

### **Advisory Editorial Board:**

DOMINIK VAN AAKEN FREDERIK AHLEMANN CHRISTOPH BODE ROLF BRÜHL JOACHIM BÜSCHKEN LEONHARD DOBUSCH RALF ELSAS DAVID FLORYSIAK **GUNTHER FRIEDL WOLFGANG GÜTTEL CHRISTIAN HOFMANN** KATJA HUTTER **LUTZ JOHANNING** STEPHAN KAISER ALFRED KIESER NATALIA KLIEWER DODO ZU KNYPHAUSEN-AUFSEß SABINE T. KÖSZEGI ARJAN KOZICA TOBIAS KRETSCHMER HANS-ULRICH KÜPPER REINER LEIDL ANTON MEYER GORDON MÜLLER-SEITZ GÜNTER MÜLLER-STEWENS **BURKHARD PEDELL** MARCEL PROKOPCZUK TANJA RABL SASCHA RAITHEL **ASTRID REICHEL** KATJA ROST MARKO SARSTEDT **DEBORAH SCHANZ** ANDREAS G. SCHERER STEFAN SCHMID **UTE SCHMIEL CHRISTIAN SCHMITZ** PHILIPP SCHRECK GEORG SCHREYÖGG LARS SCHWEIZER DAVID SEIDL THORSTEN SELLHORN ANDREAS SUCHANEK **ORESTIS TERZIDIS** ANJA TUSCHKE SABINE URNIK STEPHAN WAGNER BARBARA E. WEIßENBERGER ISABELL M. WELPE HANNES WINNER CLAUDIA B. WÖHLE THOMAS WRONA

THOMAS ZWICK

# JUNIOR MANAGEMENT SCIENCE

Randi Baum und Tatjana Egelhof, Die	
Verwendung von Emojis in der	
Konsumentenkommunikation – Eine	
stimmungsanalytische Betrachtung von	
Kurznachrichten im Social Web	1
Markus Sebastian Gebhart, Measuring Corporate Tax Avoidance – An Analysis of Different	
Measures	43
Raphael Uster, Der Einfluss von CEO Duality auf	
die Bewertung von Schweizer	
Unternehmen	61
Maximilian Krebs, Die Einstellung	
kulturinteressierter Personen gegenüber	
Sponsoren kultureller Events – Eine	
gualitative Studie	84

Published by Junior Management Science e. V.

#### Vorwort

Die vorliegende, vierte Ausgabe von Junior Management Science (JUMS) ist wieder einmal eine besondere. Gerne möchten wir das Vorwort nutzen, Sie über die aktuellen Entwicklungen unseres Journals zu informieren.

Mittlerweile publizieren wir JUMS vierteljährlich. Dies wird uns durch die großzügige Förderung der Joachim Herz Stiftung (JHS) ermöglicht, die unser wissenschaftliches Projekt professionell und finanziell begleitet. Die JHS wurde vom Hamburger Unternehmer Joachim Herz gegründet. Ihre Ziele liegen insbesondere in der Bildung junger Menschen und der Unterstützung von Nachwuchswissenschaftlern. Die Förderung der Stiftung ermöglicht es uns, JUMS in Zukunft weiter zu professionalisieren und bekannt zu machen.

Zudem freuen wir uns sehr, dass sich unser wissenschaftliches Netzwerk von Unterstützern weiterentwickelt hat. Mittlerweile können wir über 50 ProfessorenInnen von mehr als 30 Universitäten in Deutschland, Österreich und der Schweiz in unserem Advisory Editorial Board begrüßen. Um deren Arbeit noch besser begleiten zu können, übernehmen künftig unsere fünf Department-Editoren die Koordination des Review-Prozesses und die Qualitätssicherung in jeweils zwei Fachbereichen. Es freut uns, auch bei unseren Department-Editoren ein breites Spektrum an hervorragenden Wissenschaftlern gewonnen zu haben.

Im Einzelnen werden die Departments durch folgende Professoren übernommen:

- Department 1: Marketing, Technologie, Innovation und Entrepreneurship Prof. Dr. Sascha Raithel (Freie Universität Berlin)
- Department 2: Controlling, Betriebswirtschaftliche Steuerlehre Prof. Dr. Gunther Friedl (Technische Universität München)
- Department 3: Organisation und Personalwesen, Ethik und Nachhaltigkeit in der BWL Prof. Dr. Dominik van Aaken (Universität Salzburg)
- Department 4: Banken und Finanzierung, Externes Rechnungswesen und Wirtschaftsprüfung Prof. Dr. David Florysiak (University of Southern Denmark)
- Department 5: Wirtschaftsinformatik, Produktionswirtschaft und Logistik Prof. Dr. Frederik Ahlemann (Universität Duisburg-Essen)

In den letzten Monaten haben uns viele spannende Abschlussarbeiten aus der DACH-Region erreicht. Die Besten stellen wir in diesem Issue vor. Die Themen spiegeln – wie immer – die Breite unseres Fachs wieder: Von der Nutzung von Emojis in der Konsumentenkommunikation über die Auswirkungen einer CEO-Dualität in der Schweiz, die Einstellung kulturinteressierter Personen gegenüber Kultursponsoring bis hin zur Messbarkeit von Steuervermeidung. Und nun wünschen wir Ihnen viel Spaß und neue Einsichten beim Studium unserer aktuellen Ausgabe von Junior Management Science.

Ihre

Dominik van Aaken, David Florysiak und Sebastian Waic



### **Junior Management Science**

journal homepage: www.jums.academy



## Die Verwendung von Emojis in der Konsumentenkommunikation – Eine stimmungsanalytische Betrachtung von Kurznachrichten im Social Web

Randi Baum und Tatjana Egelhof *Universität Duisburg-Essen* 

#### Abstract

Social media platforms as enabler for real time and many-to-many communication play an important role in the analysis of consumers' opinions, attitudes, moods, and behaviors towards brands. Emojis as a non-verbal, explanatory and emotional component are increasingly used for a more expressive online communication. While current emotion mining tools only focus on text analysis, we are the first who conduct an automated sentiment analysis of brand-related tweets containing emojis in addition to text. We therefore analyzed 999,197 Starbucks-related and 566,597 McDonald's-related tweets. We used tweets directed at two different global brands in the fast food sector to increase generalizability. On a sentiment polarity scale, the analyzed tweets show a rather positive sentiment value towards Starbucks and a slightly negative sentiment value towards McDonald's. We also find that sentiment is classified identically across brands for 94% of emojis. We conclude that the sentiment value can be considered as an indicator for the perceived image of a brand. Our approach provides an innovative tool for companies to directly analyze emotional content on social media platforms and improves the understanding for the needs of consumers.

The research gap and the scientific approach were predetermined by the Chair of Marketing at Mercator School of Management, University of Duisburg-Essen.

Keywords: Sentiment analysis, Emoji, Twitter, brand, Stimmungsanalyse

### 1. Problemstellung und Aufbau der Arbeit

Die Digitalisierung und ihr Einfluss auf die Entwicklung der Medienlandschaft haben zu einem Wandel in der Kommunikationskultur geführt.<sup>1</sup> Neue Kommunikationskanäle wie das Social Web mit seinen vielfältigen Internetplattformen, ermöglichen es den Nutzern, mit großer Reichweite und in Echtzeit zu kommunizieren, sich auszutauschen und zu informieren.<sup>2</sup> Mit der Kommunikationsart verändert sich zunehmend auch die Kommunikationssprache. Textnachrichten im Social Web werden kürzer und sind u. a. von Neologismen, Slang und schnell als gängig empfundenen Abkürzungen geprägt.<sup>3</sup> Parallel dazu zeigt sich der Trend, begleitend Emoticons und Emojis als eine Art Parasprache<sup>4</sup>

Der Begriff Emoticon setzt sich aus den englischen Wörtern emotion (Gefühl, Gemütsbewegung) und icon (Zeichen) zusammen, ist eine kurze alphanumerische Zeichenfolge und wird als Symbol für einen Gesichtsausdruck verwendet. Ein Emoji, abgeleitet aus dem Japanischen e (Bild) und moji (Zeichen), ist die Weiterentwicklung des Emoticons und steht für ein grafisches Symbol, welches neben Mimik und Gestik auch Objekte, Konzepte und Ideen, wie bspw. Tiere, Pflanzen, Lebensmittel, Fahrzeuge oder Aktivitäten, darstellt. Obwohl Emojis bereits in den frühen 90er Jahren in Erscheinung getreten sind, können sie erst seit dem Jahr 2010 im Unicode-Zeichenformat digital abgebildet werden. Die derzeit aktuelle Unicode-Version 9.0 umfasst insgesamt 1.851 verschiedene

einzusetzen.<sup>5</sup>

<sup>&</sup>lt;sup>1</sup>Vgl. Bruhn (2014): 6f.

<sup>&</sup>lt;sup>2</sup>Vgl. Heymann-Reder (2011): 257.

<sup>&</sup>lt;sup>3</sup>Vgl. Mostafa (2013): 4241; Kiritchenko et al. (2014): 723; Ghiassi et al. (2013): 6268; Go et al. (2009): 2; Boiy und Moens (2009): 529.

<sup>&</sup>lt;sup>4</sup>Die Parasprache umfasst nonverbale Phänomene, welche in der Kommunikation eine wichtige Rolle spielen, da sie bspw. Emotionen übermitteln. Vgl. Traunmüller (2004): 653.

<sup>&</sup>lt;sup>5</sup>Vgl. Dery (1993): 560.

<sup>&</sup>lt;sup>6</sup>Vgl. Duden (2016); Vidal et al. (2016): 120; Novak et al. (2015): 1f.

<sup>&</sup>lt;sup>7</sup>Vgl. Oxford Dictionaries (2016); Novak et al. (2015): 1; Moore et al. (2013): 3.

<sup>&</sup>lt;sup>8</sup>Vgl. Novak et al. (2015): 2.

on ermöglichen. <sup>10</sup> Bereits 92 Prozent der Online-Nutzer verwenden Emojis regelmäßig. <sup>11</sup> Aufgrund der rasanten Verbreitung sowie ihrer hohen gesellschaftlichen Akzeptanz werden Emojis als die am schnellsten wachsende Sprachform der Geschichte betrachtet. <sup>12</sup> Tab. 1 zeigt, wie sich die Kommunikation im Social Web verändert hat. Kurze Kommentare in Form von Textnachrichten wurden zunächst mithilfe von Zeichenabfolgen abgekürzt und werden heute vollständig durch Emojis ersetzt.

Der Wandel der Kommunikationskultur prägt auch die Konsumentenkommunikation und wirkt sich dadurch auf die Marketingpolitik der Unternehmen aus. So hat die starke Interaktion der Konsumenten im Social Web den zuvor dominierenden digitalen Monolog seitens der Unternehmen verdrängt und eine Many-to-Many-Kommunikation etabliert.<sup>13</sup> Die Konsumenten haben eine nie dagewesene Machtstellung erzielt und fungieren als Multiplikatoren, die Meinungen, Einschätzungen sowie Kommentare schnell und weit streuen können.<sup>14</sup> Die digitale Konsumentenkommunikation, insbesondere in Microblogs<sup>15</sup>, kann somit die Wahrnehmung bzw. das Image eines Unternehmens, einer Marke bzw. eines Produkts stark beeinflussen. 16 Für Unternehmen ist daher die laufende Überwachung des Social Webs, das schnelle Erkennen von Stimmungen und Stimmungsänderungen gegenüber dem eigenen Unternehmen sowie gegenüber Marktwettbewerbern von großer Bedeutung und dient im Rahmen der Competitive Intelligence<sup>17</sup> der Erzielung strategischer Wettbewerbsvorteile.<sup>18</sup>

Vor diesem Hintergrund stellt die rasante Entwicklung des Phänomens Emoji sowohl Herausforderungen als auch Chancen für Wissenschaft und Praxis dar. Einerseits bieten Emojis, unter der Annahme, dass bei den Konsumenten ein einheitliches Verständnis für die verwendeten Emojis vorliegt und sich Emotionen so wie Stimmungen diesen eindeutig zuordnen lassen, neue Möglichkeiten und Ansätze für die Emotions- sowie die Marketingforschung. Andererseits ist im Rahmen der Markenkommunikation zu klären, ob Emojis das emotionale Erleben bspw. einer Marke beeinflussen

und einen emotionalen Zusatznutzen schaffen können, der es ermöglicht, sich von funktional gleichwertigen Marken zu differenzieren. <sup>19</sup>

Trotz weltweitem, kulturübergreifendem und sprachenunabhängigem Einsatz sind die derzeitigen Forschungsbemühungen zur Verwendung von Emojis gering. Die vorliegende Arbeit knüpft deshalb an den aktuellen Forschungsstand an und untersucht die Emotionalität ausgewählter Emojis. Sie verfolgt das Ziel zu überprüfen, ob die in Kurznachrichten verwendeten Emojis, einzeln und für sich betrachtet, Rückschlüsse auf eine jeweils gleichgerichtete zugrundeliegende Gefühlslage ermöglichen. Erstmals wird versucht, über die automatische Zuordnung von Kurznachrichten in Emotionsund Stimmungsklassen eine Emoji-Kategorisierung zu entwickeln.

