dabei ein Interaktionsterm gebildet worden sein. Diese stellen den Fokus der vorliegenden Arbeit dar, weshalb MA mithilfe von Gruppenvergleichen (d.h. Multi-gruppenanalysen) ausgeschlossen wurden. Weiterhin musste zu erkennen sein, welcher Ansatz zur Berechnung von ME verwendet wurde, da anderenfalls keine klare Einschätzung der berechneten Modelle in den Untersuchungen mithilfe des Frameworks und der Auswirkungen der Wahl eines Ansatzes auf die Ergebnisse und Interpretation dieser möglich gewesen wäre. Aus diesem Grund wäre außerdem die Vergleichbarkeit mit den restlichen Artikeln eingeschränkt. Ist demnach in einer Publikation der Ansatz nicht klar zu erkennen, liefert sie keinen Mehrwert für das Review und kann nicht eingeschlossen werden.

Es wurden weiterhin nur empirische Untersuchungen eingeschlossen, die Ergebnisse durch die Verwendung des PLS-Algorithmus generiert haben, um sicherzustellen, dass die Bildung des Interaktionsterms den zuvor erläuterten Prinzipien folgt. Dies ermöglicht die Vergleichbarkeit der inkludierten Artikel, da alle auf demselben Algorithmus zur Berechnung des Pfadmodells basieren. Insbesondere müssen die Untersuchungen dabei aus Artikeln stammen, welche in einer vom Lehrstuhl für Marketing der Otto-von-Guericke-Universität Magdeburg zur Verfügung gestellten Datenbank enthalten sind. Diese setzt sich aus 239 empirischen Forschungsartikeln zusammen, die aus einem Suchprozess stammen, welcher die Vorgehensweise im Review von Hair, Sarstedt, Ringle und Mena (2012) für die Jahre 2011 bis 2020 repliziert. Hierbei wurden die 30 Top-Marketingjournale nach Hult, Reimann und Schilke (2009) auf die Schlagwörter „PLS“ und „Partial Least Squares“ in den Datenbanken EBSCO, Scopus und Web of Science sowie deren Websites in einer Volltextsuche durchsucht. Es folgten Filtermaßnahmen, die denen von Hair et al. (2012) nachempfunden wurden und aus denen 239 empirische Marketinganwendungen von PLS-SEM hervorgingen.

Gegenstand des vorliegenden Reviews konnten somit nur Artikel werden, welche empirische Marketinganwendungen von PLS-SEM enthalten und innerhalb der vergangenen zehn Jahre in Top-Marketingjournals veröffentlicht wurden. Dabei musste mindestens eine MA in einer Untersuchung enthalten sein, die mithilfe der Bildung eines Interaktionsterms durchgeführt wurde und klar erkennen lässt, mit welchem Ansatz die Berechnung des ME erfolgt ist.

3.2. Systematische Literaturrecherche

Um die Marketinganwendungen von PLS-SEM identifizieren zu können, welche Moderationen mithilfe von Interaktionstermen analysieren, wurde die aus 239 Publikationen bestehende zur Verfügung gestellte Datenbank zunächst mithilfe der Schlagwörter „moderation“, „moderator“, „moderate“, „moderating“, „interaction“, „interacting“ und „interaction effect“ durchsucht. Aus der Volltextsuche ergaben sich 109 Artikel, die eine beliebige Form der MA enthielten. Diese wurden anschließend darauf geprüft, ob ein Interaktionsterm zur Analyse der ME gebildet wurde. Insbesondere mussten deshalb zunächst 27 Artikel ausgeschlossen werden, die ME mithilfe von Multigruppenanalysen untersuchen. Weiterhin wurden neun Artikel ausgeschlossen, welche MA mit dem PROCESS-Makro von Hayes (2013) im Rahmen von Regressionsanalysen oder mit einer ANOVA durchführen. Demnach resultierten nach der ersten Filterung der Artikel mit MA insgesamt 73 Artikel, die einen Interaktionsterm zur Analyse der Moderatoren gebildet haben.

Von diesen mussten jedoch weitere 40 Publikationen vom Review ausgeschlossen werden, da sie keine Informationen zum verwendeten Ansatz bei der MA angaben. Neben den bereits diskutierten Problemen hierdurch bei der Vergleichbarkeit mit anderen Studien, ermöglicht diese intransparente Vorgehensweise einerseits keine Einschätzung der inhaltlichen Korrektheit der Bildung des Interaktionsterms im Fall formativ gemessener Konstrukte, da nicht nachvollzogen werden kann, ob der einzige inhaltlich korrekte Zwei-Stufen-Ansatz hierbei verwendet wurde. Andererseits ist die Interpretation des Main Effects und Simple Effects nicht gänzlich nachvollziehbar, da unklar ist, ob die Effekte wie bei der Nutzung des Orthogonalisierungsansatzes als äquivalent anzusehen sind oder nach Hair et al. (2017, S. 257-258) einer getrennten Betrachtung bei der Interpretation der MA benötigen. Aus diesem Grund ist die Angabe des Ansatzes, welcher zur Bildung des Interaktionsterms herangezogen wurde, von großer Bedeutung. Anderenfalls sind die potenziell falsch analysierten und folglich fehlinterpretierten Ergebnisse der MA nicht mit den Ergebnissen anderer Artikel vergleichbar und darüber hinaus müssen deshalb die geschlussfolgerten Implikationen für die Forschung oder das Management angezweifelt werden. Die nach der letzten Filterung verbleibenden 33 Artikel, die das Ergebnis der systematischen Literaturrecherche darstellen, wurden anschließend nach den verwendeten Ansätzen zur Generierung des Interaktionsterms sortiert.

3.3. Kodierung der Moderationseffekte

In den ausgewählten 33 wissenschaftlichen Arbeiten, die in das Review zu Moderation in Marketinganwendungen von PLS-SEM aufgenommen wurden, konnten 35 Untersuchungen gefunden werden, die MA mithilfe von Interaktionstermen durchführen. Diese 35 Untersuchungen wurden im nächsten Schritt in einer Excel-Tabelle kodiert, wodurch eine Vielzahl an Datensätzen entstand, deren Analyse die Grundlage des Reviews und damit der Beantwortung der

Forschungsfrage bildet (siehe den externen Anhang A1 für eine Auflistung der final inkludierten Artikel).

Das verwendete Kodierschema wurde im Hinblick auf das Ziel der vorliegenden Arbeit und den daraus folgenden Bestandteilen des anzufertigenden Reviews erstellt. Dabei fand zunächst eine Probekodierung von fünf Artikeln mit insgesamt sechs Untersuchungen statt, die Aufschluss darüber gaben, wie das Kodierschema strukturiert werden sollte und welche Eigenschaften der Artikel und der enthaltenen MA im Kodierschema bisher fehlten. Im Anschluss an die Festlegung der Struktur des Schemas und der Ergänzung fehlender Merkmale schloss sich die erneute Kodierung dieser fünf Artikel sowie die finale Kodierung der restlichen Untersuchungen an.

Die Struktur des Kodierschemas folgt dabei grob den Vorgaben zu Kodierschemata in Metaanalysen von M. (2014, S. 31-32). Der Autor verweist darauf, mehrere Hierarchiestufen im Kodierschema zu verwenden, um zwischen der übergeordneten Ebene der Artikel und untergeordneten Ebene der Untersuchungen unterscheiden zu können. Dies wurde in dem verwendeten Schema umgesetzt, indem ein Artikel mehrere Datensätze (gruppierte Zeilen) erhalten hat, insofern er mehrere Untersuchungen enthielt. Eine Untersuchung wiederum konnte ebenfalls mehrere Zeilen umfassen, falls sie mehr als einen ME analysiert hat. Somit konnte zwischen Artikeln, Untersuchungen und ME differenziert werden, wobei letztere schließlich jeweils eine Zeile in einer Untersuchung erhalten haben.

Jeder Datensatz wurde schließlich im Hinblick auf drei wesentliche inhaltliche Abschnitte kodiert: Allgemeine Eigenschaften des Artikels, Merkmale des einzelnen ME und Hinweise zu den MA in der Untersuchung. Im finalen Kodierprozess wurden dementsprechend im ersten Abschnitt der Tabelle für jeden Artikel u.a. allgemeine Eigenschaften, wie die Autoren, Erscheinungsjahr und das Journal kodiert. Hinzukamen außerdem die Anzahl der kodierten ME und der Stichprobenumfang der betrachteten Untersuchung. Die wichtigsten Merkmale der einzelnen ME im zweiten Abschnitt wurden für jeden ME getrennt erfasst, sodass diese einzeln im Hinblick auf das Framework analysiert werden konnten. Hierfür wurden für jeden ME insbesondere die Skalierung, die Anzahl der Indikatoren sowie das Messmodell der Moderatorvariable und der exogenen und endogenen Variable erfasst. Hinzukamen der verwendete Ansatz zur Bildung des Interaktionsterms und die Datenbehandlung vor der Analyse dieses ME (Standardisierung vs. Mittelwertzentrierung vs. Verwendung nicht standardisierter Daten). Neben einer Einschätzung, ob die Analyse des ME auf Basis der vorliegenden Daten das Framework eingehalten oder gegen es in mindestens einem Aspekt verstoßen hat, wurden außerdem die Effektgröße f^2 , das Ziel der MA und inwiefern das Messmodell der Moderatorvariable evaluiert wurde erfasst. Schließlich wurde im zweiten Abschnitt ebenfalls kodiert, ob der ME im Rahmen einer a priori aufgestellten Hypothese getestet wurde. Im letzten Abschnitt der Kodierung einer Untersuchung wurden Hinweise zu allen MA in einer Untersuchung aufgenommen, wobei festgehalten wur-

de, ob und in welcher Form nach Hayes (2013) eine MA im Zusammenhang mit einer Mediationsanalyse auftritt und inwieweit HCM Bestandteil einer MA sind. Schließlich wurde erfasst, ob der Main und Simple Effect bei der Interpretation der ME fälschlicherweise verwechselt wurden, wie bei der Analyse mehrerer MV vorgegangen wurde (gleichzeitige vs. separate Betrachtung der ME) und ob einer der Übersichtsartikel von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) oder Becker et al. (2018) zitiert wurde.

Schließlich konnten jeder Untersuchung maximal 16 Merkmale im ersten, 53 Merkmale pro inkludiertem ME im zweiten und 24 Merkmale im dritten Abschnitt des Kodierschemas zugeordnet werden, wobei fehlende Werte zumeist mit „NA“ ausgefüllt wurden. Zusätzlich wurden Kommentare zu jedem ME verfasst, welche die Auswertung der insg. 183 kodierten ME erleichtern sollten. Um die Güte der Kodierung sicherstellen zu können, wurden die Ergebnisse mehrfach auf ihre Korrektheit geprüft.

4. Review von Moderationsanalysen in Marketinganwendungen von PLS-SEM

Die Frage, wie Forscher mit den Herausforderungen bei der Durchführung von einfachen MA in PLS-SEM umgehen, soll in diesem Abschnitt beantwortet werden. Hierfür wurden 33 relevante wissenschaftliche Publikationen auf den Umgang mit MA in PLS-SEM im Rahmen eines Reviews untersucht. Zunächst werden allgemeine Beobachtungen aus der Literaturrecherche dargestellt, nach denen der Umgang von Forschern mit den zuvor identifizierten potenziellen Fehlerquellen bei MA, in der u.a. auch Einhaltungen oder Verstöße gegen das zuvor aufgestellte Framework analysiert werden, folgt. Schließlich folgt die Diskussion sonstiger Beobachtungen aus dem Reviewprozess, bei denen weitere Probleme und Fehlerquellen identifiziert wurden. Summa summarum dient das Review zur Identifikation noch unbekannter Fehler- und Problemquellen sowie der Überprüfung bestehender Unsicherheiten der Forschung im Umgang mit MA für den PLS-SEM-Kontext.