Die vorliegende Arbeit gliedert sich in einen theoretischen und einen empirischen Teil. In Kap. 2 werden zur Herleitung der emotionalen Ausdruckskraft eines Emojis zunächst die theoretischen Grundlagen aus der Emotionsund Stimmungsforschung aufgezeigt. Hierfür werden Ordnungssysteme von Emotionen sowie Methoden zur Emotionsmessung vorgestellt. Anschließend wird die emotionale Bedeutung von Emojis in der computergestützten Kommunikation beleuchtet. Kap. 3 stellt Ansätze vor, um Emotionen und Stimmungen mithilfe von automatischen Verfahren zu identifizieren. Zudem wird die Konzentration der vorliegenden Forschungsarbeit auf den Microblogging-Dienst Twitter erläutert, der aufgrund seiner Popularität, der Präsenz einer großen Online-Gemeinschaft sowie der Möglichkeit der interaktiven und kostenlosen Kommunikation zunehmend auch von Unternehmensseite für die Durchführung von Markenkampagnen genutzt wird.<sup>21</sup> Danach wird auf den bisherigen Forschungsstand zur Emoji-Nutzung in Kurznachrichten, sog. Tweets<sup>22</sup>, eingegangen. Mit Kap. 4 folgt der empirische Teil der Arbeit. Es beschreibt die Aufbereitung der Datengrundlage, die Ableitung eines geeigneten Analyseverfahrens sowie die mit der Software R durchgeführte Datenanalyse. Abschließend fasst Kap. 5 die wichtigsten Ergebnisse der Arbeit zusammen, zeigt Limitationen auf und leitet Implikationen für Wissenschaft und Praxis ab.

### 2. Emojis als emotionaler Ausdruck

Dieses Kap. zeigt Ansätze der Emotionsforschung auf, welche die emotionale Bedeutung von Emojis in der computergestützten Kommunikation beleuchten. Zunächst wird dafür das Verständnis des Emotions- bzw. Stimmungsbegriffs für die vorliegende Arbeit erläutert. Darauf aufbauend werden zwei Ordnungssysteme von Emotionen sowie Methoden

<sup>&</sup>lt;sup>10</sup>Vgl. Emoji Research Team (2015): 21.

<sup>&</sup>lt;sup>11</sup>Vgl. ebd.: 4.

<sup>&</sup>lt;sup>12</sup>Vgl. o. V. (2015).

<sup>&</sup>lt;sup>13</sup>Vgl. Heymann-Reder (2011): 20, 257.

<sup>&</sup>lt;sup>14</sup>Vgl. Beutelsbacher (2011).

<sup>&</sup>lt;sup>15</sup>Microblogs sind Nachrichtendienste, in denen kurze Textnachrichten veröffentlicht werden können. Diese sind auf eine gewisse Zeichenanzahl von meist weniger als 200 Zeichen beschränkt. Die Kurznachrichten können u. a. Erfahrungen sowie Meinungen enthalten und sind einem ausgewählten Nutzerkreis zugänglich. Microblogging bezeichnet das Veröffentlichen eines Microblogs. Vgl. Mattscheck (2016); Kontopoulos et al. (2013): 4065; Jansen et al. (2009a): 3860. Weitere Informationen s. Kap. 3.2.

<sup>&</sup>lt;sup>16</sup>Vgl. Jansen et al. (2009b): 2171f.

<sup>17,</sup> Als "Competitive Intelligence" [...] wird [..] der systematische Prozess der Informationserhebung und -analyse bezeichnet, durch den aus fragmentierten (Roh-)Informationen über Märkte, Wettbewerber und Technologien den Entscheidern ein plastisches Verständnis über ihr Unternehmensumfeld und damit eine Entscheidungsgrundlage geliefert wird." Michaeli (2006): 3.

<sup>&</sup>lt;sup>18</sup>Vgl. Weinberg (2014): 66; Ghiassi et al. (2013): 6266; Kaplan und Haenlein (2011): 108; Pak und Paroubek (2010): 1320; Jansen et al. (2009a): 3864.

<sup>&</sup>lt;sup>19</sup>Vgl. Schweiger und Schrattenecker (2013): 79f; Möll und Esch (2008): 34.

<sup>34. &</sup>lt;sup>20</sup>Vgl. SwiftKey (2015): 1-18; Novak et al. (2015): 2.

<sup>&</sup>lt;sup>21</sup>Vgl. Culotta und Cutler (2016): 345; Kim und Ko (2012): 1480; Barash und Golder (2011): 146.

<sup>&</sup>lt;sup>22</sup>Ein Tweet ist eine Kurznachricht, die über den Microblogging-Dienst Twitter veröffentlicht wird. Vgl. Heymann-Reder (2011): 258. Weitere Informationen s. Kap. 3.2.

Tabelle 1: Emojis verändern die Kommunikation (Quelle: in Anlehnung an Emoji Research Team (2015): 6)

Kommentar	Kürzel	Emoji
That's funny I love you	haha/lol Ily	<b>≅</b>
I like it	Like	

zur Emotionsmessung vorgestellt, wobei die Messung des emotionalen Ausdrucksverhaltens im Vordergrund steht.

In der Literatur herrscht kein Konsens hinsichtlich der Definition des Begriffs Emotion, da je nach Strömung der Emotionsforschung inhaltlich unterschiedliche Aspekte fokussiert werden. Häufig wird jedoch auf die Arbeit von Kleinginna und Kleinginna verwiesen, aus der sich nach einer Analyse von über 90 Emotionsdefinitionen die folgende Arbeitsdefinition<sup>24</sup> herausgebildet hat: <sup>25</sup>

"Emotion ist ein komplexes Interaktionsgefüge subjektiver und objektiver Faktoren, das von neuronal/hormonalen Systemen vermittelt wird, die (a) affektive Erfahrungen, wie Gefühle der Erregung oder Lust/Unlust, bewirken können; (b) kognitive Prozesse, wie emotional relevante Wahrnehmungseffekte, Bewertungen, Klassifikationsprozesse, hervorrufen können; (c) ausgedehnte physiologische Anpassungen an die erregungsauslösenden Bedingungen in Gang setzen können; (d) zu Verhalten führen können, welches oft expressiv, zielgerichtet und adaptiv ist."<sup>26</sup>

Die vorliegende Arbeit folgt diesem viel zitierten Emotionsverständnis, da Emojis in der computergestützten Kommunikation zum einen die von Emotionen ausgelösten physiologischen Anpassungen, wie bspw. Erröten, Schwitzen oder eine Veränderung der Pupillengröße, darstellen können. Zum anderen drücken sie das durch Emotionen verursachte und beobachtbare Ausdrucksverhalten eines Menschen aus, zu dem u. a. der Gesichtsausdruck, die Gestik oder die Körperhaltung zählen. <sup>27</sup>

Zu einer Abgrenzung der Termini Emotion und Stimmung findet sich in der Literatur ebenfalls kein Konsens. Tendenziell wird davon ausgegangen, dass Stimmungen eine Gesamtbefindlichkeit wiedergeben. <sup>28</sup> Sie dauern länger an und

werden aufgrund ihrer geringeren Intensität oftmals als Low-level-Emotionen bezeichnet. Da enge Wechselwirkungen zwischen Emotionen und Stimmungen bestehen, werden die Begrifflichkeiten sowohl in der Literatur als auch in dieser Arbeit synonym verwendet.

#### 2.1. Ordnungssysteme der Emotionen

Zur Strukturierung von Emotionen haben sich zwei Ordnungsschemata herauskristallisiert, die im Folgenden vorgestellt werden. Es handelt sich dabei um den dimensionalen sowie den differentiellen Ansatz. Während der dimensionale Ansatz eher Anwendung in der Umweltpsychologie findet, dominiert der differentielle Ansatz die Emotionspsychologie und ist im Rahmen der Betrachtung der evolutionspsychologischen Emotionstheorie von Bedeutung. 32

#### Dimensionaler Ansatz

Das Ziel des dimensionalen Ansatzes ist es, allgemeine Dimensionen zu identifizieren, mit denen sich jede beliebige Emotion charakterisieren lässt. 33 Die Vertreter dieses Emotionsansatzes sind bestrebt, möglichst wenige, voneinander unabhängige Dimensionen für die Unterscheidung der Emotionen zu formulieren. 34 Die dimensionalen Ansätze unterscheiden sich in uni- und multidimensionale Modelle. Unidimensionale Modelle, welche sich nur auf eine zentrale Dimension der Emotion konzentrieren, haben in der Vergangenheit jedoch an Bedeutung verloren.<sup>35</sup> Die Vertreter des zweidimensionalen Ansatzes postulieren, dass Emotionen auf den beiden bipolaren Dimensionen Valenz und Erregung bzw. Aktivierung erlebt werden.<sup>36</sup> Die Valenz beschreibt, ob die Emotion als angenehm oder unangenehm empfunden wird und führt dementsprechend entweder zu einem Annäherungs- oder einem Vermeidungsverhalten. Die Intensität der Erregung bestimmt, wie viele objektiv messbare energetische Ressourcen für das Verhalten bereitgestellt werden und ob die Emotion dementsprechend als erregend

<sup>&</sup>lt;sup>23</sup>Vgl. Izard (2010): 363, 367, 369; Winder (2007): 25; Merten (2003): 12; Otto et al. (2000): 11f; Meyer et al. (1993): 22f; Scherer (1990): 3.

<sup>&</sup>lt;sup>24</sup>Eine Arbeitsdefinition umschreibt grob das Forschungsgebiet und stellt einen Verständigungsrahmen für die zu untersuchenden Phänomene bereit. Vgl. Otto et al. (2000): 12.

<sup>&</sup>lt;sup>25</sup>Vgl. Izard (2010): 363; Winder (2007): 25; Merten (2003): 12; Otto et al. (2000): 14; Scherer (1990): 3.

 $<sup>^{26}</sup>$ Kleinginna und Kleinginna (1981): 355. Übersetzt von Otto et al. (2000): 15.

<sup>&</sup>lt;sup>27</sup>Vgl. Meyer et al. (1993): 30.

<sup>&</sup>lt;sup>28</sup>Vgl. Ewert (1983): 399.

<sup>&</sup>lt;sup>29</sup>Vgl. Merten (2003): 11; Otto et al. (2000): 13; Meyer et al. (1993): 34; Scherer (1990): 6; Isen (1984): 185.

<sup>&</sup>lt;sup>30</sup>Vgl. Parrott (2001): 4; Schönpflug (2000): 19; Schmidt-Atzert (2000): 33; Meyer et al. (1993): 22; Scherer (1990): 6; Isen (1984): 185; Plutchik (1980): 91.

<sup>&</sup>lt;sup>31</sup>Vgl. Behrens (1991): 66.

<sup>&</sup>lt;sup>32</sup>Vgl. Merten (2003): 19; Neibecker (1985): 15.

<sup>&</sup>lt;sup>33</sup>Vgl. Neibecker (1985): 13.

<sup>&</sup>lt;sup>34</sup>Vgl. Merten (2003): 19; Neibecker (1985): 13.

<sup>&</sup>lt;sup>35</sup>Vgl. Zentner und Scherer (2000): 152f.

<sup>&</sup>lt;sup>36</sup>Vgl. Pauli und Birbaumer (2000): **78**; Bradley und Vrana (1993): **271**; Ekman et al. (1982): **50-52**; Schmidt-Atzert (1981): **39**.

oder deaktivierend empfunden wird.<sup>37</sup> Auch wenn drei- oder vierdimensionale Modelle weitere Dimensionen, wie bspw. Dominanz (überlegen/unterlegen)<sup>38</sup>, Qualität<sup>39</sup> oder Bewusstsein<sup>40</sup>, aufführen, haben sich in der Emotionsliteratur hauptsächlich die beiden Dimensionen Valenz und Erregung durchgesetzt.<sup>41</sup>

Anhand der Dimensionen lässt sich ein Koordinatensystem bilden, in dem die unterschiedlichen emotionalen Reaktionen verortet werden können.<sup>42</sup> Damit gehen Emotionen fließend ineinander über.<sup>43</sup> Die nachfolgende Tab. 2 verdeutlicht den zweidimensionalen Ansatz am Beispiel der Dimensionen Valenz und Erregung.

Der Vorteil des dimensionalen Ansatzes ist, dass die Messung, Auswertung und Interpretation von Emotionen durch die geringe Anzahl an Dimensionen vereinfacht wird. 44 Trotz ihrer Plausibilität und ihrer empirischen Fundierung gibt es jedoch deutliche Kritikpunkte. So sind die dimensionalen Modelle zur Beschreibung von Zuständen nur eingeschränkt geeignet. Die beiden Emotionen Angst und Ärger weisen, trotz deutlich verschiedener Emotionsqualität, eine negative Valenz bei hoher Erregung auf und fallen damit im zweidimensionalen Modell auf fast denselben Koordinatenpunkt. Die beiden Emotionen können jedoch unterschiedliche Ursachen haben und zu verschiedenen Reaktionen im Konsumentenverhalten führen.<sup>45</sup> Thayer kritisiert zudem die Erregungsdimension des dimensionalen Ansatzes und argumentiert, dass sie sich aus zwei unterschiedlichen Aspekten zusammensetzt und damit keine einheitliche Dimension widerspiegelt. 46 Auch die Valenzdimension steht unter Kritik, weil sie sich auf verschiedene Komponenten, wie bspw. auf die Bewertung eines Ereignisses, eines Objekts, der Handlungstendenz oder der Gefühlskomponente beziehen kann. 47 Ein weiterer Nachteil der dimensionalen Ansätze ist, dass sie im Alltag häufig auftretende Emotionen wie z. B. Liebe nicht abbilden können. 48 Trotz dieser Kritik ist der dimensionale Ansatz für die vorliegende Arbeit von Bedeutung, da hier über die bipolare Dimension Valenz Bezug auf die Polarität (positiv/negativ) einer Emotion und damit auf die möglichen Kategorisierung eines Emojis genommen werden kann.