4.1. Beobachtungen während der Literaturrecherche

Bereits bevor die ersten Beobachtungen im Rahmen des Reviews gemacht wurden, brachte die Literaturrecherche einige Erkenntnisse hervor, welche in diesem Abschnitt aufgegriffen werden. Die erste Beobachtung aus der Literaturrecherche bezieht sich auf das Verhältnis der 109 Artikel, die eine beliebige Form einer MA (sowohl einfache MA als auch MA mithilfe von Gruppenvergleichen) durchgeführt haben, zu den insg. 239 Artikeln aus der Datenbank, die keine MA enthielten. Fast die Hälfte (45,61%) der 239 Artikel in der Datenbank enthalten eine Art der Moderation, was deren Relevanz für die aktuelle Forschung hervorhebt. Bereits im Jahr 2011 haben Henseler et al. (2012) die Aussage getroffen, dass sich das Interesse von Forschern in Richtung nicht-linearer Effekte, wie Interaktionseffekte, verschoben hat, da

lineare Effekte zwischen Konstrukten in vielen Fällen bereits identifiziert wurden. Diese Aussage wird von den vorliegenden Daten gestützt.

Darüber hinaus werden von diesen 109 Artikeln in insg. 73 Publikationen Interaktionsterme gebildet (66,97%), wobei demnach in 36 Fällen (33%) eine Multigruppenanalyse, Regressionsanalyse oder ANOVA durchgeführt wurde, um ME zu analysieren. Henseler und Chin (2010) weisen darauf hin, dass Multigruppenanalysen in PLS-SEM vorrangig bei kategorischen MV anzuwenden sind. Einfache MA hingegen können nach den Autoren sowohl mit kategorischen als auch mit kontinuierlichen MV bessere Ergebnisse als Multigruppenanalysen bei der Schätzung von ME in PLS-SEM erzielen. Aus diesen Gründen ist es naheliegend, dass auch in den meisten Fällen (66,97%) Interaktionsterme gebildet und damit die einfache MA vermehrt einer anderen Form der MA in PLS-SEM vorgezogen wurde.

In der Literaturrecherche wurden weiterhin insg. 40 Publikationen herausgefiltert, die zwar eine einfache MA mit der Bildung eines Interaktionsterms durchgeführt, aber nicht angegeben haben, mit welchem Ansatz der Interaktionsterm generiert wurde. Eine Begründung für diesen Ausschluss wurde bereits dargelegt. Zwischen 2011 und 2020 wurden somit 73 Artikel im PLS-SEM-Kontext veröffentlicht, von denen über die Hälfte (54,79%) keine Informationen enthalten, wie der Interaktionsterm gebildet wurde. Abbildung 11 zeigt die Verteilung der Artikel mit und ohne entsprechende Informationen.

Letztlich stellt das finale Ergebnis der Literaturrecherche eine weitere aufschlussreiche Beobachtung dar. Hierbei ist die Verteilung der Ansätze zur Generierung des Interaktionsterms so ausgeprägt, dass der Produkt-Indikator-Ansatz in 22 Artikeln vorkommt und dadurch mit 64,71% den überwiegenden Anteil der verwendeten Ansätze darstellt⁶. Dieser ist nach dem erstellten Framework der am seltensten optimale Ansatz, da er nur die beste Wahl für Forscher darstellen sollte, wenn diese die Vorhersagegenauigkeit des ME maximieren wollen und sich bspw. viele Indikatoren (> 10) in den verwendeten Konstrukten befinden. Es wird jedoch davon ausgegangen, dass das Finden und Testen signifikanter Effekte den eigentlichen Fokus der meisten PLS-SEM-Studien darstellt und die Vorhersage von Effekten ein recht spezielles Ziel ist, welches in der Praxis nur selten beobachtet wird. Dabei wird angenommen, dass das Vorhandensein von Hypothesen in einem Artikel darauf hindeutet, dass diese getestet werden sollen und dass Forscher hierfür den Ansatz mit der höchsten statistischen Power nutzen möchten.

In diesem Sinne ist das häufige Auftreten des Produkt-Indikator-Ansatzes kontraintuitiv und deutet bereits auf viele Verstöße gegen das aufgestellte Framework hin, die im kommenden Abschnitt genauer betrachtet werden. Dieser Ansatz

⁶Da in einem Artikel (Johnston, Khalil, Jain & Cheng, 2012) sowohl der Produkt-Indikator-Ansatz als auch der Zwei-Stufen-Ansatz verwendet wurden, wurde dieser beiden Kategorien zugeordnet und die Summe der Häufigkeiten der drei gefundenen Ansätze ergibt insg. 34, obwohl das Review nur 33 Artikel umfasst.

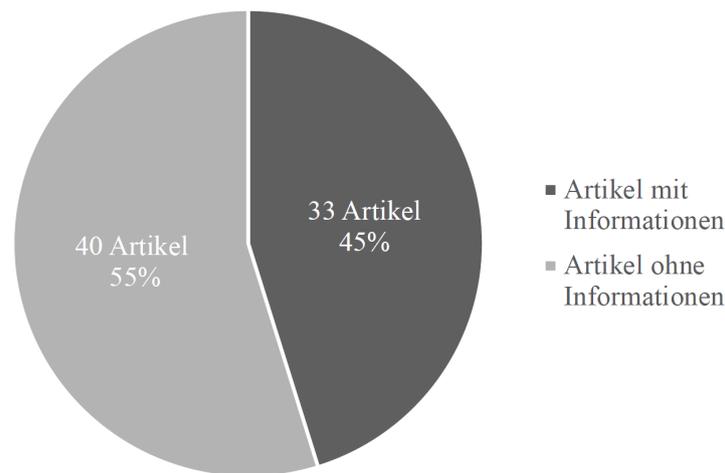


Abbildung 11: Artikel mit und ohne Informationen darüber, welcher Ansatz verwendet wurde

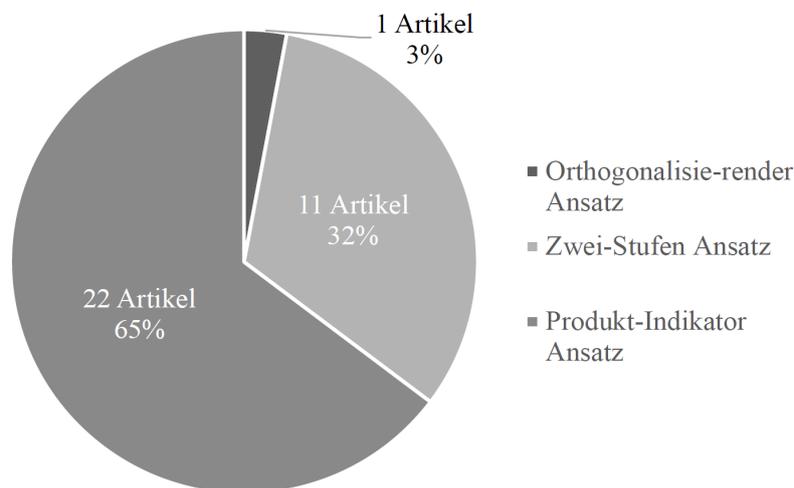


Abbildung 12: Verteilung der verwendeten Ansätze innerhalb der Artikel im Review

hätte erwartungsgemäß seltener gewählt werden müssen, wenn man davon ausgeht, dass Forscher nach den gängigen Vorgaben von Henseler und Chin (2010) den besten Ansatz für ihre MA auswählen würden. Wie im Framework gezeigt, bringen meist der Zwei-Stufen-Ansatz und ggf. der Orthogonalisierungsansatz in diesem Fall die größte statistische Power hervor, die jedoch mit elf (32,35%) innerhalb der Ergebnisse der Literaturrecherche die Minderheit der verwendeten Ansätze ausmachen. Die häufige Wahl des Produkt-Indikator-Ansatzes kann sich u.a. darauf zurückführen lassen, dass dieser den Prinzipien von MA im Regressionskontext entspringt, mit welchem multivariate Analysen insb. manifester Variablen i.d.R. durchgeführt werden (Hair et al., 2017, S. 2-3; Henseler & Fassott, 2010). Da sich die Vorgehensweise bei der Analyse von ME in Regressionsanalysen mithilfe des Produkt-Indikator-Ansatzes leicht auf PLS-SEM übertragen lassen, legt dies die Vermutung nahe, dass der Ansatz beliebt bei

Forschern ist, die zuvor im Regressionskontext ME analysiert haben. Hierbei werden infolgedessen best-practice Vorgaben zur Wahl des besten Ansatzes unter den Gegebenheiten der Messmodelle, Indikatoren in den Konstrukten und des Stichprobenumfangs nicht berücksichtigt.

Diese Erklärung wird ebenfalls von den Daten gestützt, die während des Reviews aus den 35 inkludierten Untersuchungen erfasst wurden. Hierbei wird in lediglich zehn der 35 Untersuchungen (29,41%) darauf hingewiesen, dass sich an den Empfehlungen zur Generierung von Interaktionstermen in PLS-SEM aus den Übersichtsartikeln von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) oder Becker et al. (2018) orientiert wurde. Somit weist die Mehrheit der Autoren der inkludierten Artikel (70,59%) keine Kenntnis der Vor- und Nachteile der einzelnen Ansätze zur Generierung des Interaktionsterms nach, was die unreflektierte Nutzung der Ansätze durch die Autoren erklären könnte.

4.2. Analyse der potenziellen Fehlerquellen bei Moderationsanalysen

Die in Abschnitt 2.3.5.1 herausgestellten potenziellen Fehlerquellen bei MA in PLS-SEM werden nachfolgend analysiert, um Rückschlüsse auf den aktuellen Stand der Forschung bezüglich des Umgangs mit diesen ziehen zu können. Dabei werden die Fehlerquellen inhaltlich gruppiert, sodass sich die Gebiete „Generierung des Interaktionsterms und Datenbehandlung vor der Bildung des Interaktionsterms“, „Interpretation des Moderationseffektes und Einschätzung der Effektgröße f^2 “ und „Künstliche Dichotomisierung kontinuierlicher Moderatorvariablen“ differenzieren lassen. Letzteres Gebiet besteht dabei nur aus der namensgebenden potenziellen Fehlerquelle, die keinem anderen Gebiet zugeordnet werden konnte.

4.2.1. Generierung des Interaktionsterms und Datenbehandlung vor der Analyse

Die Bildung des Interaktionsterms geht mit insg. vier potenziellen Fehlerquellen einher, welche den Einsatz falscher Ansätze bei formativ gemessenen Konstrukten, den Einsatz eines abhängig vom Ziel der Forscher suboptimalen Ansatzes bei einer gegebenen Anzahl von Indikatoren in reflektiv gemessenen Konstrukten und dem Stichprobenumfang sowie die falsche Einschätzung des Ziels der Forscher bei der MA umfassen. Hinzukommt die Datenbehandlung vor der Generierung des Interaktionsterms. Dabei werden die genannten Problembereiche u.a. im Framework aufgegriffen, das Empfehlungen zum Umgang mit den Fehlerquellen beinhaltet und demnach dazu genutzt werden kann, eine Einschätzung bzgl. des Umgangs der Forschung mit der Generierung des Interaktionsterms und auch mit der vorherigen Datenbehandlung zu ermöglichen. Hierbei werden jedoch nur die Untersuchungen evaluiert, welche die Voraussetzungen zur Verwendung des Frameworks erfüllen. Aus diesem Grund werden in den Betrachtungen dieses Abschnitts lediglich 23 Publikationen mit insg. 25 Untersuchungen einbezogen, da sieben Artikel bei der Untersuchung aller ME ein HCM untersuchen und deshalb ausgeschlossen werden mussten. Dies bedeutet gleichzeitig, dass Artikel zumindest teilweise mithilfe des Frameworks analysiert werden können, bei denen nicht jede MA mit einem HCM zusammenhängt, was bspw. bei dem Artikel von Kumar, Singh, Pereira und Leonidou (2020) der Fall ist. Der Nichteinsatz von metrischen Variablen hat darüber hinaus zum Ausschluss drei weiterer Artikel geführt, die lediglich Effekte im Zusammenhang mit kategorischen exogenen oder moderierenden Variablen beinhalten⁷.