Differentieller Ansatz

Die Vertreter des differentiellen Emotionsansatzes, in der Literatur auch als Klassifikationsansatz bezeichnet<sup>49</sup>, postulieren, dass es eine begrenzte Anzahl an klar differenzierbaren, fundamentalen Emotionen gibt. Diese sog. Primäremotionen, auch Grund- oder Basisemotionen genannt, äußern sich jeweils in einem spezifischen physiologischen, ausdrucks- und verhaltensmäßigen Reaktionsmuster. In der Literatur finden sich je nach Vertreter der Theorie zwei bis 18 Primäremotionen, wobei Angst, Ärger, Traurigkeit und Freude am häufigsten genannt werden. Stellt auszugsweise die Ansätze verschiedener Basisemotionstheoretiker in chronologischer Reihenfolge ihrer Veröffentlichung gegenüber.

Für die Heterogenität der Emotionslisten nennt die Literatur zwei mögliche Gründe. Einerseits fehlt, wie einleitend bereits erwähnt, eine einheitliche Begriffsdefinition der Emotion. Andererseits ergeben sich in Abhängigkeit von attributionstheoretischen, sprachpsychologischen, neuroanatomischen oder auch evolutionstheoretischen Begründungen Unterschiede in der Zuordnung von Emotionen zu Primäremotionen.<sup>53</sup>

Trotz der verschiedenen Auffassungen über die Art und Anzahl an Primäremotionen sind sich die Basisemotionstheoretiker in zwei Annahmen einig. Zum einen teilen sie die Auffassung, dass Basisemotionen evolutionspsychologisch verankert sind. Zum anderen gehen sie davon aus, dass Primäremotionen eine psychologische Grundlage bilden und aus ihrer Mischung komplexe Emotionen, sog. Sekundäremotionen, abgeleitet werden können. HCDougall schlägt erstmalig vor, Emotionen in einer kreisförmigen Darstellung analog zum Farbkreis abzubilden. Prämäremotionen Plutchik die von ihm identifizierten acht Primäremotionen Freude, Akzeptanz, Angst, Überraschung, Traurigkeit, Ekel, Ärger und Erwartung im sog. Rad der Emotionen an, das in Abb.1 dargestellt ist.

Das geometrische Design dient der Beschreibung der Beziehungen zwischen den Emotionen. Die Primäremotionen sind umso qualitativ unterschiedlicher, je weiter sie im Rad der Emotionen voneinander entfernt angeordnet sind. Damit verringert sich auch die Wahrscheinlichkeit einer erfolgreichen Vermischung. Im Rad gegenüberliegende Emotionen neutralisieren sich gegenseitig oder können einen schweren inneren Konflikt (C) hervorrufen. Je nach Ähnlichkeit der sich verbindenden Primäremotionen werden die Sekundäremotionen sog. Dyaden zugeordnet. <sup>56</sup> Tab. 4 stellt die Unterschiede zwischen den primären, sekundären und tertiären Dyaden übersichtlich dar.

Für die vorliegende Arbeit ist auch das Ordnungsschema des differentiellen Ansatzes von Bedeutung, da wie oben aufgeführt die Auffassung besteht, dass Basisemotionen evolutionspsychologisch verankert sind. In Anlehnung an die Evo-

<sup>&</sup>lt;sup>37</sup>Vgl. Pauli und Birbaumer (2000): 78; Weinberg (1986): 10.

<sup>&</sup>lt;sup>38</sup>Vgl. Bradley und Lang (1994): 49; Mehrabian und Russell (1974): 18f.

 $<sup>^{39} \</sup>rm Die$  Qualität drückt die Art und die subjektive Bedeutung einer Emotion, also den Erlebnisinhalt, aus. Vgl. Weinberg (1986): 10.

<sup>&</sup>lt;sup>40</sup>Vgl. Kroeber-Riel und Weinberg (1996): 105.

<sup>&</sup>lt;sup>41</sup>Grund hierfür ist u. a. die zweifelhafte bzw. extrem geringe Varianzaufklärung. Vgl. Haimerl (2007): 20; Averill (1997): 518.

<sup>&</sup>lt;sup>42</sup>Vgl. Stürmer und Schmidt (2014): 51.

<sup>&</sup>lt;sup>43</sup>Vgl. Averill (1997): 517.

<sup>44</sup>Vgl. Mau (2009): 21; Neibecker (1985): 15.

<sup>&</sup>lt;sup>45</sup>Vgl. Keltner et al. (2014): 169; Watson und Spence (2007): 498; Schmidt-Atzert (2000): 40.

<sup>&</sup>lt;sup>46</sup>Vgl. Thayer (1989): 6-8.

<sup>&</sup>lt;sup>47</sup>Vgl. Zentner und Scherer (2000): 153.

<sup>&</sup>lt;sup>48</sup>Vgl. Franke (2014): 8.

<sup>&</sup>lt;sup>49</sup>Vgl. Averill (1997): 515; Neibecker (1985): 12.

 $<sup>^{50}</sup>$ Vgl. Zentner und Scherer (2000): 154, 156; Plutchik (1980): 129.

<sup>&</sup>lt;sup>51</sup>Vgl. Zentner und Scherer (2000): 156.

<sup>&</sup>lt;sup>52</sup>Vgl. Schmidt-Atzert (2000): 30.

<sup>&</sup>lt;sup>53</sup>Vgl. Winder (2007): 50f; Schmidt-Atzert (2000): 30f.

<sup>&</sup>lt;sup>54</sup>Vgl. Winder (2007): 48, 50, 60; Schmidt-Atzert (2000): 31; Plutchik (1980): 129.

<sup>&</sup>lt;sup>55</sup>Vgl. McDougall (1924): **45**f.

<sup>&</sup>lt;sup>56</sup>Vgl. Plutchik (1980): 160-163.

Tabelle 2: Verdeutlichung des dimensionalen Ansatzes (Quelle: in Anlehnung an Stürmer und Schmidt (2014): 59)

		Val	enz
Erregung	erregend deaktivierend	angenehm Freude Entspannung	unangenehm Angst/Ärger Trauer

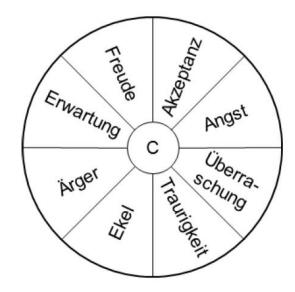


Abbildung 1: Primäremotionen im Rad der Emotionen (Quelle: in Anlehnung an Plutchik (1980): 160)

lutionstheorie nach Darwin wird postuliert, dass viele Emotionsausdrücke aufgrund ihres Reproduktionsvorteils kulturübergreifend angeboren und damit ungelernt sind.<sup>57</sup> Einige Forschungsergebnisse belegen die Universalität der Basisemotionen.<sup>58</sup> Izard, Ekman und Tomkins teilen die Auffassung, dass Basisemotionen wie Freude, Angst, Trauer, Überraschung und Ekel einen emotionsspezifischen, angeborenen Gesichtsausdruck aufweisen.<sup>59</sup> Es scheint daher auf Basis des differentiellen Ansatzes möglich zu sein, Emojis mit Gesichtsausdruck entsprechenden Emotionen zuzuordnen.

#### 2.2. Verfahren zur Emotionsmessung

In Einklang mit der in Kap. 2 aufgeführten Arbeitsdefinition nennt Izard drei Verhaltensebenen, auf denen sich Emotionen manifestieren: die neurophysiologischen Vorgänge, das subjektive Erlebnis sowie das beobachtbare Ausdrucksverhalten, insbesondere die Mimik. 60 Dementsprechend werden auch die Verfahren zur Emotionsmessung nach diesen drei Ebenen eingeteilt. Für die vorliegende Arbeit steht insbesondere die Messung des beobachtbaren Ausdrucksverhaltens im Vordergrund. Zur Vollständigkeit werden jedoch

auch die Ansätze zu den erstgenannten Verhaltensebenen skizziert

Anhand der (neuro-)physiologischen Messung wird die Intensität der emotionalen Erregung bestimmt. Dabei werden oftmals bioelektrische Signale gemessen, die von der Körperoberfläche abgeleitet werden.<sup>62</sup> Die gängigsten Verfahren sind die Messung der Herzfrequenz, des Blutdrucks und Blutvolumens, die Messung der Atmung, des Hautwiderstandes und der Gehirnwellen sowie computer- und kernspintomographische Verfahren. 63 Die (neuro-)physiologischen Verfahren weisen jedoch einige Nachteile auf. Ihre Durchführung erfordert technisches Know-how, wird überwiegend im Labor und damit nicht in einem lebensnahen Umfeld durchgeführt und verursacht bei oftmals nur kleinen Stichprobengrößen hohe Kosten.<sup>64</sup> Zudem wird kritisiert, dass die Bestimmung der Intensität der inneren Erregung nicht ausreicht, um eine Emotion zu beschreiben.<sup>65</sup> Dennoch werden psychobiologische Verfahren aufgrund ihrer Objektivität zur Validierung anderer Messverfahren, vor allem von Skalen zur subjektiven Einschätzung der Gefühle, eingesetzt und finden zudem Anwendung in der apparativen Werbeforschung. 66

<sup>&</sup>lt;sup>57</sup>Vgl. Winder (2007): 36; Schmidt-Atzert (2000): 31; Euler (2000): 47; Plutchik (1980): 1, 5; Darwin (1871).

<sup>&</sup>lt;sup>58</sup>Vgl. Kaiser und Wehrle (2000): 420.

<sup>&</sup>lt;sup>59</sup>Vgl. Izard (1991), Ekman (1972) und Tomkins (1962) zitiert von Winder (2007): 51f.

<sup>60</sup>Vgl. Izard (1994): 20.

<sup>&</sup>lt;sup>61</sup>Vgl. Kroeber-Riel und Weinberg (1999): 101, 106.

<sup>&</sup>lt;sup>62</sup>Vgl. Weinberg (1986): 11.

<sup>&</sup>lt;sup>63</sup>Einen guten Überblick über die einzelnen Verfahren gibt Kenning (2014): 87f.

 $<sup>^{64}\</sup>mbox{Vgl.}$  Kroeber-Riel und Gröppel-Klein (2013): 141; Weinberg (1986): 11.  $^{65}\mbox{Vgl.}$  Weinberg (1986): 11.

<sup>&</sup>lt;sup>66</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 76, 89f; Wildner und Jäncke (2010): 95f; Schiel (2007): 171.

Tabelle 3: Übersicht der Primäremotionen nach verschiedenen Autoren (Quelle: in Anlehnung an Liu (2015): 33; Kroeber-Riel und Gröppel-Klein (2013): 114; Plutchik (2003): 73 und Weinberg (1986): 41)

Basisemotionstheorie nach							Basis	Basisemotionen							Weitere aufgeführ- te Basisemotionen
McDougall (1021)	Angst	Ärger	Ekel	Traurigkeit	Freude	Überraschung	Verachtung	Interesse	Scham	Schuld/Reue	Erwartung	Liebe	Sorge	Zufriedenheit	Unterwirfickeit
McDongan (1721)	>	Þ	>	>		>									Zärtlichkeit, Hoch- gefühl
Plutchik (1962, 1980)	×	×	×	×	×	×					×				Akzeptanz
Tomkins (1962, 1963)	×	×	×	×	×	×	×	×	×						
Osgood (1966)	×	×	×		×	×		×			×			×	Langeweile
Arieti (1970)	×	×												×	Anspannung, Ver-
Izard (1972, 1977, 1994)	×	×	×	×	×	×	×	×	×	×					ım Şen
Ekman (1973)	×	×	×	X	×	×									
Ekman/Friesen (1975)	×	×	×	×	×	×		×							
Sroufe (1979)	×	×			×										
Etude (1980)	×	×	×	×	×	×		×	×	×					Schüchternheit,
Scott (1980)	×	×			×							×	×		Einsamkeit, Neu-
Fromme/O'Brien (1982)	×	×		×	×									×	giei de Begeisterung, Schook
Malatesta/ Haviland (1982)	×	×		×	×			×							Schmerz
Panksepp (1982)	×	×									×				Panik
Epstein (1984)	×	×		×	×							×			
Shaver/Schwartz (1984)	×	×		×	×							×			
Trevarthen (1984)	×	×		×	×										
Fehr/Russell (1985)	×	×		×	×							×			
Plutchik (1991)	×	×	×	×	×	×									
Parrott (2001)	×	×													

Tabelle 4: Übersicht der Sekundäremotionen (Quelle: in Anlehnung an Plutchik (1980): 161f)

	Primäre Dyaden	Sekundäre Dyaden	Tertiäre Dyaden
Sekundäremotionen	Liebe, Unterwerfung, Ehrfurcht, Enttäuschung, Reue, Verachtung, Aggressivität, Optimismus	Schuld, Neugierde, Verzweiflung, Neid, Zynismus, Stolz, Fatalismus	Vergnügen, Scham, Resignation, Empörung, Pessimismus, Krank- haftigkeit, Dominanz, Feigheit
Entstehen durch die Vermischung von	Zwei unmittelbar benachbarten Primäremotionen.	Zwei Primäremotionen, die durch eine dazwischenliegende Emotion getrennt sind.	Zwei Primäremotionen, die durch zwei dazwischenliegende Emotio- nen getrennt sind.
Beispiel	Freude + Akzeptanz $\rightarrow$ Liebe	Freude + Angst $\rightarrow$ Schuld	Freude + Überraschung $\rightarrow$ Vergnügen

Die subjektive Erlebnismessung kann verbal oder nonverbal erfolgen und bezieht sich auf die Messung bewusster Emotionen. Verbale Verfahren werten sprachliche Äußerungen von Personen über ihre Gefühle aus. Dazu zählen Protokolle lauten Denkens, Tagebücher sowie verbale Skalen. Erstere haben jedoch den Nachteil, dass Probanden ihre wahren Emotionen, insbesondere bei Themen mit hohem sozialen Potenzial, wie bspw. Diskriminierung oder Sexualität, verschleiern oder kognitiv filtern und damit das Messergebnis verzerren. Zudem treten durch die verbale Beschreibung von Emotionen oftmals Artikulationsschwierigkeiten auf, da die Sprache über eine Vielzahl an unterschiedlichen Wörtern verfügt, um eine Emotion auszudrücken. Letztlich lassen die verbalen Beschreibungen zudem Raum für die Ergebnisinterpretation.

Mithilfe verbaler Skalen wird gemessen, wie wünschenswert ein Proband ein gewisses Ziel einschätzt und den Zielerreichungsgrad bewertet. Daraus konzeptualisiert sich die Emotion. Zu den in der Marktforschung sowie in der biologisch orientierten Emotionsforschung bekanntesten Skalen zählen

- die Pleasure/Arousal/Dominance-Skala (PAD) von Merhabian und Russell (1974),
- die Differenzielle Emotionsskala (DES) von Izard (1977),
- die PANAS-Skala von Watson, Clark und Tellegen (1988),
- der Emotions Profile Index (EPI) von Plutchik (2003),
- das Consumption Emotion Set (CES) von Richins (1997) sowie

Der Vorteil der Selbstauskunftsskalen ist, dass sie die Stärke, Richtung und Qualität der Emotion ermitteln können.  $^{72}$  Zudem ist ihre Durchführung wenig zeitintensiv und kostengünstig. Trotz ihrer Popularität finden sich in der Literatur jedoch auch Kritikpunkte und Verbesserungspotenziale.  $^{73}$ 

Nonverbale Messungen werden durchgeführt, wenn Emotionen sprachlich nicht formulierbar sind oder dem Probanden zunächst in das Bewusstsein gebracht werden müssen. Programmanalysator verwendet, der spontane emotionale Eindrücke während einer Programmdarbietung, wie bspw. eines Werbespots, feststellt. Zum anderen finden Bilderskalen Anwendung. Hier werden die Items zur Messung von Emotionen nicht verbal, sondern bildhaft vorgegeben. Tab. 5 stellt die drei populärsten Bilderskalen übersichtlich dar.

In der Literatur finden sich zahlreiche Vorteile von Bilderskalen zur Emotionsmessung:

- Bilder schaffen einen leichteren Zugang zu verbal schwierig ausdrückbaren und nicht unmittelbar bewussten Empfindungen.
- Nonverbale Messverfahren reduzieren die Gefahr von sozial erwünschten Antworten, da sie von den Probanden weniger leicht durchschaut werden.
- Bilderskalen erfassen schwach empfundene Eindrücke besser und können daher in Low-Involvement-Situationen eingesetzt werden.<sup>77</sup>
- Bildbasierte Skalen können das Interesse und die Motivation der Probanden positiv beeinflussen und damit auch die Antwortquote erhöhen.
- Bilderskalen können interkulturell angewendet werden.<sup>78</sup>

<sup>&</sup>lt;sup>67</sup>Vgl. Kroeber-Riel und Weinberg (1999): 105, 107.

<sup>&</sup>lt;sup>68</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 128.

<sup>&</sup>lt;sup>69</sup>Vgl. Weinberg (1986): 11.

<sup>&</sup>lt;sup>70</sup>Vgl. Leahu et al. (2008): 425f.

<sup>&</sup>lt;sup>71</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 128f.

<sup>&</sup>lt;sup>72</sup>Vgl. ebd.: 141.

<sup>&</sup>lt;sup>73</sup>Mehr dazu in Ciuk et al. (2015): 1; Richins (1997): 129f.

<sup>&</sup>lt;sup>74</sup>Vgl. Kroeber-Riel und Weinberg (1999): 110.

<sup>&</sup>lt;sup>75</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 130f.

<sup>&</sup>lt;sup>76</sup>Vgl. Weinberg (1986): 32f.

<sup>&</sup>lt;sup>77</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 133.

<sup>&</sup>lt;sup>78</sup>Vgl. Dieckmann et al. (2008): 77.

**Tabelle 5:** Übersicht populärer Bilderskalen (Quelle: in Anlehnung an Kroeber-Riel und Gröppel-Klein (2013): 131-134; Dieckmann et al. (2008): 83f; Berger und Heath (2007): 123; Bradley und Lang (1994): 53-58)

Skala	Deskriptoren und Dimensionen bzw. Emotionen	Vor- (+) und Nachteile (-)
Self-Assessment Manikin- Skala (SAM) von Lang (1980) und Morris (1995)	Deskriptoren: 15 Piktogramme (je fünf pro Dimension). Dimensionen: Valenz, Erregung, Dominanz.	<ul><li>+ Kürzere Erhebungszeit als verbales</li><li>Pendant PAD.</li><li>- Piktogramme werden von Probanden teilweise nicht verstanden.</li></ul>
PrEmo-Verfahren von Desmet (2002)	Deskriptoren: 14 animierte Piktogramme (für sieben positive und negative Emotionen). Emotionen: Belustigung, Bewunderung, Inspiration, Zufriedenheit, Faszination, Verlangen, angenehme Überraschung, unangenehme Überraschung, Langeweile, Unzufriedenheit, Enttäuschung, Ärger, Ekel, Verachtung.	<ul> <li>+ Animation verdeutlicht die Entwicklung der Emotion.</li> <li>- Nur beschränkte Anzahl an Emotionen messbar.</li> <li>- Nur computergestützte Anwendung.</li> </ul>
GfK EMO Sensor von Dieck- mann2008beyond	Deskriptoren: 1-3 Bilder inkl. Bezeichnung für jede der 22 Emotionen. Emotionen: Behaglichkeit, Interesse, Neugier, Freude, Überraschung, Begehren, Leidenschaft, Stolz, Vertrauen, Liebe, Langeweile, Enttäuschung, Scham, Traurigkeit, Mitleid, Ablehnung, Neid, Ekel, Angst, Wut, Aggressivität, Überlegenheit.	- Aufgrund von Eigen- und Fremd- gruppeneffekten keine kulturüber- greifende Anwendung.

 Im Gegensatz zu Verbalskalen können Bilderskalen bei sprachlich gering entwickelten Bevölkerungsgruppen, wie bspw. Kindern, angewendet und ohne Übersetzungsfehler in andere Sprachräume übertragen werden.<sup>79</sup>

Jedoch weisen Bilderskalen im Allgemeinen den Nachteil auf, dass die Mimik der Bilder bzw. das Zusammenspiel zwischen dem Gesichtsausdruck und der Gestik mehrdeutig wirken kann. Damit ergeben sich Risiken der Fehlinterpretation und der verzerrten Messung von Emotionen. B1

Die Messung des emotionalen Ausdrucksverhaltens umfasst die gesamte Körpersprache eines Menschen. Während die Gestik, Körperhaltung und der stimmliche Ausdruck jedoch kaum untersucht werden, findet die Gesichtssprache besondere Beachtung. Neben der kommunikativen Funktion der Mimik, Informationen auszutauschen und Interaktionen zu regulieren, kann das menschliche Gesicht emotionale Vorgänge in schneller Abfolge widerspiegeln.82 Ergänzend ist hier auch die Wechselbeziehung zwischen Emotionen und Mimik anzumerken. So vermutete bereits Darwin, dass der Gesichtsausdruck für die Entstehung von Emotionen eine entscheidende Rolle spielt.<sup>83</sup> Im Rahmen von Experimenten zur Facial-Feedback-Hypothese konnte bspw. nachgewiesen werden, dass ein bewusst erzeugter emotionaler Gesichtsausdruck eine entsprechende emotionale Reaktion auslöst.84 Zudem verstärkt bzw. verringert ein bewusst

freundlicher Gesichtsausdruck die Bewertung von Freudeinduzierenden bzw. Ärger-induzierenden Reizen. <sup>85</sup> Die in der Emotionsforschung populärsten Verfahren zur Messung des emotionalen Ausdrucksverhaltens in Bezug auf die Gesichtssprache sind in Tab. 6 dargestellt.

Besondere Aufmerksamkeit in der Literatur erfährt das FACS, ein von Ekman und Friesen entwickeltes umfassendes Kategoriensystem. Es basiert auf der Kernidee, dass jede Emotion mit typischen Muskelbewegungen im Gesicht einhergeht. Die theoretisch fundierte und zuverlässige Methode ermöglicht eine detaillierte Analyse der Gesichtsmotorik. Dafür werden alle visuell unterscheidbaren Aktivitäten der Stirn, Wangen, Nase und des Kinns, vor allem aber der Augenlider, Augenbrauen und Lippen in 33 anatomisch bedingte, kleinste Bewegungseinheiten (Action Units 2) zerlegt. Im Jahr 2002 erweiterten die Autoren das FACS auf 44 AU's.

AU's können sowohl von einem als auch von mehreren Gesichtsmuskeln ausgelöst werden und entweder nur in einer oder in beiden Gesichtshälften auftreten. Einigen der in Tab. 3 genannten Basisemotionen nach Ekman und Friesen werden Kombinationen von AU's zugeordnet, während z. B. Freude durch ein einziges AU ausgedrückt werden kann. Sekundäremotionen hingegen können nicht aus den AU's ab-

<sup>&</sup>lt;sup>79</sup>Vgl. Bradley und Lang (1994): 50.

<sup>&</sup>lt;sup>80</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 134.

<sup>&</sup>lt;sup>81</sup>Vgl. Dieckmann et al. (2008): 77.

 $<sup>^{82}</sup>$ Vgl. Kroeber-Riel und Gröppel-Klein (2013): 134; Kaiser und Wehrle (2000): 419f.

<sup>83</sup> Vgl. Darwin (1871) zitiert in Pauli und Birbaumer (2000): 76.

<sup>&</sup>lt;sup>84</sup>Vgl. Mori und Mori (2009): 76; Rutledge und Hupka (1985): **219**f; Ekman et al. (1983): **1208**.

<sup>&</sup>lt;sup>85</sup>Vgl. Strack et al. (1988): 768.

<sup>&</sup>lt;sup>86</sup>Vgl. Kaiser und Wehrle (2000): 423.

<sup>&</sup>lt;sup>87</sup>Vgl. Kroeber-Riel und Gröppel-Klein (2013): 135.

<sup>&</sup>lt;sup>88</sup>Vgl. Weinberg (1986): 37-39.

<sup>&</sup>lt;sup>89</sup>Die einzelnen AU's können nachgelesen werden in Ekman und Friesen (1978): 8.

<sup>&</sup>lt;sup>90</sup>Vgl. Ekman und Friesen (2003): 1; Kaiser und Wehrle (2000): 423; Weinberg (1986): 37; Ekman und Friesen (1978): 7f.

<sup>&</sup>lt;sup>91</sup>Vgl. Ekman et al. (2002).

<sup>&</sup>lt;sup>92</sup>Vgl. Kaiser und Wehrle (2000): 423; Ekman und Friesen (1978): 7.

**Tabelle 6:** Überblick der Verfahren zur Messung des Ausdrucksverhaltens (Quelle: in Anlehnung an Merten (2003): 31; Kaiser und Wehrle (2000): 421-426; Pauli und Birbaumer (2000): 76; Weinberg (1986): 43-45)

Methode	Beschreibung	Vor- (+) und Nachteile (-)
Elektromyographische Verfahren (EMG) nach Schwatz (1977)	Messung der nichtsichtbaren mimischen Muskelaktivität, die bei der mentalen Vorstellung emotionaler Erlebnisse auftritt. Die Versuchsperson ist dabei verkabelt.	<ul> <li>+ Die emotionale Reaktion kann weniger willentlich beeinflusst werden als der offen beobachtbare Gesichtsausdruck.</li> <li>+ Echtes und vorgetäuschtes Lächeln lassen sich differenzieren.</li> <li>- Messung beschränkt sich auf die beiden Muskeln "Lächelmuskel" und "Augenbrauen-zusammenziehen".</li> <li>- Kleine Bewegungen des Probanden führt zu Störsignalen.</li> </ul>
Kodierverfahren	Mimisches Verhalten wird per Video aufgezeichnet und in Zeitlupe analysiert.	+ Durch die detaillierte Analyse des mimischen Verlaufsmusters können objektiv beobachtbare Beschreibungsmaße definiert und die Kodiererübereinstimmung bestimmt werden Zeitintensives Verfahren, daher sind die Kodiersysteme selektiv.
Facial Action Coding System (FACS) nach Ekman/Friesen (1978)	Analyse der Gesichtsmotorik anhand einzelner Bewegungseinheiten, sog. Action Units (genauere Beschreibung folgt ab S. 15)	<ul> <li>+ Theoretische Fundierung.</li> <li>+ Zuverlässiges Messverfahren.</li> <li>- Hoher Erhebungsaufwand.</li> <li>- Intensität der Mimik wird nicht getrennt für beide Gesichtshälften erfasst.</li> </ul>
Emotional Facial Action Coding System (EM-FACS) nach Friesen/Ekman (1984)	Auf Basis des FACS werden nur die- jenigen Gesichtsausdrücke erfasst, die am stärksten zur Identifikation einer Emotion beitragen.	<ul> <li>+ Reduzierte Kodierungszeit.</li> <li>+ Schnelles und einfaches Erkennen von Emotionskategorien durch reduzierte Datenmenge.</li> <li>- Ungenauere Ergebnisse als FACS, da bestimmte AU's vernachlässigt werden.</li> <li>- Begrenzte Transparenz der theoretischen Fundierung für die Zuordnung von mimischen Ausdrucksmustern zu den Basisemotionen.</li> </ul>
Maximally Discriminative Facial Movement Coding System nach Izard (1979)	Kodierung mimischen Verhaltens bei Säuglingen und Kleinkindern.	- Es werden nur diejenigen Ausdrucksmuster untersucht, die Izard aufgrund interkultureller Studien als prototypisch er- klärt hat.
Facial Expression Analysis Tool (FEAT) nach Kaiser/Wehrle (1992)	Auf Basis des FACS erfolgt die automatische Kodierung der Gesichtsmimik in einer Mensch-Computer-Interaktion.	<ul> <li>+ Automatische und damit objektive Mimikkodierung.</li> <li>+ Intensität der Mimik wird für beide Gesichtshälften getrennt erfasst.</li> <li>- Für den therapeutischen Bereich ungeeignet, da das Verfahren auf künstliche Markierungen angewiesen ist.</li> </ul>
Self-Evaluative Emotions Coding System (SEECS) nach Geppert et al. (1997)	Messung selbstreflexiver Emotionen von Kindern in einer Wettbewerbs- situation. Neben der Mimik werden auch Blickverhalten, Gestik und Kör- perhaltung erhoben.	- Anwendung nur in spezifischer experimenteller Umgebung.