Insgesamt werden in 20 Untersuchungen Modelle bestehend aus lediglich reflektiv gemessenen Konstrukten analysiert. Dabei musste jedoch berücksichtigt werden, dass in einer Untersuchung verschiedene Konstellationen von MA vor-

kommen können, die sowohl Modelle aus ausschließlich reflektiv gemessenen Konstrukten als auch Modelle mit mindestens einem formativ gemessenem exogenen oder moderierenden Konstrukt beinhalten. Demzufolge kann eine Untersuchung zur Menge von Untersuchungen mit ausschließlich reflektiv gemessenen Konstrukten, aber auch zur Menge mit mindestens einer entsprechenden formativen Messung gehören. Inklusive der Modelle mit zumindest teilweise formativ gemessenen Konstrukten (7 mal) ergeben sich durch die doppelte Betrachtung zweier Untersuchungen, die beide Modelle enthalten⁸, somit 27 individuelle zu analysierende Fälle. Diese werden nachfolgend auf die ersten vier potenziellen Fehlerquellen überprüft, wobei die Herausforderungen bei der Generierung des Interaktionsterms, deren Vorkommen und passende Beispiele in Tabelle 3 zusammengefasst sind.

Zunächst wird mit der Verwendung eines falschen Ansatzes bei formativ gemessenen exogenen oder moderierenden Konstrukten der schwerwiegendste Fehler untersucht, der bei der Generierung des Interaktionsterms auftreten kann und demnach einen Verstoß gegen das Framework darstellt. Hierbei geht die inhaltliche Bedeutung des Interaktionsterms verloren, wenn dieser mithilfe des Orthogonalisierungsansatzes oder des Produkt-Indikator-Ansatzes gebildet wird, woraufhin die gesamte MA ihre Aussagekraft verliert (Henseler & Fassott, 2010). Insgesamt weisen Untersuchungen in Mahr et al. (2014) und Rollins et al. (2012) sowie eine MA in Matsuno et al. (2014) einen Verstoß gegen das Framework auf. Die erstgenannten Autoren nutzen formativ gemessene MV und die zweit- sowie drittgenannten Autoren verwenden eine formativ gemessene exogene Variable, um den Interaktionsterm fälschlicherweise mithilfe des Produkt-Indikator-Ansatzes zu generieren. Prozentual bildet nur ein kleiner Teil mit 11,11 Prozent aller 27 mit dem Framework analysierbaren Fälle den Interaktionsterm konzeptionell falsch. Insgesamt wird jedoch lediglich in sieben der 27 Fälle (25,93%) mindestens ein formatives Konstrukt im Zusammenhang mit einer MA genutzt. Demnach ist auffällig, dass mit drei von sieben Publikationen insg. 42,86% der Artikel den Interaktionsterm aus mindestens einem formativen Konstrukt nicht korrekt mit dem Zwei-Stufen-Ansatz bildet. Hierbei wird in keiner der gegen das Framework verstoßenden Untersuchungen auf einen der Übersichtsartikel von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) oder Becker et al. (2018) hingewiesen. Dies legt die Vermutung nahe, dass die Verstöße mit einem allgemeinen Unwissen von Forschern über die Einschränkungen des Produkt-Indikator-Ansatzes und Orthogonalisierungsansatzes begründet werden könnten.

Darüber hinaus besteht auch bei der Wahl des Ansatzes in Modellen mit ausschließlich reflektiv gemessenen Konstrukten Unsicherheit in der aktuellen Forschung

⁷Die ME, die innerhalb einer moderierten Mediation auftraten, wurden nicht im Kontext der Mediation betrachtet, sondern als herkömmliche MA angesehen. Diese Vorgehensweise diente der Vereinfachung des Reviewprozesses und hat dennoch allgemeingültige Erkenntnisse des Umgangs von Forschern mit MA hervorgebracht.

⁸Die Artikel von Matsuno et al. (2014) und Johnston et al. (2012) weisen sowohl Modelle mit ausschließlich reflektiv gemessenen Konstrukten als auch Modelle mit mind. einem formativ gemessenem Konstrukt auf, wobei letzterer Artikel darüber hinaus zwei verschiedene Ansätze (Produkt-Indikator-Ansatz und Zwei-Stufen-Ansatz) in den MA verwendet.

Tabelle 3: Zusammenfassung der Beobachtungen zu bekannten Fehlerquellen bei der Generierung des Interaktionsterms, deren Vorkommen und Beispiele

Bekannte Fehlerquelle	Spezifizierung der Beobachtung	Vorkommen (absolut)	Vorkommen (relativ)	Beispiele
kein Zwei-Stufen-Ansatz (nur Untersuchungen mit mindestens einer formativen Messung)	-	3 von 27 Untersuchungen	11,11% der Untersuchungen	Mahr, Lievens und Blazevic (2014); Matsuno, Zhu und Rice (2014); Rollins, Bellenger und Johnston (2012)
	7 formative Messungen von Konstrukten in 27 Untersuchungen	3 von 7 Untersuchungen	42,86% der Untersuchungen	
Wahl eines suboptimalen Ansatzes nach dem Framework (nur reflektive Messungen)	Ziel der MA: 16 mal Herausstellen der Signifikanz und 4 mal Vorhersage von Effekten	14 von 20 Untersuchungen	70% der Untersuchungen	Chen, Peng und Hung (2015); Coelho und Henseler (2012) (beide Untersuchungen)
	Ziel der MA korrigiert: 20 mal Herausstellen der Signifikanz	16 von 20 Untersuchungen	80% der Untersuchungen	Kumar et al. (2020); Teller, Gittenberger und Schnedlitz (2013)
Zitierung eines Übersichtsartikels (nur reflektive Messungen)	gegen das Framework verstoßende Fälle; unveränderte Ziele	3 von 14 Untersuchungen	21,43% der Untersuchungen	Chen et al. (2015); Johnston et al. (2012); Wolter und Cronin (2016)
	gegen das Framework verstoßende Fälle; korrigierte Ziele	4 von 16 Untersuchungen	25% der Untersuchungen	Teller et al. (2013); Wolter und Cronin (2016)
Wahl eines suboptimalen Ansatzes nach dem Framework	Ziel der MA: 23 mal Herausstellen der Signifikanz und 4 mal Vorhersage von Effekten	17 von 27 Untersuchungen	62,96% der Untersuchungen	Matsuno et al. (2014); Mullins, Agnihotri und Hall (2020)
	Ziel der MA korrigiert: 27 mal Herausstellen der Signifikanz	19 von 27 Untersuchungen	70,37% der Untersuchungen	Mahr et al. (2014); Kumar et al. (2020)
Zitierung eines Übersichtsartikels	gegen das Framework verstoßende Fälle; unveränderte Ziele	3 von 17 Untersuchungen	17,65% der Untersuchungen	Chen et al. (2015); Johnston et al. (2012); Wolter und Cronin (2016)
	gegen das Framework verstoßende Fälle; korrigierte Ziele	4 von 19 Untersuchungen	21,05% der Untersuchungen	Johnston et al. (2012); Teller et al. (2013)

(Henseler & Chin, 2010). Es können hierbei prinzipiell alle gängigen Ansätze genutzt werden, um den Interaktionsterm zu bilden. In diesem Fall sollten Forscher jedoch stets den Ansatz wählen, der nach Henseler und Chin (2010) die beste Leistung im Hinblick auf das Ziel der MA bei der Generierung des Interaktionsterms hervorbringt. Das Ziel stellt dabei erfahrungsgemäß nur in Ausnahmefällen die Maximierung der Vorhersagegenauigkeit oder der Parametergenauigkeit dar. Aus diesem Grund wurde im Review bei der Identifikation des Ziels von Forschern bei MA konservativ vorgegangen und das Ziel im Zweifel stets als Maximierung der statistischen Power kodiert. Nur wenn eindeutig von Autoren herausgestellt wurde, dass der Fokus eines Artikels die Vorhersage oder Parametergenauigkeit umfasst, wurde nicht die Maximierung der Power als Ziel der MA anerkannt. Dabei zeigt sich bei der Analyse der 27 Fälle, die mit dem Framework evaluiert werden können, dass bei der Ausprägung der Ziele eindeutig das Testen von Effekten und damit die Maximierung der statistischen Power dominiert. So wird sich in insg.

23 Fällen (85,19%) auf das Testen der Signifikanz von Effekten und gleichzeitig in nur vier Fällen (14,81%) auf die Vorhersage von Effekten fokussiert.

Von diesen 27 Untersuchungen treten wie bereits erwähnt in 20 Untersuchungen Modelle mit ausschließlich reflektiv gemessenen Konstrukten auf. Betrachtet man nun die Wahl des Ziels der Forscher und zieht die Indikatoren der Konstrukte⁹ innerhalb der MA sowie die Stichprobenumfänge heran, lässt sich erkennen, dass in 14 der 20 Untersuchungen (70%) mit ausschließlich reflektiv gemessenen Konstrukten der nach dem aufgestellten Framework suboptimale Ansatz verwendet wurde. Diese niedrige Anzahl der Verwendungen des besten Ansatzes hängt, wie bereits im

⁹In dem Fall, dass die Indikatoren in den Konstrukten unterschiedlich ausgeprägt sind und eine damit zusammenhängende im Framework abgebildete Grenze nicht eindeutig von allen Indikatoren unter- oder überschritten wird, wurde der Mittelwert aller Indikatoren gebildet und anhand dessen beurteilt, ob die Indikatorenanzahl in den Konstrukten die Verwendung eines Ansatzes begünstigt oder nicht.

Vorfeld antizipiert, vorrangig mit dem häufigen Einsatz des Produkt-Indikator-Ansatzes zusammen (15 mal), da das Ziel von Forschern meistens die Maximierung der statistischen Power darstellt (16 mal) und hierbei dieser Ansatz stets suboptimal ist. Insbesondere kann die Wahl des Produkt-Indikator-Ansatzes mit insg. 13 der 14 Verstöße in Verbindung gebracht werden (92,86%), während die Wahl des Orthogonalisierungsansatzes ein Mal gegen das Framework verstößt. Verglichen hiermit wird nur selten beobachtet, d.h. bei drei der 14 Verstöße (21,43%), dass die im Framework aufgestellten Grenzen bzgl. der Indikatoren pro Konstrukt oder des Stichprobenumfangs verletzt und damit ein anderer Ansatz als der im Artikel genutzte der beste Ansatz gewesen wäre. Demnach ist die hohe Anzahl der Verstöße gegen das Framework mit der häufigen Anzahl der Verwendungen des Produkt-Indikator-Ansatzes bei dem Ziel der Maximierung statistischer Power zu erklären. Darüber hinaus hält jeder der vier Einsätze des Zwei-Stufen-Ansatzes und auch der zwei des Produkt-Indikator-Ansatzes das Framework ein. Letzteres ist darauf zurückzuführen, dass in insg. vier Untersuchungen die Maximierung der Vorhersagegenauigkeit das Ziel der MA darstellt, wobei der Produkt-Indikator Ansatz in zwei Fällen den besten Ansatz darstellt und das Framework einhält.

Bei der Suche möglicher Gründe für die häufige Wahl eines suboptimalen Ansatzes fiel auf, dass nur in sechs der 20 Publikationen (30%) die genannten Übersichtsartikel zitiert wurden, wobei die Mehrheit der Publikationen (70%) folglich keinen Verweis der verschiedenen Vor- und Nachteile der Ansätze enthält. Hierbei wurde in drei der sechs Publikationen (50%), die einen der Übersichtsartikel zitieren, der Ansatz nicht korrekt gewählt. Dies ist bspw. bei [Wolter und Cronin \(2016\)](#) zu beobachten, die den Orthogonalisierungsansatz nach [Henseler und Chin \(2010\)](#) als den in jedem Fall besten Ansatz beschreiben, wobei der Zwei-Stufen-Ansatz aber die größere statistische Power hätte hervorbringen können. Da neben den genannten drei Fehldeutungen der Übersichtsartikel jedoch auch drei korrekte Deutungen auftreten, ist unklar, ob trotz der Zitierung eines Übersichtsartikels tendenziell der falsche Ansatz gewählt wird oder ob eine Zitierung eher mit der Wahl des besten Ansatzes in Verbindung gebracht werden kann. Demnach wird kein Zusammenhang zwischen Unsicherheit bei der Verwendung der Informationen aus den Übersichtsartikeln und einer häufigen Wahl gewählter suboptimaler Ansätze beobachtet, der von [Hair et al. \(2019\)](#) vermutet wurde.

Vielmehr könnte bei der Betrachtung der 14 Publikationen, in denen im Falle reflektiver latenter Variablen ein nach dem Framework suboptimaler Ansatz gewählt wurde, die geringe Anzahl an Zitationen einer der Übersichtsartikel erneut einen möglichen Grund für die hohe Anzahl der Verstöße liefern. Insgesamt erwähnen nur drei der 14 gegen das Framework verstößende Artikel (21,43%) einen der Übersichtsartikel. Demnach könnte hier davon ausgegangen werden, dass nicht die falsche Anwendung der Vorgaben in den Übersichtsartikeln für die Unsicherheit der Forscher und häufige Wahl des Produkt-Indikator-Ansatzes sorgt, sondern Uninformiert-

heit über die Existenz der Übersichtsartikel und Unwissen darüber, dass die Ansätze eine unterschiedliche Leistung hervorbringen können. Forscher verwenden in diesem Fall unter Umständen den Ansatz, den sie aus dem Regressionskontext zur Durchführung von MA kennen, weshalb der Produkt-Indikator-Ansatz so häufig gewählt worden sein könnte.

Fast man die 20 Modelle mit ausschließlich reflektiv gemessenen Konstrukten und die sieben mit zumindest teilweise formativ gemessenen Konstrukten zusammen, wird nun in insg. 17 Fällen gegen das Framework bei der Wahl des Ansatzes verstoßen (62,96%). Das bedeutet auch, dass in zehn Fällen der korrekte Ansatz bzgl. der Messmodelle gewählt wird (37,04%). Hierbei wird erneut der Zusammenhang zwischen der Wahl des falschen bzw. suboptimalen Ansatzes und dem nicht vorhandenen Verweis auf einen der besprochenen Übersichtsartikel deutlich, da in 14 der 17 Verstöße keine Zitation auftritt (82,35%). Demnach wird weiterhin vermutet, dass die Wahl des falschen Ansatzes mit dem Unwissen der Forscher über die Existenz der Empfehlungen in den Übersichtsartikeln zusammenhängt. Gleichzeitig kann die Einhaltung der Empfehlungen wie zuvor nicht eindeutig mit den erfolgten Zitierungen der Übersichtsartikel in Verbindung gebracht werden, da bei sechs der zehn korrekt gewählten Ansätze (60%) ein Übersichtsartikel zitiert wurde und bei vier Einhaltungen kein Verweis gefunden werden konnte.

Dabei sollte jedoch auch die dritte potenzielle Fehlerquelle bei MA in PLS-SEM einbezogen werden, um ein realistischeres Bild der Verstöße gegen die Vorgaben des Frameworks zur Bildung des Interaktionsterms erhalten zu können. Das Finden signifikanter Effekte stellt den primären Fokus einer Studie in PLS-SEM dar, insofern diese zu testende Hypothesen enthält. Dies macht die Differenzierung von [Henseler und Chin \(2010\)](#) nach den Zielen von MA obsolet, wobei höchstens in absoluten Ausnahmefällen die Maximierung des Bestimmtheitsmaßes R^2 und damit die Vorhersagegenauigkeit einen Fokus von Forschern darstellt. Dennoch wird in vier der 27 Fälle (14,81%) fälschlicherweise darauf verwiesen, dass der Fokus auf der Vorhersage von Effekten liegt. Korrigiert man diese vier Fälle dahingehend, dass das eigentliche Ziel in jedem Fall die Maximierung der statistischen Power ist, erhöht sich die Anzahl von Verstößen gegen das Framework von 17 auf 19 von 27 (70,37%) und die Anzahl der Einhaltungen verringern sich von zehn auf acht von 27 (29,63%). Dies kann damit begründet werden, dass zwei der vier korrigierten Fälle ohnehin bereits gegen das Framework verstoßen haben, da ein suboptimaler Ansatz gewählt wurde, um die Vorhersagegenauigkeit zu maximieren. Die verbleibenden zwei Fälle ([Kumar et al. \(2020\)](#); [Teller et al. \(2013\)](#)) wurden demnach korrigiert und verstoßen nun ebenfalls gegen das Framework, da beide den Produkt-Indikator-Ansatz nutzen, der nach dem Framework niemals der beste Ansatz zur Generierung des Interaktionsterms bei dem Ziel der Maximierung statistischer Power darstellt. Zusammengefasst verstößt der Einsatz des Produkt-Indikator-Ansatzes immer gegen das Framework, da die Wahl des besten Ansatzes bei dem Ziel der meisten Forscher vorrangig in dem Auffinden signifikanter Effekte besteht und hierbei in

den meisten Fällen der Zwei-Stufen-Ansatz und seltener der Orthogonalisierungsansatz bei der Maximierung der statistischen Power dominieren.

Werden nun die 19 fehlerhaften Wahlen der Ansätze näher betrachtet, kann erkannt werden, dass hierbei in 15 Fällen (78,95%) keiner der genannten Übersichtsartikel zitiert wurde. Darüber hinaus sind die acht Fälle mit korrekter Wahl des Ansatzes auch in diesem Fall fast gleichverteilt, da in fünf Fällen (60%) einer der Übersichtsartikel erwähnt wird und in drei nicht. Demnach kommt man auch nach der Korrektur mutmaßlich falsch vermuteter Ziele der MA von Forschern zu derselben Vermutung bezüglich der Begründung der häufigen Verstöße gegen das Framework wie zuvor. So kann das Unwissen über die verschiedenartige Leistungsfähigkeit der Ansätze als plausibler Grund für die Wahl des falschen Ansatzes herangezogen werden, anstatt des falschen Umgangs mit den bekannten Empfehlungen der Übersichtsartikel.

Letztlich stellt die Entscheidung über die Art der Behandlung der Indikatordaten eine weitere potenzielle Fehlerquelle dar, die mit der Bildung des Interaktionsterms zusammenhängt. Wie bereits dargelegt, sollten die Daten der Indikatoren vor der Analyse standardisiert werden, was insbesondere bei dem Orthogonalisierungsansatz und dem Produkt-Indikator-Ansatz erfolgen sollte, da der Zwei-Stufen-Ansatz die Daten bei der Generierung des Interaktionsterms ohnehin standardisiert (Becker et al., 2018). Insgesamt weisen lediglich zwölf der 27 betrachteten Fälle (44,44%) eine Angabe zur Datenbehandlung auf, wobei die Daten hierbei in elf Fällen standardisiert und in einem Fall (Matsuno et al., 2014) mittelwertzentriert. Dementsprechend wird in 15 Fällen keine Information dazu gegeben, wie mit den Daten vor der Analyse verfahren wurde.

Dies stellt jedoch nicht in jedem Fall einen Verstoß gegen die Vorgaben des Frameworks dar, da die Standardisierung der Daten vor der Analyse bei der Verwendung des Zwei-Stufen-Ansatzes freiwillig erfolgen kann. Die Angabe keiner Informationen, kann jedoch bei der Verwendung des Produkt-Indikator-Ansatzes oder Orthogonalisierungsansatzes die Einschätzung von Problemen, wie einer Erschwerung der Interpretation der Ergebnisse oder einer potenziellen Verzerrung der Höhe der Pfadkoeffizienten, verhindern (Becker et al., 2018; Hair et al., 2017, S. 257). Schließt man die sechs Fälle bei dem Einsatz des Produkt-Indikator-Ansatzes und den einen Fall bei der Verwendung des Orthogonalisierungsansatzes, in denen keine Informationen zur Datenbehandlung gegeben werden, von der Evaluation mithilfe des Frameworks aus, wird deutlich, dass lediglich 20 der 27 Fälle für die Analyse verbleiben. Werden diese nun unter der Annahme, dass die Autoren stets das Ziel der Maximierung der statistischen Power verfolgen, ebenfalls auf die Korrektheit der verwendeten Ansätze geprüft, ergeben sich noch acht das Framework einhaltende Fälle (40%). Der relative Anteil verstoßender Fälle sinkt somit von 70,37 Prozent auf 60 Prozent, da sämtliche dem Framework folgende Artikel erhalten werden konnten, und der relative Anteil dem Framework folgender Fälle erhöht sich von 29,63 auf 40 Prozent. Wird jedoch davon ausgegangen, dass die sieben ausgeschlossenen Fälle

auch weiterhin Verstöße darstellen, erhöht sich die Anzahl von Verstößen wieder auf 19, während die Anzahl der dem Framework folgenden Fälle bei acht verbleibt und das Verhältnis optimaler und suboptimaler bzw. falscher Generierungen des Interaktionsterms sich wie zuvor ausprägt (70,37% zu 29,63%). In dem vorliegenden Review hat die Datenbehandlung demnach keinen Einfluss auf die Einhaltung oder Verstöße gegen das aufgestellte Framework, da fast alle Artikel die Daten standardisieren und die keine Angaben zur Verfügung stellenden Untersuchungen sowie der Artikel mit der Mittelwertzentrierung bereits aus den zuvor diskutierten Gründen gegen das Framework verstoßen.

Zusammenfassend verstößt mit 19 Fällen die Mehrheit der mit dem Framework untersuchbaren 27 Fälle (70,37%) gegen die Empfehlungen zur Bildung des Interaktionsterms (und der vorangehenden Datenbehandlung). Dies könnte sich insb. auf die geringe Anzahl von Zitationen der Übersichtsartikel von Henseler und Chin (2010), Henseler und Fassott (2010), Hair et al. (2017) sowie Becker et al. (2018) und damit auf das Unwissen vieler Forscher zu den Stärken und Schwächen der Ansätze zurückführen lassen.

4.2.2. Evaluation der Messmodelle und Interpretation der Ergebnisse

Der zweite Bereich potenzieller Fehlerquellen umfasst neben der Interpretation der Ergebnisse einer MA zunächst die Evaluation der Messmodelle der Variablen, die innerhalb der MA geschätzt werden. Dabei sind die gemachten Beobachtungen mit der absoluten und relativen Anzahl der Vorkommnisse und passenden Beispielen sowohl für diesen Abschnitt als auch für den nächsten Abschnitt, d.h. für Beobachtungen zur künstlichen Dichotomisierung kontinuierlicher Variablen, in Tabelle 4 dargestellt.

Während die Evaluation der Messmodelle der latenten moderierenden, exogenen und endogenen Variable in herkömmlicher Art und Weise ablaufen sollte, weisen Forscher widersprüchliche Ansichten bezüglich der Evaluation des Interaktionsterms auf (Hair Jr, Hult, Ringle & Sarstedt, 2017, S. 255). Dabei konnte im Rahmen der vorliegenden Arbeit keine Empfehlung gegeben werden, wie mit der Evaluation insb. der Reliabilität des Interaktionsterms umzugehen ist.