geleitet werden. <sup>93</sup> Bei einzelnen AU's werden je nach Ausprägung fünf unterschiedliche Intensitätsgrade kodiert. Es wird jedoch kritisiert, dass der Intensitätsgrad nicht für beide Gesichtshälften getrennt erfasst wird. <sup>94</sup> Dieses Defizit wird beim FEAT-Verfahren behoben (s. Tab. 6). Ein weiterer Nachteil von FACS ist die große Datenmenge und der damit verbundene hohe Erhebungsaufwand. <sup>95</sup> Hier schafft das EMFACS-Verfahren Abhilfe (s. Tab. 6). In der Literatur wird zudem die ökologische Validität des FACS hinterfragt. Es wird kritisiert, dass die Gesichtsausdrücke stilisiert und übertrieben

sind und in alltäglichen Interaktionen nur in abgeschwächter Form auftreten.  $^{96}$ 

Da AU's auf der menschlichen funktionalen Anatomie basieren, kann das FACS selbst bei sehr unterschiedlichen Gesichtsformen angewendet werden. Trotz der universellen Gültigkeit der Gesichtsmimik von Basisemotionen nehmen jedoch kultur- und familienspezifische Darstellungsregeln, sog. display rules, Einfluss auf den Ausdruck von Gefühlen. Auch wenn der biologisch angelegte Gesichtsausdruck durch die aktivierte Emotion automatisch angeregt wird,

<sup>&</sup>lt;sup>93</sup>Vgl. Weinberg (1986): 41f.

<sup>&</sup>lt;sup>94</sup>Vgl. Kaiser und Wehrle (2000): 423-426.

<sup>&</sup>lt;sup>95</sup>Vgl. Weinberg (1986): 43.

<sup>&</sup>lt;sup>96</sup>Vgl. Keltner et al. (2014): 91; Kaiser und Wehrle (2000): 420.

<sup>&</sup>lt;sup>97</sup>Vgl. Kaiser und Wehrle (2000): 423.

Die nachfolgende Tab. 7 zeigt, dass AU's Ähnlichkeiten mit Emojis aufweisen, die einen Gesichtsausdruck darstellen. Zu hinterfragen ist daher, ob Emojis das emotionale Ausdrucksverhalten in der schriftlichen Kommunikation ersetzen.

### 2.3. Emojis als emotionales Ausdrucksverhalten in der schriftlichen Kommunikation

In der direkten, persönlichen Kommunikation übermitteln Sender und Empfänger automatisch nonverbale Signale wie räumliches Verhalten, Gestik, Mimik, Blickkontakt und Tonfall. 100 Diese nonverbalen Signale übernehmen neben der Informationsvermittlung auch soziale Funktionen. 101 Sie regulieren die Interaktion der Gesprächspartner und bringen Vertrautheit zum Ausdruck. 102 Sowohl die Theorie der sozialen Präsenz<sup>103</sup> als auch die Medienreichhaltigkeitstheorie<sup>104</sup> postulieren, dass kein Medium nonverbale Signale so gut übermitteln kann, wie es bei der persönlichen Kommunikation der Fall ist. 105 Die Cuelessness-Theorie 106 schließt sich diesem Verständnis an und führt auf, dass ein Mangel an sozialen Hinweisen zu einer psychologischen Distanz zwischen den Gesprächspartnern und damit zu einer weniger freundlichen, sachlicheren Kommunikation führt. 107 Damit kann die computergestützte Kommunikation, wozu auch Kurznachrichten des Microblogging-Dienstes Twitter zählen, die o.g. sozialen Funktionen nicht erfüllen. 108 Nach Auffassung der Theorien gilt die computervermittelte Kommunikation damit als ineffektiv. 109

Die bisherigen Ausführungen zu den Grundlagen der Emotionsforschung erlauben die Schlussfolgerung, dass Emojis diese sozialen Funktionen zumindest teilweise übernehmen können. Emojis unterstützen den Sender der Textnachricht, Emotionen und Stimmungen auszudrücken, wodurch der Empfänger den Bedeutungskontext der Aussage besser verstehen kann. Somit ergibt sich für die vorliegende Arbeit die Überlegung, Emoji-enthaltende Kurznachrich-

<sup>98</sup>Vgl. Stürmer und Schmidt (2014): 62; Ekman und Friesen (2003): 23f; Merten (2003): 126; Euler (2000): 47; Kaiser und Wehrle (2000): 420. ten mithilfe automatischer Emotions- und Stimmungsanalysen zu untersuchen und darauffolgend Rückschlüsse über die emotionale Ausdruckskraft eines Emojis zu ziehen.

### 3. Stimmungsanalyse zur Messung des verbalen Ausdrucksverhaltens im Social Web

Kap. 3 beschäftigt sich mit der textbasierten Analyse von Emotionen und Stimmungen, die grundsätzlich manuell aber zunehmend auch automatisch oder kombiniert möglich sind. 112 Mit dem schnellen Wachstum des Social Webs, das Wissenschaft und Praxis einen kostenlosen Zugang zu einer Vielzahl an heterogenen Datenmengen ermöglicht, wächst das Forschungsinteresse an automatischen Stimmungs- und Emotionsanalysen als Alternative zu aufwendigen manuellen Verfahren. 113 Die Stimmungsanalyse, im Englischen unter dem Begriff Sentiment Analysis oder auch Opinion Mining bekannt, bezeichnet die automatische Analyse verbaler Stimmungsäußerungen in Textbeiträgen und ist ein Forschungsfeld in den Bereichen des Text Minings sowie der natürlichen Sprachverarbeitung.<sup>114</sup> Sie verfolgt vorrangig das Ziel, Texte und Äußerungen hinsichtlich ihrer Polarität – negativ, positiv oder neutral – zu klassifizieren, um so ein Stimmungsbild für die angesprochenen, kommentierten sowie diskutierten Objekte, z. B. Produkte, Unternehmen oder auch Personen, festzulegen. 115 Hiervon abzugrenzen sind die bisher weniger erforschten Emotionsanalyseverfahren, im Englischen als Emotion Mining bekannt, die eine Zuordnung von Texten nach konkreten Emotionen anstreben. 116 Liu führt an, dass solche Klassifizierungen deutlich anspruchsvoller als Stimmungsanalysen sind, da eine eindeutige Differenzierung aufgrund der höheren Anzahl an Emotionsklassen sowie der bestehenden Parallelen zwischen den Emotionen erschwert wird. 117 Die Ergebnisse solcher Klassifizierungen werden derzeit i. d. R. als nicht zufriedenstellend gewertet. 118

Die vorliegende Arbeit stellt deshalb die automatischen Stimmungsanalyseverfahren als bewährtere Verfahren im Rahmen des jungen Forschungsgebiets in den Fokus. Damit verbunden ist die Konsequenz, dass die zu entwickelnde Emoji-Kategorisierung nicht auf Basis konkreter Emotionen erfolgt. Als Ergebnis der stimmungsanalytischen Betrachtung wird nun eine Kategorisierung der Emojis nach negativen, positiven und neutralen Stimmungen angestrebt. In Kap. 3.1 werden die Grundlagen der automatischen Stimmungsanalyse skizziert und heutige Anwendungsbereiche dargestellt. In

<sup>99</sup>Vgl. Ekman und Friesen (2003): 23f.

<sup>100</sup> Vgl. Kiesler et al. (1984): 1125; Plutchik (1980): 96.

<sup>&</sup>lt;sup>101</sup>Vgl. Derks et al. (2007): 843; Ekman und Friesen (1969): 56.

 $<sup>^{102}</sup>$ Vgl. Derks et al. (2007): 843; Kaiser und Wehrle (2000): 420; Ekman und Friesen (1969): 56.

<sup>&</sup>lt;sup>103</sup>Vgl. Short et al. (1976).

<sup>&</sup>lt;sup>104</sup>Vgl. Draft und Lengel (1986).

<sup>&</sup>lt;sup>105</sup>Vgl. Byrne et al. (2012): 56.

<sup>&</sup>lt;sup>106</sup>Vgl. Rutter (1987).

<sup>&</sup>lt;sup>107</sup>Vgl. Byrne et al. (2012): 56; Rice und Love (1987): 88.

<sup>&</sup>lt;sup>108</sup>Vgl. Walther und D'Addario (2001): 324.

<sup>&</sup>lt;sup>109</sup>Vgl. Byrne et al. (2012): 56.

 $<sup>^{110}</sup>$ Vgl. Walther und D'Addario (2001): 325; Rezabek und Cochenour (1998): 201f; Thompsen und Foulger (1996): 230.

<sup>&</sup>lt;sup>111</sup>Vgl. Novak et al. (2015): 1; Huang et al. (2008): 466; Thompsen und Foulger (1996): 226, 229f.

<sup>112</sup> Vgl. Web Analytics Tools (2014).

<sup>&</sup>lt;sup>113</sup>Vgl. Liu (2015): xi; Ghiassi et al. (2013): 6267; Pang und Lee (2008): 7

<sup>7. &</sup>lt;sup>114</sup>Vgl. Liu (2015): 1; Zhang et al. (2014): 262; Kontopoulos et al. (2013): 4065; Guenther und Furrer (2013): 328.

<sup>&</sup>lt;sup>115</sup>Vgl. Liu (2015): 1; Kontopoulos et al. (2013): 4065.

<sup>&</sup>lt;sup>116</sup>Vgl. Aman und Szpakowicz (2007): 196.

<sup>&</sup>lt;sup>117</sup>Vgl. Liu (2015): 67.

<sup>&</sup>lt;sup>118</sup>Vgl. ebd.: 87.

Tabelle 7: Vergleich einiger AU's mit der Mimik der Emojis (Quelle: in Anlehung an o. V. (2016a))

Beschreibung der AU	Gesichtspartie	Beispiel	Emoji
Zwinkern	Augenpartie	-	6
		00	60
Heben des oberen Augenlides	Augenpartie		00
Kussmund	Mundpartie	100	3

Kap. 3.2 wird anschließend die Eignung des Microblogging-Dienstes Twitter als Datengrundlage für stimmungsanalytische Betrachtungen von Kurznachrichten begründet. Zum Abschluss wird in Kap. 3.3 ein Überblick zu dem aktuellen Forschungsstand hinsichtlich der emotionalen Aussagekraft von Emojis in der computergestützen Kommunikation gegeben.

### 3.1. Stimmungsanalytische Verfahren und ihre Anwendungsbereiche

Liu unterscheidet drei Komplexitätsebenen, auf denen Stimmungsanalysen durchgeführt werden können: die dokumentbasierte, die satzbasierte und die objekt- bzw. attributbasierte Stimmungsanalyse. <sup>119</sup> Diese werden im Folgenden kurz erläutert.

Im Rahmen der dokumentbasierten Stimmungsanalyse wird untersucht, ob eine Stellungnahme oder auch Meinungsäußerung im Ganzen betrachtet, wie z. B. eine Produktrezension, eine positive oder negative Stimmung gegenüber einem Objekt zum Ausdruck bringt. Hierbei wird die Annahme getroffen, dass die im Dokument implizierte Stimmung auf ein einziges Objekt ausgerichtet ist. Für Stellungnahmen bzw. Meinungsäußerungen, in denen mehrere Objekte evaluiert oder verglichen werden, ist diese Art der stimmungsanalytischen Betrachtung somit nicht ausreichend. 120

Die satzbasierte Stimmungsanalyse verfolgt das Ziel, die in einem Satz implizierte Stimmung zu untersuchen und eine entsprechende Klassifizierung vorzunehmen. Im Gegensatz zu der dokumentbasierten Stimmungsanalyse wird bei der satzbasierten Stimmungsanalyse neben den Klassen Positiv und Negativ meistens auch eine neutrale Klasse berücksichtigt. Es gilt die Annahme, dass der zu analysierende Satz entweder eine polare oder eine neutrale Stimmung ausdrückt. Sätze, die der neutralen Klasse zugeordnet werden, implizieren i. d. R. keine Stimmung und können als objektiv betrachtet werden. Für komplexe Sätze, in denen mehrere unterschiedliche Stimmungen zum Ausdruck kommen, ist dieser

Ansatz nicht geeignet. Des Weiteren können die in einem Satz enthaltenen Stimmungen nicht objektspezifisch zugeordnet werden.  $^{121}\,$ 

Im Gegensatz zu den ersten beiden Komplexitätsebenen versucht die objekt- bzw. attributbasierte Stimmungsanalyse Objekte und deren Attribute in Stellungnahmen oder Meinungsäußerungen zu identifizieren und darauf aufbauend zu bestimmen, ob die mit diesem Objekt bzw. Attribut verbundene Stimmung als negativ, positiv oder neutral einzuordnen ist. Sie ermöglicht somit eine stimmungsanalytische Betrachtung auf einer deutlich höheren Komplexitätsebene. Aufgrund der differenzierten Ergebnisse ist dieser Ansatz der Stimmungsanalyse für eine Vielzahl von Forschungsaufgaben besser geeignet und findet auch in der Praxis breite Anwendung. 122

Für die vorliegende Arbeit ist die Verwendung eines objekt- bzw. attributbasierten Stimmungsanalyseverfahrens nicht erforderlich. Wie bereits in Kap. 1 erwähnt, soll eine stimmungsanalytische Betrachtung von Tweets durchgeführt werden. Diese sind i. d. R. wenig komplex und umfassen nicht mehr als einen Satz. Eine Stimmungsanalyse auf Satzebene scheint daher ausreichend zu sein. Da zudem im Folgenden der Arbeit weitestgehend sichergestellt werden kann, dass sich die zu analysierenden Tweets auf ein bestimmtes, vorher definiertes Objekt beziehen, ist davon auszugehen, dass die für einen Tweet identifizierte Stimmung die Stimmung gegenüber dem entsprechenden Objekt widerspiegelt. 123

Unabhängig von der Komplexitätsebene lassen sich die Verfahren der automatischen Stimmungsanalyse in Abhängigkeit von der Idee und Konzeption des zugrundeliegenden Algorithmus in die Ansätze des maschinellen Lernens sowie die lexikonbasierten Ansätze unterscheiden. 124

Bei den Ansätzen des maschinellen Lernens werden Textklassifikatoren trainiert, um zu erkennen, ob eine negative, positive oder neutrale Stimmung vorliegt. Hierfür erlernen sie klassenspezifische Merkmale mithilfe von gelabelten

<sup>&</sup>lt;sup>119</sup>Vgl. Liu (2015): 9.