Das Review aller 35 Untersuchungen¹⁰ gibt jedoch Aufschluss über den Umgang der einschlägigen Literatur mit der Evaluation des Interaktionsterms bei einfachen MA in PLS-SEM. Es fällt auf, dass im Rahmen der Evaluation des Messmodells des Interaktionsterms lediglich die diskriminante Validität des Interaktionsterms geprüft wurde. Hierbei überprüfen dies nur Calantone und Rubera (2012) sowie Köhler et al. (2011), wobei in keinem der zwei Artikel ein Grund hierfür angegeben ist.

Insbesondere die Reliabilität des Interaktionsterms soll jedoch nach Aguinis et al. (2017) und Li et al. (2019) beachtet werden, um Verzerrungen der Ergebnisse durch Messfehler zu vermeiden. Die Autoren differenzieren dabei nicht

¹⁰Da eine Analyse mithilfe des Frameworks nun nicht mehr nötig ist, können wieder alle 35 Untersuchungen herangezogen werden.

Tabelle 4: Zusammenfassung der Beobachtungen bekannter Fehlerquellen bei der Evaluation der Messmodelle und bei der Interpretation der Ergebnisse

Bekannte Fehlerquelle	Vorkommen (absolut)	Vorkommen (relativ)	Beispiele
Evaluation des Interaktionsterms (diskriminante Validität)	2 von 35 Untersuchungen	5,71% der Untersuchungen	Calantone und Rubera (2012); Köhler, Rohm, De Ruyter und Wetzels (2011)
Evaluation des Interaktionsterms (interne Konsistenzreliabilität)	0 von 35 Untersuchungen	0% der Untersuchungen	-
Evaluation der reflektiv gemessenen exogenen und moderierenden Variablen (interne Konsistenzreliabilität)	27 von 27 Untersuchungen	100% der Untersuchungen	Heidenreich, Wittkowski, Handrich und Falk (2015); Pelser et al. (2015)
keine Trennung des Main Effects und Simple Effects	12 von 35 Untersuchungen	34,29% der Untersuchungen	Matsuno et al. (2014); Nakata, Zhu und Izberk-Bilgin (2011); Pelser et al. (2015)
Prüfung der MV auf Signifikanz	1 von 35 Untersuchungen	2,86% der Untersuchungen	Krishen, Leenders, Muthaly, Ziółkowska und LaTour (2018)
Angabe der Effektgröße f^2 (individuelle und kollektive ME)	10 von 35 Untersuchungen	28,57% der Untersuchungen	Johnston et al. (2012); Ahrholdt, Guderger und Ringle (2019) (beide Untersuchungen)
davon Angabe der Effektgröße f^2 (individuelle ME)	7 von 10 Untersuchungen	70% der Untersuchungen	Johnston et al. (2012); Teller et al. (2013); Zobel (2017); Coelho und Henseler (2012) (beide Untersuchungen)
dabei Einschätzung der Höhe der Effektgröße f^2 nach Kenny (2018); (hiervon 3 mal ohne Angabe)	2 von 10 Untersuchungen	20% der Untersuchungen	Ahrholdt et al. (2019) (beide Untersuchungen)
künstliche Dichotomisierung kontinuierlicher Variablen	0 von 33 Untersuchungen	0% der Untersuchungen	-

zwischen den Messmodellen, wobei die Evaluation der internen Konsistenzreliabilität formativ gemessener Konstrukte im Allgemeinen nicht empfohlen wird (Hair et al., 2017, S. 137-138). Es wurden dennoch alle 35 Untersuchungen auf die Reliabilität des Interaktionsterms geprüft, da dieser nicht formativ gemessen wird. Jedoch gibt keine analysierte Untersuchung einen Wert für Cronbachs Alpha oder die Composite Reliability des Interaktionsterms an, welche die interne Konsistenzreliabilität von Konstrukten messen können (Hair et al., 2017, S. 106). Aguinis et al. (2017) empfehlen, dass zumindest die Angabe der Reliabilität aller exogenen Variablen und der MV, die Interaktionsterme bilden, damit sich im Fall nicht-signifikanter ME bei der Interpretation der Ergebnisse der MA hierauf bezogen werden kann. Im PLS-SEM-Kontext sollte eines der gängigen Reliabilitätsmaße demnach zumindest bei reflektiv gemessenen Konstrukten mit mehr als einem Indikator erfasst werden. Dem kommen sämtliche Untersuchungen im Review nach, da in jedem Fall reflektiver gemessener Konstrukte mindestens ein Reliabilitätsmaß angegeben wurde. In acht Fällen wurden keine Angaben zur Reliabilität gemacht, wobei hier mit formativ gemessenen Konstrukten oder latenten Variablen mit nur einem Indikator verfahren wurde und die Evaluation der Reliabilität nicht wie bei reflektiven Messungen möglich ist (Hair et al., 2017, S. 108-

109, 137-138). Dies steht den Resultaten von Aguinis et al. (2017) gegenüber, die in ihrem Review herausfanden, dass in der Mehrheit der von ihnen untersuchten Artikel keine Angaben zur Reliabilität der Messungen machen. Demzufolge wird die Evaluation und insb. die Reliabilität der Konstrukte in Marketinganwendungen von PLS-SEM in allen Fällen berichtet, weshalb sie in diesem Kontext nicht als Problemquelle angesehen werden sollte.

Die Interpretation des Main Effects und Simple Effects stellt die nächste potenzielle Problemquelle dar, da hierbei ebenfalls gegensätzliche Meinungen in der Forschung vorherrschen. Insbesondere ist unklar, ob die Effekte nach Hair et al. (2017, S. 257-258), Chin et al. (2003) und Henseler und Fassott (2010) in getrennten Modellen, d.h. jeweils ohne und mit der latenten MV im Modell, oder nach Aguinis et al. (2017) und Li et al. (2019) in demselben Modell analysiert werden sollen. Letztere Autoren sprechen sich dabei dafür aus, den Main Effect bei Vorliegen eines signifikanten ME vollständig aus der Interpretation zu exkludieren und lediglich den Simple Effect zu interpretieren. Alles in allem wurden in den vorliegenden 35 Untersuchungen zwölf Mal der Main und Simple Effect nicht getrennt (34,29%), wobei in den meisten Fällen eine getrennte Berechnung des Main Effects- und Simple Effects-Modells erfolgt (65,71%). Bei-

spielsweise wurde im Artikel von Matsuno et al. (2014) nur ein Modell berechnet, in dem alle Effekte simultan geschätzt und berichtet wurden und der Main Effect demnach nicht vorkommt. Der Simple Effect wird in diesem Fall zum Testen einer Hypothese genutzt, was nach Hair et al. (2017, S. 257-258) nicht der inhaltlich korrekten Vorgehensweise bei der Interpretation der Ergebnisse entspricht, nach Aguinis et al. (2017) jedoch genau richtig ist.

Darüber hinaus beinhaltet die nächste potenzielle Fehlerquelle im Rahmen der Interpretation der Ergebnisse von MA die Prüfung der MV anstatt des Interaktionsterms auf Signifikanz. Nach Henseler und Fassott (2010) sollte stets der Interaktionsterm zur Überprüfung der Signifikanz eines ME herangezogen werden. In einer der 35 in das Review eingeschlossenen Untersuchungen (2,86%) konnte diesbezüglich festgestellt werden, dass der Pfadkoeffizient p_2 von einer MV auf eine endogene Variable auf Signifikanz überprüft wurde. So stellt der Artikel von Krishen et al. (2018) einen Fall dar, in dem der direkte Effekt der MV auf endogene Konstrukte getestet und interpretiert wurde, die mitunter als exogene latente Variablen in der MA hätten dienen müssen. Insbesondere wird hierbei keine MA durchgeführt, sondern eine Mediationsanalyse, da direkte Effekte der MV auf zwei endogene Konstrukte und direkte Effekte des einen endogenen Konstrukts auf das andere endogene Konstrukt hypothetisiert und schließlich geschätzt werden. Weiterhin finden sich in den Ergebnissen der Studie generell keine Interaktionsterme, wobei diese nach den Autoren dennoch gebildet wurden. Die Interpretation der MA, u.a. bestehend aus der Betrachtung des linearen Einflusses der „MV“ auf endogene Variablen, ist folglich inhaltlich nicht korrekt, da diese ebenfalls zu einer Mediationsanalyse passen würde. Die Ergebnisse der „MA“ sind demnach zwar im Rahmen einer Mediationsanalyse verwertbar, aber nicht für die Analyse von ME. Im Rahmen einer korrekt hypothetisierten und konzeptionell richtig modellierten MA wurde allerdings kein Fehler im Sinne der fälschlichen Prüfung der MV auf Signifikanz gefunden.

Schließlich können bei der Interpretation auch Fehler auftreten, die mit der Bildung und Einschätzung der Effektgröße f^2 zusammenhängen. Wie im Vorfeld vermutet wurde, weisen auch die 35 Untersuchungen im Review lediglich in zehn Untersuchungen (28,57%) die Angabe der Effektgröße des ME auf. Bei der Betrachtung der angegebenen zehn Effektgrößen konnte beobachtet werden, dass in den meisten Fällen (70%) herausgestellt wird, wie sehr ein individueller ME zur Erklärung der endogenen latenten Variable beiträgt. Demnach wurden in sieben Untersuchungen für jede MV einzeln die Größe des ME bestimmt. Drei Untersuchungen (30%) bestimmen die Effektgröße aller MV im Modell jedoch simultan, wobei nicht zu erkennen ist, wie sehr die einzelnen MV zur Erklärung der endogenen Variable beitragen. Dieses Vorgehen kann kritisch betrachtet werden, da bei der Interpretation der ME die Einschätzung der Höhe der Effektgröße in diesem Fall nicht für jeden ME einzeln möglich ist. So ist bspw. in dem Fall, dass eine MV auf mehrere Beziehungen wirkt, die alle mit derselben endogenen Variable zusammenhängen, unklar, wie hoch die Effektgröße der ME in

jeder Beziehung tatsächlich ist. Unter anderem kann in beiden Untersuchungen im Artikel von Ahrholdt et al. (2019) nicht erkannt werden, wie hoch die Effektgrößen der fünf ME und damit die Größe der fünf Effekte ist, die auf die Beziehungen der fünf exogenen Variablen auf die endogene Variable wirken. Die Autoren Teller et al. (2013) hingegen prüfen bspw. jeden ME einzeln auf ihren Beitrag zur Erklärung der Varianz der endogenen Variable. Auch hier wirkt dieselbe MV auf mehrere Pfade im Modell.

Neben der Bildung der Effektgröße f^2 stellt auch die Einschätzung der Höhe der Effektgröße im Rahmen der Interpretation des ME eine Problemquelle dar. Diese kann nach Cohen (1988, S. 412-414) oder Kenny (2018) erfolgen, wobei in in der vorliegenden Arbeit die Einschätzung nach Kenny (2018) empfohlen wird, nach der ein kleiner Effekt zumindest 0,005 betragen sollte. Lediglich Ahrholdt et al. (2019) nutzen in ihren beiden Untersuchungen die Einteilung von f^2 nach Kenny (20%), wogegen in fünf Untersuchungen noch auf Cohen (1988, S. 412-414) zurückgegriffen wird (50%). Schließlich sind in drei Untersuchungen keine Informationen bezüglich der Einschätzung von Effektgrößen der ME angegeben (30%). Dies zeigt ein gewisses Unwissen der aktuellen Forschung darüber, dass ME i.d.R. sehr klein ausfallen. Schließlich könnte sich dieses Unwissen in Form von voreiligen Ausschlüssen einzelner ME aus der Analyse manifestieren, die keinen zumindest kleinen Effekt ausüben. Wird in einer Untersuchung mithilfe von Cohen (1988, S. 412-414) versucht f^2 einzuschätzen, ist denkbar, dass ein ME infolge eines zu kleinen Effekts von der Interpretation ausgeschlossen wird. Da Effektgrößen jedoch nach Aguinis et al. (2005) häufig unterhalb der unteren Grenze von Cohen (1988, S. 412-414) liegen, sollten diese nicht verfrüht bei einer Unterschreitung der unteren Grenze ausgeschlossen werden. Keine der inkludierten Untersuchungen weist bei den ME eine Unterschreitung der unteren Grenze nach Cohen (1988, S. 412-414) auf. Die Vermutung, dass es zu verfrühten Ausschlüssen von ME von der Interpretation gekommen ist, kann im Rahmen der vorliegenden Arbeit weder widerlegt noch bestärkt werden.

4.2.3. Künstliche Dichotomisierung kontinuierlicher Moderatorvariablen

Letztlich kann neben der Generierung des Interaktionsterms und der Evaluation der Ergebnisse auch der Umgang mit der Skalierung der MV bei einfachen MA eine potenzielle Problemquelle darstellen. Bei der Skalierung von MV unterscheiden Henseler und Fassott (2010) zwischen kategorischen und kontinuierlichen MV, wobei kontinuierliche Variablen durch künstliche Dichotomisierung in Kategorien eingeteilt und somit zu kategorischen MV umgewandelt werden können. Nach Aguinis et al. (2017) reduziert diese Vorgehensweise jedoch die Varianz der MV, wodurch die Analyse dieser beeinträchtigt wird. Im PLS-SEM-Kontext konnte bei der Betrachtung von 33 der 35 Untersuchungen (verbleibende zwei Untersuchungen beinhalten bereits ausschließlich kategorische MV) im Review keine Untersuchung identifiziert werden, die kontinuierliche Variablen künstlich di-

chotomisiert. Hierbei steht das vorliegende Ergebnis somit der Feststellung von [Aguinis et al. \(2017\)](#) gegenüber, die in 10,24 Prozent der von den Autoren untersuchten Artikel eine künstliche Dichotomisierung vorfanden. Da im aktuellen Review keine künstlichen Dichotomisierung gefunden wurde, folgen sämtliche inkludierte Artikel den Vorgaben von [Henseler und Fassott \(2010\)](#) und [Aguinis et al. \(2017\)](#), kontinuierliche MV nicht zu dichotomisieren sondern sie metrisch bzw. quasi-metrisch zu belassen.

4.3. Sonstige Beobachtungen

Der aktuelle Abschnitt umfasst weitere Beobachtungen, die erst während des Reviewprozesses gemacht werden konnten und deshalb im Vorfeld nicht theoretisch hergeleitet wurden. Aus diesem Grund beziehen sich die nachfolgenden Beobachtungen nicht auf bekannte Fehlerquellen, sondern repräsentieren neue Erkenntnisse bezüglich potenzieller und unerforschter Fehlerquellen bei Moderation in Marketinganwendungen von PLS-SEM. Diese Erkenntnisse sind ebenfalls in Tabelle 5 zusammengetragen und wurden darüber hinaus mit passenden Beispielen aus den in das Review eingeschlossenen Artikeln ergänzt.

Die erste Auffälligkeit in den kodierten Daten der 35 Untersuchungen im Review umfasst das häufige Vorkommen mehrerer ME, die innerhalb einer Untersuchung getestet wurden. Diese multiplen ME können dabei auf dieselbe oder verschiedene Beziehungen im Modell wirken. Alles in allem haben in 21 der 35 Untersuchungen (60%) mehrere MV einen Einfluss auf eine und/oder mehrere Beziehungen. PLS-SEM ist nach [Sarstedt, Ringle et al. \(2020\)](#) geeignet, um auch komplexe Modelle mit mehreren MV zu schätzen. Dennoch weisen [Frazier et al. \(2004\)](#) darauf hin, dass die simultane Analyse multipler MV in demselben Modell Probleme hervorrufen kann. Nach den Autoren kann demnach bei der Durchführung vieler MA in einem Modell die Fehlerquote vom Typ 1 erhöht werden. Die Betrachtung der insg. 21 Untersuchungen, in denen mehr als eine MV analysiert wird, gibt Aufschluss darüber, wie häufig die MV simultan bzw. separat, d.h. in demselben bzw. in unterschiedlichen Modellen, geschätzt wurden. Insgesamt wirken verschiedene MV in 16 Untersuchungen auf dieselbe Beziehung im Pfadmodell, wobei die ME dabei in 14 Fällen (87,5%) simultan berechnet wurden. Darüber hinaus haben in acht Untersuchungen multiple MV Einflüsse auf mehrere Beziehungen im Modell, wovon fünf (62,5%) auf einmal geschätzt wurden.

In diesem Fall erfolgte verhältnismäßig häufiger eine simultane Berechnung multipler ME, wenn die MV auf dieselbe Beziehung im Pfadmodell gewirkt haben. Hierbei wird allgemein deutlich, dass die Mehrheit der Untersuchungen mehrere MV analysiert und diese vorrangig simultan geschätzt werden. Infolgedessen bedarf es spezifischer Richtlinien für den Fall, dass mehrere MV in eine MA inkludiert werden. Hier ist bspw. fraglich, ob Forscher die Schätzung der ME simultan, durch sukzessives Hinzufügen der Interaktionsterme oder in gänzlich eigenen Modellen vollziehen sollten.

Weiterhin fiel auf, dass zehn der 35 Untersuchungen (28,57%) mindestens eine MV enthalten, die mit nur einem

Indikator gemessen wird. Dies steht der Empfehlung von [Hair et al. \(2017, S. 246\)](#) gegenüber, MV mit mehr als einem Indikator zu messen. Dabei gehen die Autoren darauf ein, dass eine Messung mit nur einem Indikator weniger prädiktive Power (ausgedrückt in einem kleinen R^2) nach sich zieht als eine Messung mit mehreren Indikatoren und dies insb. das Auffinden kleiner, aber dennoch in der Population signifikanter, ME erschwert. Die Untersuchung von [Teller et al. \(2016\)](#) bspw. enthält die zwei MV „Attraktivität des Standorts des Ladens“ und „Kraft eines Ladens Laufkundschaft zu generieren“, die mit nur einem Indikator gemessen werden. Beide MV sind nicht als nicht signifikant herausgestellt worden, wobei ein Fehler vom Typ 2 aufgetreten sein könnte. Dies kann neben der verringerten prädiktiven Power ebenfalls u.a. mit einem zu geringen Stichprobenumfang oder zu kleinen Effektgrößen der ME zusammenhängen. Allerdings ist unklar, ob ein Fehler vom Typ 2 tatsächlich zu diesem Ergebnis geführt hat oder ob der ME tatsächlich nicht existiert. Werden hierbei auch exogene und endogene Variablen mit einbezogen, die mit nur einem Indikator gemessen werden, erhöht sich die Anzahl von Untersuchungen mit mindestens einer solchen Messung auf 15. Vor dem Hintergrund der potenziellen Nachteiligkeit solcher Messungen, entspricht dies mit 42,86 Prozent einem recht hohen Anteil an den inkludierten Untersuchungen im Review.

Bei der anschließenden Betrachtung der Anzahl signifikanter Ergebnisse aller inkludierter Untersuchungen fällt auf, dass nur zwei Studien keine signifikanten ME berichten (5,71%). Die demnach hohe Zahl signifikanter Ergebnisse ist dabei erwartungswidrig, da ME nach [Aguinis et al. \(2005\)](#) klein ausfallen und damit schwerer zu erkennen sind. Dies hätte zur Folge, dass ein größerer Anteil von nicht signifikanten Ergebnissen hätte gefunden werden müssen. Aus diesem Grund kann vermutet werden, dass die in das Review eingeschlossenen Artikel dem sog. Publication Bias unterliegen. Dieser stellt nach [Rothstein, Sutton und Borenstein \(2005, S. 1-3\)](#) die verzerrte Darstellung von Daten in wissenschaftlichen Publikationen dar. Dabei werden nach den Autoren u.a. selektiv vorrangig signifikante Ergebnisse veröffentlicht, was auch die geringe Anzahl nicht signifikanter und das häufige Vorkommen signifikanter Ergebnisse erklären könnte.

Darüber hinaus treten in nur zwei Fällen kategorische MV auf, wobei diese stets mit einem Indikator gemessen werden ([Henseler & Fassott, 2010](#)). Die Untersuchung im Artikel von [Krishen et al. \(2018\)](#) beinhaltet mit dem Geschlecht und dem Land zwei kategorische MV, während [Rollins et al. \(2012\)](#) mit der Kategorisierung nach verschiedenen hohen Anzahlen von Jahren, die ein Unternehmen Erfahrung mit Customer-Relationship-Management-Systemen hat, nur eine enthält. Somit kommen kategorische Variablen gegenüber kontinuierlichen Variablen jedoch wesentlich seltener (5,71%) in den 35 analysierten Marketinganwendungen vor.

Neben den bisherigen sonstigen Beobachtungen sollten außerdem die neun Untersuchungen aufgegriffen werden, die mindestens einen ME innerhalb einer Mediation getestet haben (25,71%). Hierbei wird der Vorteil von PLS-SEM genutzt, auch sehr komplexe Modellbeziehungen wie moderier-

Tabelle 5: Sonstige Beobachtungen aus dem Review, deren Vorkommen und Beispiele

Beobachtung	Vorkommen (absolut)	Vorkommen (relativ)	Beispiele
multiple ME wirken auf dieselbe bzw. mehrere Beziehungen im Modell	21 von 35 Untersuchungen	60% der Untersuchungen	Ali et al. (2020); Mahr et al. (2014); Mullins et al. (2020)
Messung mind. eines Moderators mit nur einem Indikator	10 von 35 Untersuchungen	28,57% der Untersuchungen	Rollins et al. (2012); Teller, Alexander und Floh (2016); van der Borgh und Schepers (2014)
Messung mind. einer Variable mit nur einem Indikator	15 von 35 Untersuchungen	42,86% der Untersuchungen	Ashill und Jobber (2014); Teller et al. (2016); Wolter und Cronin (2016)
Anzahl nicht signifikanter Ergebnisse	2 von 35 Untersuchungen	5,71% der Untersuchungen	Jean, Sinkovics und Hiebaum (2014); Teller et al. (2016)
Anzahl kategorischer MV	2 von 35 Untersuchungen	5,71% der Untersuchungen	Krishen et al. (2018); Rollins et al. (2012)
mind. eine moderierte Mediation	9 von 35 Untersuchungen	25,71% der Untersuchungen	Kumar et al. (2020); Johnston et al. (2012)
mind. ein HCM im Zusammenhang mit MA	8 von 35 Untersuchungen	22,86% der Untersuchungen	Lam, Ahearne und Schillewaert (2012); Griffith, Lee, Yeo und Calantone (2014)
Hypothesisierung des ME	34 von 35 Untersuchungen	97,14% der Untersuchungen	Miao und Evans (2013); Sharma und Jha (2017); Siahtiri (2018)

te Mediationen ohne Probleme schätzen zu können (Sarstedt, Hair Jr et al., 2020). Diese wurden wie bereits angedeutet jedoch als einfache MA betrachtet und im vorliegenden Review so behandelt, als würden sie auf eine Beziehung zweier Konstrukte wirken, die nicht in einem Mediationsmodell beinhaltet sind. Dies zieht ohnehin vorrangig Konsequenzen für die Schätzung und Interpretation der mit der Mediation einhergehenden Effekte nach sich, da die Moderation auf die Mediation wirkt und nicht umgekehrt. Insbesondere die Struktur der Mediation und der Pfad bzw. die Pfade auf welchen eine oder mehrere MV wirken haben hierauf einen Einfluss (Hair et al., 2017, S. 261). Werden die Modelle von Hayes (2013) betrachtet, die diverse Arten von MA im Zusammenhang mit Mediationen darstellen, können insg. fünf der neun Mediationen einem Modell zugeordnet werden. So wirkt bspw. im Artikel von Kumar et al. (2020) die MV auf beide Pfade des indirekten Effektes, also den Pfaden zu der Mediatorvariable hin und von ihr weg, wobei Modell 58 von Hayes (2013) identifiziert werden konnte.