<sup>&</sup>lt;sup>120</sup>Vgl. ebd..

<sup>&</sup>lt;sup>121</sup>Vgl. ebd.: 9, 70f.

<sup>&</sup>lt;sup>122</sup>Vgl. Liu (2015): 9f.

<sup>&</sup>lt;sup>123</sup>Vgl. ebd.: 71.

<sup>&</sup>lt;sup>124</sup>Vgl. Ortigosa et al. (2014): 529; Zhang et al. (2014): 262f.

<sup>&</sup>lt;sup>125</sup>Vgl. Zhang et al. (2014): 262.

Daten, bspw. mit einer großen Menge an Tweets, die bereits manuell den Klassen Negativ, Positiv und Neutral zugeordnet wurden. 126 Da bei diesen Ansätzen die Klassenzuordnungen der entsprechenden Trainingsdaten bekannt sind, wird auch von einem überwachten Lernen gesprochen.<sup>127</sup> Die erlernten Merkmale werden im Englischen als Features bezeichnet und können in lexikalische sowie syntaktische Features unterschieden werden. 128 Features, die bei der Klassifizierung zum Einsatz kommen, sind u. a. einzelne Wörter, Wortarten, spezifische Stimmungswörter (z. B. gut oder schlecht), Redewendungen oder auch Negierungen. 129 Bereits vor dem Trainieren des Klassifikators wird bestimmt, welche Arten von Features erlernt und somit für die spätere Stimmungsklassifizierung berücksichtigt werden sollen. 130 Die Auswahl der Features beeinflusst die Qualität der Klassifikation wesentlich. 131 Nach Abschluss der Trainingsphase findet die Testphase statt. Hierbei werden Testdaten, die ebenfalls gelabelt und nicht identisch mit den Trainingsdaten sind, mithilfe des trainierten Klassifikators in die entsprechenden Klassen eingeordnet. 132 Anschließend wird die Qualität der Klassifikation mithilfe von bestimmten Kennziffern evaluiert. 133 Hierzu zählen üblicherweise die Korrektklassifikationsrate sowie die Genauigkeit. 134 Bei Bedarf kann das Klassifikationsergebnis durch Anpassungen in der Datenbasis, bspw. durch das Entfernen von Stoppwörtern sowie durch die Hinzunahme weiterer Features, optimiert werden. <sup>135</sup> Zu den in der Stimmungsanalyse meist verwendeten Klassifikatoren zählen der Naive-Bayes-Klassifikator, die Stützvektormaschine sowie der Maximum-Entropie-Klassifikator. 136

Bei den lexikonbasierten Ansätzen wird für die Durchführung einer Stimmungsklassifizierung auf sog. Emotions- bzw. Stimmungslexika zurückgegriffen. Diese führen eine Vielzahl von Emotions- bzw. Stimmungswörtern auf und geben Hinweise auf deren Stimmungsorientierung sowie -intensität. 137 Bei der stimmungsanalytischen Betrachtung der zu untersuchenden Daten, z. B. eines Tweets, kann es zur Identifizierung mehrerer Emotions- bzw. Stimmungswörter mit verschiedenen Stimmungsorientierungen bzw. -intensitäten kommen. Um eine eindeutige Stimmungsklassifizierung vorzunehmen, erfordert der lexikonbasierte Ansatz für diese Erscheinung eine polarisierende Abwägung. So kann zur Ermittlung der Polarität z. B. die Summe der Stimmungsintensitäten der negativen Stimmungswörter mit der Summe der Stimmungsintensitäten der positiven Stimmungswörter verglichen werden. Überwiegt die positive (negative) Stim-

 $^{126}$ Vgl. Han et al. (2012): 328; Chaovalit und Zhou (2005): 3.

mungsintensität, findet eine Zuordnung in die Klasse Positiv (Negativ) statt. Ergibt sich ein ausgeglichenes Verhältnis, wird in die neutrale Klasse eingeordnet. 138

Vor- und Nachteile beider Ansätze

Beide Klassifikationsansätze weisen Vor- und Nachteile auf (s. Tab. 8).

Welche der beiden Methode zum Einsatz kommt, ist von den entsprechenden Datengrundlagen sowie den spezifischen Anwendungsbereichen abhängig. Grundsätzlich gilt jedoch, dass der Einsatz von lexikonbasierten Methoden vor allem als sinnvoll erscheint, wenn das Erheben von Trainingsdaten nicht möglich ist. 140 Falls allerdings eine große Menge an Trainingsdaten zur Verfügung steht, erzielen die Methoden des maschinellen Lernens generell deutlich bessere Klassifikationsergebnisse. 141 Da mit der wachsenden Bedeutung des Social Webs eine Vielzahl an heterogenen, domänenübergreifenden Daten zum Trainieren und Testen von Klassifikatoren bereitsteht, messen Forscher den Methoden des maschinellen Lernens größere Bedeutung für die Zukunft bei. 142 Im späteren Verlauf der vorliegenden Arbeit kommt sowohl ein Verfahren, das auf dem Ansatz des maschinellen Lernens beruht, als auch ein lexikonbasiertes Verfahren zur Anwendung. Im Rahmen eines Pretests (s. Kap. 4.2) erfolgt die Entscheidung, welches Verfahren sich für die Stimmungsklassifizierung der zugrundeliegenden Datenbasis besser eignet.

Anwendungsbereiche der Stimmungsanalysen

Stimmungsanalysen finden bereits vielfältige Anwendung in verschiedenen Bereichen der Computer-, Betriebssowie Sozialwissenschaften. 143 So werden sie in der Politik u. a. dazu genutzt, um Stimmungen und Meinungen der Bevölkerung zu innen- und außenpolitischen Themen zu identifizieren und zu untersuchen. Mit ihrer Hilfe können Politiker auf Veränderungen des gesellschaftlichen Klimas reagieren. 144 In der Finanzbranche können Stimmungsanalysen genutzt werden, um bspw. Kursentwicklungen am Aktienmarkt zu prognostizieren. 145 Im Bereich des Marketings kommen stimmungsanalytische Betrachtungen insbesondere im Rahmen des Social Media Monitorings<sup>146</sup> zum Einsatz.<sup>147</sup> Die Kommunikation der Konsumenten bei Microblogging-Diensten wie Twitter kann die Wahrnehmung bzw. das Image einer Marke stark beeinflussen. 148 Das Markenimage 149 gilt als verhaltenswirksames Konstrukt, das Präferenzen schafft

<sup>&</sup>lt;sup>127</sup>Vgl. Han et al. (2012): 330.

<sup>&</sup>lt;sup>128</sup>Vgl. Becker et al. (2013): 333; Chaovalit und Zhou (2005): 3.

<sup>&</sup>lt;sup>129</sup>Vgl. Liu (2015): 49f; Pang und Lee (2008): 35-39.

<sup>&</sup>lt;sup>130</sup>Vgl. Chaovalit und Zhou (2005): 3.

<sup>&</sup>lt;sup>131</sup>Vgl. Sharma und Dey (2012): 1.

<sup>&</sup>lt;sup>132</sup>Vgl. Han et al. (2012): 330.

 $<sup>^{133}</sup>$ Vgl. Chaovalit und Zhou (2005): 3.

<sup>&</sup>lt;sup>134</sup>Vgl. Han et al. (2012): 364.

<sup>&</sup>lt;sup>135</sup>Vgl. Chaovalit und Zhou (2005): 3.

<sup>&</sup>lt;sup>136</sup>Vgl. Rosenthal et al. (2015): 461; Pang et al. (2002): 81.

<sup>&</sup>lt;sup>137</sup>Vgl. Liu (2015): **59**; Zhang et al. (2014): **263**; Mohammad et al. (2013): **322**.

<sup>&</sup>lt;sup>138</sup>Vgl. Liu (2015): 59.

<sup>&</sup>lt;sup>140</sup>Vgl. Ortigosa et al. (2014): 529.

<sup>&</sup>lt;sup>141</sup>Vgl. Liu (2015): 61; Gezici et al. (2013): 472.

<sup>&</sup>lt;sup>142</sup>Vgl. Rosenthal et al. (2015): 461; Liu (2015): 97.

<sup>&</sup>lt;sup>143</sup>Vgl. Liu (2015): xi.

<sup>&</sup>lt;sup>144</sup>Vgl. ebd.: 4f.

<sup>&</sup>lt;sup>145</sup>Vgl. Bollen et al. (2011): 91.

<sup>&</sup>lt;sup>146</sup>Im Social Media Monitoring werden die Dialoge in sozialen Netzwerken, die Erwähnungen der Marke eines Unternehmens und die von Social Media ausgehenden Verlinkungen überwacht." Heymann-Reder (2011): 256.

<sup>&</sup>lt;sup>147</sup>Vgl. Aßmann und Röbbeln (2013): **366**.

<sup>&</sup>lt;sup>148</sup>Vgl. Jansen et al. (2009b): 2172.

<sup>&</sup>lt;sup>149</sup>"Brand image is consumer perceptions of and preferences for a brand, measured by the various types of brand associations held in memory." Es besteht aus "[...] lower-level considerations, related to consumer perceptions

**Tabelle 8:** Vor- und Nachteile der Ansätze des maschinellen Lernens sowie der lexikonbasierten Ansätze (Quelle: Liu (2015): 61, 96f; Zhang et al. (2014): 262; Marchand et al. (2013): 421f; Ortega et al. (2013): 501; Chaovalit und Zhou (2005): 2)

Ansätze	Vor- (+) und Nachteile (-)
Ansätze des maschinellen Lernens	<ul> <li>+ Eine Vielzahl von Merkmalen kann erlernt und für die Klassifizierung berücksichtigt werden.</li> <li>+ Über die Quantität der Trainingsdaten lassen sich bessere Klassifikationsergebnisse erzielen.</li> <li>- Merkmale, die nur selten in den Trainingsdaten auftreten, können i. d. R. nicht verlässlich erlernt werden.</li> <li>Daher kann es zu Problemen bei der Stimmungsidentifizierung kommen.</li> <li>- Existierende Regeln und gelernte Merkmale können nur schwer angepasst bzw. verbessert werden.</li> <li>- Bei der Verwendung domänenspezifischer<sup>139</sup> Trainingsdaten kann es zum Erlernen von domänenspezifischen Merkmalen kommen. Ein domänenübergreifender Einsatz des Klassifikators ist somit nicht bedenkenlos möglich.</li> </ul>
Lexikonbasierte Ansätze	<ul> <li>+ Trainingsdaten und somit eine zeitaufwendige Labelung der Daten sind nicht notwendig.</li> <li>+ Existierende Regeln und Merkmale können unkompliziert und schnell durch den Nutzer manuell korrigiert bzw. ergänzt werden. Daher gelten diese Ansätze als flexibler.</li> <li>- Neben den aufgeführten Emotionswörtern können weitere Regeln und Merkmale kaum bzw. nur schwer berücksichtigt werden.</li> <li>- Der Auf- und Ausbau von lexikonbasiertem Wissen bedarf einiger Zeit.</li> <li>- Grundsätzlich liegt eine geringere Domänenabhängigkeit vor. Jedoch gibt es eine Vielzahl an domänenbzw. kontextabhängigen Emotions- und Stimmungswörtern, die eine bedenkenlose Übertragung des Klassifikators in eine andere Domäne nicht zulassen.</li> </ul>

und Kaufentscheidungen der Konsumenten beeinflusst. <sup>150</sup> Für Unternehmen ist es daher relevant, die in den Textbeiträgen geäußerten Stimmungen zu ihren Marken und Produkten mithilfe von Stimmungsanalysen zu erfassen, auszuwerten und im Zeitverlauf zu beobachten. So können negative und kritische Äußerungen der Konsumenten identifiziert und gegensteuernde Kommunikationsmaßnahmen eingeleitet werden, die den Aufbau eines positiven Images fördern. <sup>151</sup> Automatische Stimmungsanalysen stellen somit eine neue Methode für die Überwachung und Messung des Markenimages dar. Sie ermöglichen es zudem, die Kundenzufriedenheit zu untersuchen und wertvolle Informationen für den Aufbau sowie den Erhalt von Kundenbeziehungen zu erlangen. <sup>152</sup>

### 3.2. Kurznachrichten des Microblogging-Dienstes Twitter als geeignete Datenquelle für Stimmungsanalysen

Twitter ist ein im Jahr 2006 gegründeter, kostenloser Microblogging-Dienst. Sonsumenten nutzen die computergestützte Informations- und Kommunikationsplattform, um ihre persönlichen und spontan empfundenen Emotionen zu kommunizieren sowie Einschätzungen zu weltweiten Ereignissen in Echtzeit zu diskutieren. Nachrichten auf Twitter werden als Tweets bezeichnet. Der Begriff leitet sich

aus dem Englischen "to tweet", zu Deutsch "zwitschern", ab und steht für eine kompakte, auf 140 Zeichen beschränkte Kurznachricht. Schätzungsweise werden pro Minute 347.222 Tweets gesendet. Dies impliziert ein bedeutsames Datenvolumen, welches für umfangreiche Analysen genutzt werden kann. Da die Nutzer aufgrund des vorgegebenen Formats gezwungen sind, kurze und präzise Einschätzungen abzugeben und ihre Gefühle zu pointieren, sind Tweets in besonderem Maße für Stimmungsanalysen geeignet.