Mit acht Fällen konnten fast genauso häufig wie Mediationsmodelle auch HCM im Zusammenhang mit MA beobachtet werden. So ist ein HCM in acht Untersuchungen in mindestens einer MA zu finden (22,86%). Hierbei treten HCM jeweils zwei Mal an der Stelle der moderierenden, exogenen, endogenen sowie gleichzeitig der exogenen und endogenen Variable auf. Dabei haben Teller et al. (2016), Köhler et al. (2011), Lam et al. (2012) und Ali et al. (2020) HCM als endogene Variablen verwendet. Darüber hinaus wird nur in den Artikeln von Zobel (2017) und Ali et al. (2020) der Zwei-Stufen-Ansatz zur Repräsentation des HCM im Messmodell genutzt, während Heidenreich, Spieth und Petschnig (2017), Köhler et al. (2011), Kumar et al. (2020) und Lam

et al. (2012) hierfür den Repeated Indicators Approach verwenden. In den Artikeln von Teller et al. (2016) und Griffith et al. (2014) wird keine Angabe zum verwendeten Ansatz gemacht.

Schließlich findet eine Hypothesisierung der ME in fast allen Fällen statt, da nur ein Mal im Fall in Wolter und Cronin (2016) beobachtet werden konnte, dass nicht versucht wurde, einen ME zu hypothesisieren. Somit wird in den meisten Fällen (97,14%) den Forderungen von Andersson et al. (2014) und (Field, 2018, S. 486) nach der Schaffung einer hinreichenden theoretischen Grundlage des vermuteten ME nachgekommen. Im Gegensatz dazu steht die Beobachtung, dass bei der Interpretation nach Memon et al. (2019) Simple Slope Plots verwendet werden sollen, um die Interpretation zu erleichtern und nachvollziehbarer zu machen. Es wurden jedoch nur in neun Fällen (25,71%) Simple Slope Plots abgebildet und für die Interpretation genutzt.

5. Diskussion

5.1. Zusammenfassung

Die vorliegende Arbeit hat den Umgang der aktuellen Forschung mit Moderationsanalysen in Marketinganwendungen von PLS-SEM untersucht. Die Wahlmöglichkeiten zwischen verschiedenen Ansätzen zur Generierung des Interaktionsterms, der zur Modellierung eines ME auf die spezifische Beziehung zweier Variablen benötigt wird, stellen Forscher dabei u.a. vor Herausforderungen. Das Ziel war es deshalb, anhand eines Reviews von 35 empirischen Untersuchungen aus 33 einschlägigen Artikeln herauszufinden, inwiefern bekannte potenzielle Fehlerquellen auftreten, wie mit diesen in der

Literatur umgegangen wird und wo weitere Probleme vorkommen. Hierfür wurde zunächst ein Framework aufgestellt, das Forscher bei der Generierung des Interaktionsterms unterstützt. Nach einer systematischen Literaturrecherche wurde dieses anschließend als eine Grundlage für das Review genutzt.

Die Recherche brachte zunächst die Erkenntnis hervor, dass in den meisten Fällen keine Angaben zum verwendeten Ansatz gemacht werden und der Produkt-Indikator-Ansatz der am häufigsten verwendete Ansatz ist. Im Review konnte die Unsicherheit der Forschung bei der Generierung des Interaktionsterms bestätigt werden, da insg. 19 von 27 mit dem Framework untersuchbare Fälle auftraten, die einen suboptimalen Ansatz verwenden. Darüber hinaus stellen die Behandlung der Indikatordaten vor der Analyse sowie die Evaluation der Konsistenzreliabilität der exogenen und moderierenden Variable ein geringes Problem für Forscher dar. Demgegenüber spiegelt die Interpretation der Main Effects und Simple Effects eine größere Unsicherheit wider, da zwölf der 35 Untersuchungen diese nicht voneinander trennen. Weiterhin werden nur in zehn Untersuchungen Angaben zur Effektgröße f^2 des ME gemacht. Schließlich fielen außerdem eine hohe Anzahl mit nur einem Indikator gemessene Variablen, ein geringes Aufkommen ausschließlich nicht signifikanter Ergebnisse und einige HCM und moderierte Mediationen im Zusammenhang mit MA auf. Die Beobachtungen aus dem Review lassen den Schluss zu, dass trotz vorhandener Übersichtsartikel zur Durchführung von MA in PLS-SEM wie Henseler und Chin (2010) oder Henseler und Fassott (2010) die Empfehlungen häufig nicht in der aktuellen Forschung umgesetzt werden.

5.2. Implikationen

Die Ergebnisse des Reviews in dieser Arbeit tragen in erster Linie zum Verständnis von Herausforderungen der aktuellen Forschung zu PLS-SEM im Umgang mit MA bei. Dabei wird zunächst die Vermutung bestätigt, dass eine Vielzahl von Artikeln die Empfehlungen zur Generierung des Interaktionsterms nicht einhalten. Dies kann vorrangig darauf zurückgeführt werden, dass nur vier der 19 gegen die Vorgaben verstoßenden Fälle einen Übersichtsartikel zitieren. Hierbei kann eine gewisse Unkenntnis der aktuellen Forschung darüber erkannt werden, dass die verschiedenen Ansätze nicht bei allen Messmodellen gleichermaßen genutzt werden bzw. unterschiedliche Ergebnisse hervorbringen können. Einerseits kann dabei die Nutzung eines Ansatzes mit wenig statistischer Power dazu führen, dass Fehler vom Typ 2 auftreten. Andererseits kann im Fall formativ gemessener Konstrukte innerhalb der MA die Verwendung eines anderen anstelle des Zwei-Stufen-Ansatzes zu inhaltlich falschen Ergebnissen und infolgedessen inkorrekten Implikationen führen. Demnach sollten Forscher vermehrt auf die Vorteile und Probleme der Ansätze aufmerksam gemacht werden, was durch weitere Übersichtsartikel in der Art von Henseler und Fassott (2010) oder Zusammenfassungen in zugänglichen Werken wie Hair et al. (2017) erfolgen kann. Dies kann auch den seltenen Angaben des verwendeten Ansatzes, der

bei der Interpretation eine Rolle spielt, der Effektgröße von ME, die zeigen, ob tatsächlich Implikationen aus den signifikanten ME folgen, und dem häufigen Einsatz von Konstrukten, welche mit nur einem Indikator gemessen werden und dadurch ebenfalls die statistische Power verringern können, entgegenwirken.

Im Gegensatz dazu stellt die Überprüfung der Konsistenzreliabilität im Rahmen der Evaluation reflektiv gemessener latenter Variablen nach den Ergebnissen des Reviews eine gängige Praxis in der aktuellen Forschung dar. In sämtlichen Artikeln wurden dementsprechend die Messmodelle der verwendeten Konstrukte evaluiert. Allerdings bestehen noch immer Unsicherheiten darüber, ob eine Evaluation der Reliabilität des Interaktionsterms erfolgen sollte, was sich zunächst an unterschiedlichen Meinungen in der Literatur erkennen ließ. Im Review konnte diese Unsicherheit nicht erkannt werden, da die Evaluation der Reliabilitäten in keinem Artikel durchgeführt wurde. Dies impliziert, dass in der Praxis i.d.R. den Vorgaben (u.a.) von Hair et al. (2017, S. 255) zur Evaluation von Messmodellen gefolgt wird und hier verglichen mit anderen Problemen bei MA in PLS-SEM kaum Aufklärung benötigt wird.

Mehr Unsicherheit in der Forschung besteht demgegenüber bei der Interpretation der Main Effects und Simple Effects. Die Ergebnisse des Reviews deuten hierbei darauf hin, dass Forscher teilweise unterschiedlich mit den Effekten umgehen und sie auf verschiedene Arten generieren. Hierbei wird deutlich, dass die Interpretation der verschiedenen Effekte im Zusammenhang mit MA weiter erforscht werden muss, um eindeutige Empfehlungen für die Auswertung empirischer Studien zur Verfügung stellen zu können.

Letztlich wurde ein relativ großer Anteil von Studien beobachtet, die komplexe MA durch die Inklusion von HCM oder in Form von moderierten Mediationen durchführen, deren Analyse und Interpretation ein hohes Maß an Wissen und Erfahrung fordert (Hair et al., 2017, S. 285). Deshalb sind die Ergebnisse einer entsprechend hohen Anzahl an Studien mit Vorsicht zu betrachten, was auch für die Durchführung eigener Untersuchungen mit komplexen MA gilt.

5.3. Limitationen und zukünftige Forschung

Die Durchführung des Reviews unterlag gewissen Limitationen, die nun diskutiert werden. Zunächst ist die Grundgesamtheit aller Artikel, die in das Review aufgenommen wurden, dadurch eingeschränkt, dass eine zur Verfügung gestellte Datenbank aus 239 empirischen Artikeln verwendet wurde. Es kann hierbei nicht ausgeschlossen werden, dass Artikel in der Datenbank fehlen, die in das vorliegende Review hätten mit aufgenommen werden können. Hinzukommt, dass die Datenbank lediglich Publikationen der Jahre 2011 bis 2020 enthält. Einerseits werden somit ältere Artikel ausgeschlossen, andererseits fehlen die Artikel aus dem Jahr 2021, welche ein aktuelleres Bild des Umgang mit MA in PLS-SEM hätten vermitteln können.

Darüber hinaus basiert das generierte Framework auf verschiedenen Simulationsstudien, die der Limitation unterliegen, dass Parameter, wie Effektgrößen oder Korrelationen

nen zwischen Konstrukten, bei unterschiedlichen Einstellungen verschiedenartige Ergebnisse generieren. Dies wirkt sich demnach auf die nach dem Framework zu wählenden Ansätze aus und hat einen Einfluss auf die Beobachtungen im Review bzgl. der Befolgung oder des Verstoßes des Frameworks bei der Generierung des Interaktionsterms. Weiterhin konnten ausschließlich Untersuchungen ohne HCM und mit kontinuierlichen Variablen mithilfe des Frameworks ausgewertet werden, was den Umfang von einbezogenen Artikeln reduziert hat. Die Generierung des Interaktionsterms wurde bei Untersuchungen, die in jeder MA ein HCM bzw. kategoriale Variablen beinhalten, nicht analysiert. Außerdem wurden ME, die innerhalb einer Mediation aufgetreten sind, nicht im Kontext der Mediation betrachtet sondern als einfache MA ausgewertet.

Insgesamt beleuchtet das vorliegende Review bereits eine Vielzahl von Herausforderungen der aktuellen Forschung, wobei eine Wiederholung des Reviews auf der Basis einer erneuten systematischen Literaturrecherche verschiedenster Datenbanken in Zukunft weitere bzw. andersartige Ergebnisse hervorbringen kann. Hierbei sollten ebenfalls Untersuchungen mit HCM und kategorischen Variablen auf die Vorgehensweise bei der Generierung des Interaktionsterms überprüft werden. Letztlich kann ein Review von Artikeln mit MA durch Multigruppenanalysen in PLS-SEM auf dieselbe Weise durchgeführt werden, um diesbezügliche Herausforderung aufzeigen und bessere Vergleiche beider Arten von MA zu ermöglichen.