Zudem ist Twitter eine öffentlich zugängliche Datenquelle, die mithilfe der Twitter API<sup>157</sup> relativ einfach und kostenlos für Stimmungsanalysen verwendet werden kann. <sup>158</sup> Schätzungen zufolge sind rund 92 Prozent der Twitter-Accounts nicht durch Privatsphäre-Einstellungen geschützt und damit zugänglich. <sup>159</sup> Ein ReTweet ermöglicht die Weiterverbreitung eines, durch einen anderen Nutzer bereits abgesendeten und inhaltlich unveränderten Tweets im Twitter-Netzwerk. <sup>160</sup> Die @reply Fkt. ermöglicht es einem Nutzer, sich mit seinem Tweet direkt an einen anderen Nutzer zu wenden. Diese öffentlich sichtbaren Twitter-Aktivitäten gelten als Indikatoren für soziale Verbindungen. Hashtags ermöglichen hingegen Orientierung und ein schnelles Auffinden von Informationen zu spezifischen Themen oder Begrif-

of specific performance and imagery attributes and benefits, and higher-level considerations related to overall judgments, feelings, and relationships." Keller (2008): 379, 636.

<sup>&</sup>lt;sup>150</sup>Vgl. Schweiger (1995): 919.

<sup>&</sup>lt;sup>151</sup>Vgl. Heymann-Reder (2011): 100.

 $<sup>^{152}</sup>$ Vgl. Kiritchenko et al. (2014): 723; Mostafa (2013): 4241; Go et al. (2009): 1.

<sup>&</sup>lt;sup>153</sup>Vgl. Fischer (2015): **38**; Weinberg (2014): **173**; Barash und Golder (2011): **143f**; Jansen et al. (2009a): **3861**.

<sup>&</sup>lt;sup>154</sup>Vgl. Weinberg (2014): 173; Hassan et al. (2013): 357; Pak und Paroubek (2010): 1320; Jansen et al. (2009a): 3860.

<sup>155</sup> Vgl. Twitter Leitfaden (2015b); Ghiassi et al. (2013): 6268; Barash und Golder (2011): 144; Go et al. (2009): 2.

<sup>&</sup>lt;sup>156</sup>Vgl. o. V. (2016b): 11.

<sup>&</sup>lt;sup>157</sup>Eine API (Application-Programming-Interface) ist eine Schnittstelle, die von einem Softwaresystem bereitgestellt wird. Mithilfe dieser Schnittstelle ist es anderen Programmen möglich, sich an das Softwaresystem anzubinden. APIs werden vermehrt von Webdiensten verwendet und ermöglichen es, Daten und Inhalte zu extrahieren, auszutauschen und weiterzuverbreiten. Vgl. Gründerszene (2016).

 $<sup>^{158} \</sup>mathrm{Vgl.}$  Culotta und Cutler (2016): 343, 345; Barash und Golder (2011): 164.

<sup>&</sup>lt;sup>159</sup>Vgl. Cha et al. (2010): 12.

<sup>&</sup>lt;sup>160</sup>Vgl. Twitter Leitfaden (2015a); Barash und Golder (2011): 148f.

fen. Hashtags sind Schlüssel- oder Schlagwörter, die durch das Hash-Zeichen (#) eingeleitet werden und als eine Art Verschlagwortung von Tweets fungieren. <sup>161</sup>. Follower sind Nutzer, die Tweets anderer Twitter-Accounts abonnieren. Die Anzahl der Follower einer Marke deutet auf den sozialen Status und die Beliebtheit einer Marke hin und gilt daher auch als Indikator für das Markenimage. <sup>162</sup>

Ein weiteres Argument für die Nutzung von Twitter als Datenquelle ist die mediale Reichweite. 163 Twitter verzeichnet im Jahr 2016 weltweit rund 310 Millionen aktive Nutzer, Tendenz steigend, und belegt damit Rang neun der größten sozialen Netzwerke und Nachrichtendienste. 164 Allein knapp 20 Prozent der volljährigen US-Amerikaner waren im Jahr 2014 auf Twitter aktiv. 165 Die Twitter-Nutzer weltweit verteilen sich mit je rund 20 Prozent gleichmäßig auf die Altersklassen zwischen 18 und 54 Jahren bzw. älter als 54 Jahre. 166 Neben der hohen Altersdiversität sind auf Twitter auch eine Vielzahl von Nutzern unterschiedlicher sozialer Schichten und Interessensgruppen vertreten. 167 Die Heterogenität der Twitter-Nutzer ermöglicht es, die Ergebnisse der mit Tweets durchgeführten Stimmungsanalyse zu verallgemeinern.

Es gibt jedoch sprachliche Besonderheiten in der Twitter-Kommunikation, die als Herausforderung für eine automatische Stimmungsanalyse zu betrachten sind. Der sprachliche Ausdruck in den Tweets ist informell und entspricht meist nicht den lexikalischen und syntaktischen Normen. 168 Rechtschreib- und Grammatikfehler, Umgangssprache, Neologismen, Abkürzungen, Akronyme und Hashtags sind keine Seltenheit. 169 Dies begründet sich zum einen durch die Zeichenbeschränkung. Die Kenntlichmachung von ReTweets verringert die verbleibende Zeichenanzahl um ein weiteres und hat wiederum fehlende Artikel und neue Abkürzungen zur Folge. 170 Zum anderen sind Tweets oft spontaner Natur und werden von verschiedenen Geräten abgesetzt. 171 Bspw. sind 82 Prozent der Nutzer per Smartphone auf Twitter aktiv. 172 Die sprachlichen Besonderheiten können die Zuordnung von Tweets in die Klassen Negativ, Positiv und Neutral erschweren und zu Verzerrungen der Ergebnisse führen. Daher gilt es, sie bei der Entwicklung und Anwendung von Stimmungsanalyseverfahren zu berücksichtigen.

3.3. Forschungsstand zur emotionalen Aussagekraft von Emojis in Tweets

Wie bereits in den vorherigen Kapiteln angeführt, werden Emojis in der computergestützten Konsumentenkommunikation immer häufiger genutzt. Allein auf Twitter wurden seit 2014 mehr als 110 Milliarden Emojis verwendet. Die wissenschaftliche Erforschung der Verwendung von Emojis im Social Web ist allerdings noch recht jung. Erstmalig analysierten Novak et al. Stimmungen von Tweets und leiteten daraus Stimmungswerte für die 751 am häufigsten auftretenden Emojis ab. Als Datengrundlage dienten 1,6 Millionen Tweets in 13 verschiedenen europäischen Sprachen, die von professionellen Bewertern manuell in die Stimmungskategorien Negativ, Positiv und Neutral eingeordnet wurden. Die Autoren erzielten folgende Kernergebnisse:

- Nutzungshäufigkeit: Nur vier Prozent der Tweets enthalten Emojis.
- Auftreten der Emojis: Das mit Abstand am häufigsten verwendete Emoji ist das Face with tears of joy <sup>(2)</sup>, gefolgt von dem Emojis Heavy Black Heart <sup>(3)</sup> und Black Heart Suit <sup>(3)</sup>.
- Verständlichkeit: Emojis erleichtern das Textverständnis, sodass die Zuordnung von Tweets mit Emojis in eine Stimmungskategorie eine höhere Beurteiler-Übereinstimmung erzielt als Tweets ohne Emojis.
- Textstruktur: Emojis treten vorwiegend nach zwei Dritteln der Satzlänge auf.
- Stimmungswerte: Die meisten Emojis weisen einen positiven Stimmungswert auf, insbesondere die am häufigsten verwendeten Emojis. Die Polarität des Stimmungswertes eines Emojis steigt mit zunehmender Distanz zum Satzanfang.
- Interkulturalität der Ergebnisse: In Bezug auf unterschiedliche Sprachräume existieren keine signifikanten Unterschiede in den einzelnen Stimmungswerten der Emojis.<sup>175</sup>

Analog zum Vorgehen von Novak et al. schließen Vidal et al. in ihrer Arbeit von der manuell zugeordneten Stimmungskategorie eines Tweets auf den Stimmungswert des in dem Tweet vorkommenden Emojis. Jedoch beziehen sich die Tweets im Gegensatz zu der Datenbasis von Novak et al. inhaltlich ausschließlich auf den Lebensmittelbereich. So werden für die Analyse 12.260 Tweets verwendet, die eines der englischen Schlagwörter "Frühstück", "Mittagessen", "Snack" oder "Abendessen" enthalten und ungefähr gleichverteilt auf die vier Esssituationen sind. Die Autoren beziehen in ihre Analyse auch Emoticons ein.

<sup>&</sup>lt;sup>161</sup>Vgl. Heymann-Reder (2011): 130

<sup>&</sup>lt;sup>162</sup>Vgl. Naylor et al. (2012): 105; Barash und Golder (2011): 146, 149.

<sup>&</sup>lt;sup>163</sup>Vgl. Heymann-Reder (2011): 127.

<sup>&</sup>lt;sup>164</sup>Vgl. Statista (2016): 7, 18.

<sup>&</sup>lt;sup>165</sup>Vgl. Duggan et al. (2015).

<sup>&</sup>lt;sup>166</sup>Vgl. Statista (2016): 26.

<sup>&</sup>lt;sup>167</sup>Vgl. Hassan et al. (2013): 357f; Pak und Paroubek (2010): 1320.

<sup>&</sup>lt;sup>168</sup>Vgl. Hassan et al. (2013): 357; Boiy und Moens (2009): 529.

<sup>&</sup>lt;sup>169</sup>Vgl. Mostafa (2013): 4241; Kiritchenko et al. (2014): 723; Ghiassi et al. (2013): 6268; Go et al. (2009): 2; Boiy und Moens (2009): 529.

<sup>&</sup>lt;sup>170</sup>Vgl. Barash und Golder (2011): **149**.

<sup>&</sup>lt;sup>171</sup>Vgl. Go et al. (2009): 1f.

<sup>&</sup>lt;sup>172</sup>Vgl. Twitter (2016).

<sup>&</sup>lt;sup>173</sup>Vgl. Novak et al. (2015): 1.

<sup>&</sup>lt;sup>174</sup>Vgl. Shah (2016).

<sup>&</sup>lt;sup>175</sup>Vgl. Novak et al. (2015): 1-13.

- Nutzungshäufigkeit: 24 Prozent der Tweets enthalten Emojis/Emoticons.
- Auftreten der Emojis: Von den 631 zu untersuchenden Emojis wurden rund 40 Prozent (254 Emojis) in den Tweets identifiziert. Die am häufigsten verwendeten Emojis sind das Face savoring delicious food , das Smiling face with heart-shaped eyes und das Face with tears of joy . Bzgl. der Nutzungshäufigkeit ist jedoch keine Dominanz bestimmter Emojis zu erkennen.
- Textstruktur: Mehrfachnennungen sind selten. In 76
  Prozent der Fälle tritt nur ein Emoji/Emoticon pro
  Tweet auf.
- Stimmungswerte: 67 Prozent der Emojis bzw. Emoticons weisen einen positiven Stimmungswert auf.
- Zweck der Emojis/Emoticons: 92 Prozent der Emojis/Emoticons übermitteln zusätzliche Informationen, die nicht sprachlich im Tweet ausgedrückt werden.
- Acht Prozent der Emojis/Emoticons haben den Zweck, Informationen zu betonen, die im Text enthalten sind.<sup>176</sup>

Auch Twitter führte in den Jahren 2014 und 2015 eine Studie bzgl. der Nutzung von Emojis durch. Untersucht wurde deren Verwendung in Tweets zu ausgestrahlten TV-Sendungen.

- Nutzungshäufigkeit: Rund 14 Prozent der Tweets enthalten mindestens ein Emoji. Damit ist ein leichter Anstieg um vier Prozentpunkte seit der Einführung von Emojis auf Twitter zu verzeichnen. In Bezug auf verschiedene TV-Genres zeigen sich Unterschiede in der Nutzungshäufigkeit von Emojis.
- Auftreten der Emojis: Das mit Abstand am häufigsten verwendete Emoji ist das Face with tears of joy (24 Prozent), gefolgt vom Smiling face with heart-shaped eyes (zwölf Prozent) und dem Loudly crying face (acht Prozent).
- Emoji-Nutzer: 86 Prozent der Emoji-Nutzer sind 24 Jahre oder jünger. Knapp mehr als die Hälfte aller Emoji-Nutzer sind weiblich.<sup>177</sup>

Die vorliegende Arbeit knüpft an die hier vorgestellten Studien an. Anders als bisher soll die Emotionalität der Emojis jedoch mithilfe einer automatischen Stimmungsanalyse untersucht werden. Der Fokus liegt somit auf Auswahl und Konfiguration eines Verfahrens zur Ermittlung von Stimmungswerten bzw. der Entwicklung einer Emoji-Stimmungskategorisierung. Allgemeine Erkenntnisse zur Verwendung von Emojis in der Konsumentenkommunikation, die sich auf die Textstruktur, die Interkulturalität der Emojis sowie ihren Verwendungszweck beziehen, sind für die vorliegende Arbeit von geringerer Bedeutung.