Literatur

- Aguinis, H., Beaty, J. C., Boik, R. J. & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: a 30-year review. *Journal of Applied Psychology*, 90 (1), 94–107.
- Aguinis, H., Edwards, J. R. & Bradley, K. J. (2017). Improving our understanding of moderation and mediation in strategic management research. *Organizational Research Methods*, 20 (4), 665–685.
- Ahrholdt, D. C., Gudergan, S. P. & Ringle, C. M. (2019). Enhancing loyalty: When improving consumer satisfaction and delight matters. *Journal of Business Research*, 94, 18–27.
- Ali, I., Ali, M., Salam, M. A., Bhatti, Z. A., Arain, G. A. & Burhan, M. (2020). How international SME's vicarious learning may improve their performance? The role of absorptive capacity, strength of ties with local SMEs, and their prior success experiences. *Industrial Marketing Management*, 88, 87–100.
- Andersson, U., Cuervo-Cazurra, A. & Nielsen, B. B. (2014). From the Editors: Explaining interaction effects within and across levels of analysis. *Journal of International Business Studies*, 25, 1063–1071.
- Ashill, N. J. & Jobber, D. (2014). The effects of the external environment on marketing decision-maker uncertainty. *Journal of Marketing Management*, 30 (3), 268–294.
- Baron, R. M. & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51 (6), 1173–1182.
- Becker, J.-M., Klein, K. & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: guidelines for using reflective-formative type models. *Long Range Planning*, 45, 359–394.
- Becker, J.-M., Ringle, C. M. & Sarstedt, M. (2018). Estimating moderating effects in PLS-SEM and PLSc-SEM: Interaction term generation* data treatment. *Journal of Applied Structural Equation Modeling*, 2 (2), 1–21.
- Calantone, R. & Rubera, G. (2012). When should RD & E and marketing collaborate? The moderating role of exploration–exploitation and environmental uncertainty. *Journal of Product Innovation Management*, 29 (1), 144–157.
- Carte, T. A. & Russell, C. J. (2003). In pursuit of moderation: Nine common errors and their solutions. *MIS Quarterly*, 27 (3), 479–501.
- Chen, A., Peng, N. & Hung, K.-P. (2015). Strategic management of salespeople when promoting new products: Moderating effects of sales-related organizational psychological climate. *European Journal of Marketing*, 49 (9), 1616–1644.
- Chin, W. W., Marcolin, B. L. & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a monte carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14 (2), 189–217.
- Coelho, P. S. & Henseler, J. (2012). Creating customer loyalty through service customization. *European Journal of Marketing*, 46 (3/4), 331–356.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Publishers.
- Echambadi, R., Arroniz, I., Reinartz, W. & Lee, J. (2006). Empirical generalizations from brand extension research: How sure are we? *International Journal of Research in Marketing*, 23 (3), 253–261.
- Esposito Vinzi, V., Trinchera, L. & Amato, S. (2010). PLS path modeling: from foundations to recent developments and open issues for model assessment and improvement. In V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Hrsg.), (S. 47–82). Berlin, Springer.
- Fassott, G., Henseler, J. & Coelho, P. S. (2016). Testing moderating effects in PLS path models with composite variables. *Industrial Management & Data Systems*, 116 (9), 1887–1900.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5. Aufl.). Newcastle, Sage.
- Frazier, P. A., Tix, A. P. & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51 (1), 115–134.
- Griffith, D. A., Lee, H. S., Yeo, C. S. & Calantone, R. (2014). Marketing process adaptation: Antecedent factors and new product performance implications in export markets. *International Marketing Review*, 31 (3), 308–334.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M. & Thiele, K. O. (2017). Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science*, 45, 616–632.
- Hair, J. F., Ringle, C. M. & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19 (2), 139–151.
- Hair, J. F., Sarstedt, M. & Ringle, C. M. (2019). Rethinking some of the rethinking of partial least squares. *European Journal of Marketing*, 53 (4), 566–584.
- Hair, J. F., Sarstedt, M., Ringle, C. M. & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the academy of marketing science*, 40 (3), 414–433.
- Hair Jr, J. F., Hult, G. T. M., Ringle, C. M. & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Newcastle, Sage.
- Hall, J. & Sammons, P. (2014). Mediation, Moderation & Interaction. In T. Teo (Hrsg.), *Handbook of quantitative methods for educational research* (S. 267–286). Rotterdam: SensePublishers. Zugriff auf https://doi.org/10.1007/978-94-6209-404-8_13 doi: 10.1007/978-94-6209-404-8_13
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, Guilford.
- Heidenreich, S., Spieth, P. & Petschnig, M. (2017). Ready, steady, green: Examining the effectiveness of external policies to enhance the adoption of eco-friendly innovations. *Journal of Product Innovation Management*, 34 (3), 343–359.
- Heidenreich, S., Wittkowski, K., Handrich, M. & Falk, T. (2015). The dark side of customer co-creation: exploring the consequences of failed co-created services. *Journal of the Academy of Marketing Science*, 43, 279–296.
- Henseler, J. & Chin, W. W. (2010). A comparison of approaches for the analysis of interaction effects between latent variables using partial least squares path modeling. *Structural equation modeling*, 17 (1), 82–109.
- Henseler, J. & Fassott, G. (2010). Testing moderating effects in PLS Path Models: An illustration of available procedures. In V. Esposito Vinzi, W. W. Chin, J. Henseler & H. Wang (Hrsg.), *Handbook of partial least squares: Concepts, methods and applications* (S. 713–735). Berlin, Heidelberg: Springer Berlin Heidelberg. Zugriff auf https://doi.org/10.1007/978-3-540-32827-8_31 doi: 10.1007/978-3-540-32827-8_31
- Henseler, J., Fassott, G., Dijkstra, T. K. & Wilson, B. (2012). Analysing quadratic effects of formative constructs by means of variance-based structural equation modelling. *European Journal of Information Systems*, 21 (1), 99–112.
- Hult, G. T. M., Reimann, M. & Schilke, O. (2009). Worldwide faculty perceptions of marketing journals: rankings, trends, comparisons, and segmentations. *Trends, Comparisons, and Segmentations*, 3 (3), 1–23.
- Jean, R.-J. B., Sinkovics, R. R. & Hiebaum, T. P. (2014). The effects of supplier involvement and knowledge protection on product innovation in customer–supplier relationships: A study of global automotive suppliers in china. *Journal of Product Innovation Management*, 31 (1), 98–113.
- Johnston, W. J., Khalil, S., Jain, M. & Cheng, J. M.-S. (2012). Determinants of joint action in international channels of distribution: The moderating role of psychic distance. *Journal of International Marketing*, 20 (3), 34–49.
- Kenny, D. A. (2018). *Moderation*. Zugriff auf <http://davidakenny.net/cm/moderation.htm>
- Köhler, C. F., Rohm, A. J., De Ruyter, K. & Wetzels, M. (2011). Return on interactivity: The impact of online agents on newcomer adjustment. *Journal of Marketing*, 75 (2), 93–108.
- Krishen, A. S., Leenders, M. A., Muthaly, S., Ziolkowska, M. & LaTour, M. S. (2018). Social networking from a social capital perspective: a cross-cultural analysis. *European Journal of Marketing*, 53 (6), 1234–1253.
- Kumar, P., Singh, S. K., Pereira, V. & Leonidou, E. (2020). Cause-related marketing and service innovation in emerging country healthcare:

- Role of service flexibility and service climate. *International Marketing Review*, 37 (5), 803–827.
- Lam, S. K., Ahearne, M. & Schillewaert, N. (2012). A multinational examination of the symbolic-instrumental framework of consumer-brand identification. *Journal of International Business Studies*, 43 (3), 306–331.
- Li, M., Sharp, B. M., Bergh, D. D. & Vandenberg, R. (2019). Statistical and methodological myths and urban legends in strategic management research: The case of moderation analysis. *European Management Review*, 16 (1), 209–220.
- Little, T. D., Bovaird, J. A. & Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling*, 13 (4), 497–519.
- M., E. (2014). *Metaanalyse* (1. Aufl.). München, Rainer Hampp Verlag.
- Mahr, D., Lievens, A. & Blazevic, V. (2014). The value of customer cocreated knowledge during the innovation process. *Journal of Product Innovation Management*, 31 (3), 599–615.
- Matsuno, K., Zhu, Z. & Rice, M. P. (2014). Innovation process and outcomes for large Japanese firms: Roles of entrepreneurial proclivity and customer equity. *Journal of Product Innovation Management*, 31 (5), 1106–1124.
- Memon, M. A., Cheah, J.-H., Ramayah, T., Ting, H., Chuah, F & Cham, T. H. (2019). Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, 3 (1), 1–11.
- Miao, C. F. & Evans, K. R. (2013). The interactive effects of sales control systems on salesperson performance: a job demands-resources perspective. *Journal of the Academy of Marketing Science*, 41, 73–90.
- Mullins, R., Agnihotri, R. & Hall, Z. (2020). The ambidextrous sales force: aligning salesperson polychronicity and selling contexts for sales-service behaviors and customer value. *Journal of Service Research*, 23 (1), 33–52.
- Nakata, C., Zhu, Z. & Izberk-Bilgin, E. (2011). Integrating marketing and information services functions: a complementarity and competence perspective. *Journal of the Academy of Marketing Science*, 39 (5), 700–716.
- Pelser, J., de Ruyter, K., Wetzels, M., Grewal, D., Cox, D. & van Beuningen, J. (2015). B2B channel partner programs: Disentangling indebtedness from gratitude. *Journal of Retailing*, 41 (9), 660–678.
- Rollins, M., Bellenger, D. N. & Johnston, W. J. (2012). Does customer information usage improve a firm's performance in business-to-business markets? *Industrial Marketing Management*, 41 (6), 984–994.
- Rothstein, H. R., Sutton, A. J. & Borenstein, M. (2005). *Publication bias in Meta-Analysis: Prevention, assessment and adjustments*. New Jersey, John Wiley & Sons Ltd.
- Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O. & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, 69, 3998–4010.
- Sarstedt, M., Hair Jr, J. F., Nitzl, C., Ringle, C. M. & Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: use of PLS-SEM for mediation analyses! *International Journal of Market Research*, 62 (3), 288–299.
- Sarstedt, M., Ringle, C. M., Cheah, J.-H., Ting, H., Moisescu, O. I. & Radomir, L. (2020). Structural model robustness checks in PLS-SEM. *Tourism Economics*, 26 (4), 531–554.
- Sharma, R. & Jha, M. (2017). Values influencing sustainable consumption behaviour: Exploring the contextual relationship. *Journal of Business Research*, 76, 77–88.
- Siahtiri, V. (2018). Innovation at the service encounter in knowledge intensive business services: antecedents and boundary conditions. *Journal of Product Innovation Management*, 35 (5), 742–762.
- Teller, C., Alexander, A. & Floh, A. (2016). The impact of competition and cooperation on the performance of a retail agglomeration and its stores. *Industrial Marketing Management*, 52, 6–17.
- Teller, C., Gittenberger, E. & Schnedlitz, P. (2013). Cognitive age and grocery-store patronage by elderly shoppers. *Journal of Marketing Management*, 29 (3-4), 317–337.
- van der Borgh, M. & Schepers, J. J. (2014). Do retailers really profit from ambidextrous managers? The impact of frontline mechanisms on new and existing product selling performance. *Journal of Product Innovation Management*, 31 (4), 710–727.
- Wold, H. (1975). Path models with latent variables: The NIPALS approach, *International Perspectives on Mathematical and Statistical Modeling*. In *Quantitative sociology* (S. 307–357).
- Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Hrsg.), *Systems under direct observations: Part II* (S. 1–54). Amsterdam, North-Holland.
- Wolter, J. S. & Cronin, J. J. (2016). Re-conceptualizing cognitive and affective customer-company identification: the role of self-motives and different customer-based outcomes. *Journal of the Academy of Marketing Science*, 44 (3), 397–413.
- Zobel, A.-K. (2017). Benefiting from open innovation: A multidimensional model of absorptive capacity. *Journal of Product Innovation Management*, 34 (3), 269–288.