### 4. Entwicklung einer Stimmungskategorisierung für Emojis

Im Zentrum des folgenden Kap. steht die Entwicklung einer Stimmungskategorisierung. Hierfür werden zunächst die Datengrundlage sowie das methodische Vorgehen im Rahmen der durchzuführenden Stimmungs- und Datenanalyse beschrieben. Im Anschluss sollen die Ergebnisse vorgestellt und validiert werden. Für die Erhebung der Datengrundlage sowie für ausgewählte statistische Berechnungen und Auswertungen kommen die frei zugängliche Software R und die hierzugehörige open-source Entwicklungsumgebung RStudio zur Anwendung.<sup>178</sup> Mithilfe der Softwaremodule twitteR, tm, Sentiment140 und Syuzhet wird die Software für die analysespezifischen Bedürfnisse konfiguriert.

#### 4.1. Bestimmung einer Datenbasis

Wie bereits in Kap. 3.2 aufgezeigt, nutzen Sozialforscher für die Identifikation sowie für die Analyse von affektiven Konsumenteneinstellungen zunehmend den Microblogging-Dienst Twitter. 179 Eines der populärsten Themen bei Twitter ist der Konsum von Lebensmitteln. Wöchentlich werden weltweit mehrere Millionen Tweets zu diesem Themenbereich abgesetzt. Nutzer berichten darüber, was, wo und mit wem sie essen, sowie welche Lebensmittelmarken bzw. -händler sie bevorzugen. 180 Die Literatur geht von einer reziproken Beziehung zwischen Emotionen und dem Konsum von Lebensmitteln aus. Demnach beeinflusst das emotionale Befinden der Konsumenten ihren Lebensmittelkonsum und der Konsum von Lebensmitteln wiederum das emotionale Befinden der Konsumenten.  $^{181}$  Dies legt die Vermutung nahe, dass insbesondere Tweets zu Marken bzw. Unternehmen aus diesem Themenbereich einen stark emotionalen Charakter aufweisen und sich daher für die Durchführung einer Stimmungsanalyse besonders gut eignen. In Tab. 9 werden, gemessen an der Anzahl der Follower, die zehn erfolgreichsten Twitter-Profile in der Kategorie Lebensmittelhandel dargestellt. 182 Das Unternehmen Starbucks führt mit seinem Twitter-Profil Starbucks Coffee das Ranking an und erst mit deutlichem Abstand folgen Marken wie McDonald's, Subway oder auch Burger King.

Die Starbucks Corporation, ansässig in Seattle (USA), ist ein international agierender Systemgastronom, der sich auf den Vertrieb von Kaffeeprodukten in konzerneigenen und lizenzierten Kaffeehäusern spezialisiert hat. Mit einem Markenwert von 29,31 Milliarden US-Dollar ist Starbucks nach McDonald's die wertvollste Fast-Food-Marke weltweit. 183 Mittlerweile betreibt das Unternehmen über 23.000 Filialen, wobei sich mehr als die Hälfte davon in den USA befindet. Mit einem jährlichen Nettoumsatz von rund 19 Milliarden

<sup>&</sup>lt;sup>176</sup>Vgl. Vidal et al. (2016): 119-127.

<sup>&</sup>lt;sup>177</sup>Vgl. Cruse (2015).

<sup>&</sup>lt;sup>178</sup>Vgl. RStudio (2016); R Foundation (2016).

<sup>&</sup>lt;sup>179</sup>Vgl. Ghiassi et al. (2013): 6266.

<sup>&</sup>lt;sup>180</sup>Vgl. Vidal et al. (2015): 1; MacMillan (2014).

<sup>&</sup>lt;sup>181</sup>Vgl. Macht (2008): 1; Canetti et al. (2002): 157.

<sup>&</sup>lt;sup>182</sup>Vgl. Socialbakers Twitter Statistics (2016b).

<sup>&</sup>lt;sup>183</sup>Vgl. Schept (2015): 123.

**Tabelle 9:** Top 10 der Twitter-Profile in der Kategorie Lebensmittelhandel gemessen an der Anzahl der Twitter-Follower (Quelle: Socialbakers Twitter Statistics (2016b))

Rang	Twitter-Profile	Anzahl der Follower
1.	Starbucks Coffee (@Starbucks)	11.740.031
2.	McDonald's (@McDonalds)	3.279.112
3.	スターバックス コーヒー (@Starbucks_J)	3.168.463
4.	Subway (@SUBWAY)	2.397.256
5.	Starbucks Indonesia (@SbuxIndonesia)	1.843.329
6.	Taco Bell (@tacobell)	1.761.468
7.	Nando's (@NandosUK)	1.555.472
8.	Pizza Hut (@pizzahut)	1.461.648
9.	Burger King (@BurgerKing)	1.417.247
10.	マクドナルド (@McDonaldsJapan)	1.119.608

US-Dollar im Jahr 2015 ist Starbucks heute einer der größten Kaffeeröster und -anbieter der Welt.  $^{\rm 184}$ 

Auch im Social Web ist es Starbucks gelungen, eine starke Markenpräsenz mit einer hohen medialen Reichweite aufzubauen. Selbst in einem branchenübergreifenden Vergleich der erfolgreichsten Twitter-Profile belegt das Profil Starbucks Coffee hinter Chanel, Samsung Mobile und PlayStation den vierten Platz. 185 Twitter-Nutzer weltweit verfolgen die Nachrichten und Neuigkeiten über das Unternehmen und werden angeregt, selbst polarisierende Tweets zum Thema Starbucks zu publizieren. Die Vielzahl und Vielfalt dieser abgesetzten Tweets bietet somit eine breite sowie solide Datengrundlage und stellt eine hinreichende Aussagekraft für die zu entwickelnde Emoji-Stimmungskategorisierung sicher.

Da im Rahmen der vorliegenden Arbeit nicht sämtliche Emojis berücksichtigt werden können ist die Auswahl relevanter Emojis erforderlich. Hierfür wird auf die von Twitter unterstützte Version Twemoji  $1.0^{186}$  zurückgegriffen. Diese wurde im November 2014 veröffentlicht und stellt den Nutzern 874 verschiedene Emojis zur Verfügung. Laut des SwiftKey Emoji Reports aus dem Jahr 2015 enthalten die Emoji-Gruppen Happy faces, Sad faces, Hearts und Hand gestures die am häufigsten verwendeten Emojis (s. Abb. 3).

Die Emoji-Gruppen Happy faces und Sad faces bilden bereits 59 Prozent der meist verwendeten Emojis ab. Als eine Form des Gesichtsausdrucks werden, bis auf zwei Ausnahmen<sup>189</sup>, alle Happy faces und Sad faces der Version Twemoji 1.0 in die Auswahl aufgenommen. Auch drei Herz-Emojis,

die in den Top 10 der meist genutzten Emojis gelistet sind, werden mit einbezogen. Hierzu gehören die Emojis Two Hearts ♥, Black Heart Suit ♥ und Heavy Black Heart ♥.

Zusätzlich finden die Handgesten Thumbs up und Thumbs down sowie die Emojis See-no-evil monkey , Hear-no-evil monkey und Speak-no-evil monkey zum Ausdruck von Stimmungen Berücksichtigung. Insgesamt beinhaltet die Auswahl somit 66 verschiedene Emojis, die im Anhang A (S. 85) eingesehen werden können.

Extraktion von Tweets zur Marke Starbucks

Zur Extraktion und späteren Analyse der Tweets zur Marke Starbucks kommt das Modul twitteR zur Anwendung. Dieses ermöglicht den Aufbau einer Verbindung zwischen der Software R und der Twitter API. Über diese Verbindung können, nach Authentifizierung bei Twitter (Skript B.2, S. 90), Tweets des Vortages mithilfe der Fkt. searchTwitter gezielt gesucht und abgerufen werden. Abb. 4 erläutert die Vorgehensweise.

Mithilfe des Skripts B.3 (S. 91) werden Tweets extrahiert, die den Suchtext Starbucks<sup>192</sup> enthalten und in englischer Sprache<sup>193</sup> formuliert sind. Da Twitter eine gewisse Zugriffsbegrenzung für das Abrufen von Tweets vorsieht, kann es beim Übersteigen dieser Begrenzung zum Funktionsabbruch kommen. Um den Abfragezeitraum vollständig zu erfassen und sicherzustellen, dass alle Tweets eines Tages abgerufen und gespeichert sind, müssen daher mehrere Suchabfragen durchgeführt werden. Dies erleichtert eine hierfür programmierte While-Schleife. Diese ermöglicht, Tweets eines definierten Abfragedatums rückwärtsschreitend, also von 24:00 Uhr bis 00:00 Uhr, abzurufen. Beim Ausführen des Skripts wird zunächst geprüft, ob bereits eine Datei für das vorher definierte Abfragedatum besteht. Ist dies der Fall, wird die

<sup>&</sup>lt;sup>184</sup>Vgl. Statista (2015): 13, 16.

<sup>&</sup>lt;sup>185</sup>Vgl. Socialbakers Twitter Statistics (2016a).

<sup>&</sup>lt;sup>186</sup>Es sei angemerkt, dass Twitter seit Dezember 2015 auch die Version Twemoji 2.0 unterstützt. Diese enthält insgesamt 1.657 verschiedene Emojis. Allerdings werden die entsprechenden Neuerungen nicht von allen online zur Verfügung stehenden Add-on-Tastaturen unterstützt, die erforderlich sind, um Emojis zu twittern. Es wird davon ausgegangen, dass eine Vielzahl der Nutzer die neuen Emojis nicht verwendet. Daher werden die Emojis der Version Twemoji 2.0 in der vorliegenden Arbeit nicht berücksichtigt.

<sup>&</sup>lt;sup>187</sup>Vgl. Emojipedia (2016b).

<sup>&</sup>lt;sup>188</sup>Vgl. SwiftKey (2015): 2.

 $<sup>^{189}\</sup>rm{Es}$ entfallen lediglich die Emojis Slightly smiling face und Slightly frowning face, da diese Emojis auf der verwendeten Add-on-Tastatur nicht zur Verfügung stehen.

<sup>&</sup>lt;sup>190</sup>Vgl. Emojitracker (2016).

<sup>&</sup>lt;sup>191</sup>Vgl. Gentry (2015): 1.

<sup>192</sup>Der Suchtext Starbucks ermöglicht die Extraktion von Tweets, die das Wort Starbucks in Groß- oder Kleinschreibung enthalten sowie die Extraktion von Tweets die #Starbucks oder @Starbucks aufweisen.

 $<sup>^{193}\</sup>mathrm{Da}$  die später verwendeten Stimmungsanalyseverfahren vor allem für englischsprachige Texte entwickelt wurden, werden lediglich in Englisch formulierte Tweets berücksichtigt.



Abbildung 2: Logo der Starbucks Corp. (Quelle: Wikipedia (2016))

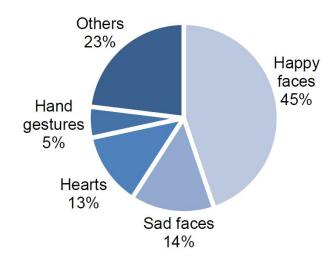


Abbildung 3: Die am häufigsten verwendeten Emoji-Gruppen (Quelle: in Anlehnung an SwiftKey (2015))

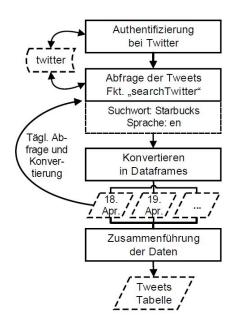


Abbildung 4: Extraktion der Starbucks-Tweets

ID des ältesten abgespeicherten Tweets ausgelesen. Von dieser ausgehend werden nun weitere 5.000 Tweets abgerufen und der bestehenden Datei hinzugefügt. 194 Dieser Vorgang wird so lange wiederholt, bis für das gewünschte Datum keine weiteren Tweets gefunden werden bzw. die Zugriffsbegrenzung von Twitter erreicht ist. 195 Für den Fall, dass noch keine Datei für das Abfragedatum erstellt wurde, wird der jüngste Tweet des Tages zum Suchwort Starbucks bei Twitter identifiziert. Hiervon ausgehend werden rückwärtsschreitend Tweets in 5.000er-Schritten so lange extrahiert, bis keine weiteren Tweets gefunden bzw. die Zugriffsbegrenzung erreicht ist. Abschließend werden die jeweiligen tagesspezifischen Datensätze, die im sog. Listenformat vorliegen, in Dataframes konvertiert (Skript B.4, S. 92) und in einem Datensatz zusammengefasst (Skript B.5, S. 92), um so in den folgenden Bearbeitungsschritten einen wiederholten Zugriff sowie eine schnelle Be- und Verarbeitung der Daten zu ermöglichen.

Im Zeitraum vom 18. April 2016 bis zum 18. Mai 2016 konnten mithilfe dieses Verfahrens 996.197 Tweets extrahiert werden. Neben den Textnachrichten stellt Twitter den Nutzernamen, ID, Datum, Uhrzeit und Geocode des Tweets sowie die Information, ob es sich um einen ReTweet handelt, zur Verfügung. Da diese zusätzlichen Informationen jedoch häufig nicht vollständig vorliegen, bleiben sie bei der späteren Auswertung weitgehend unberücksichtigt.

Selektion der Tweets mit Emojis

Für die vorliegende Arbeit werden ausschließlich die Tweets selektiert, die mindestens einen der zuvor ausgewählten Emojis enthalten. Die Software R bildet Emojis jedoch nicht als Piktogramme, sondern in Form von spezifischen Zeichenabfolgen ab, nachfolgend als Emoji-Code bezeichnet. Es bedarf somit einer Zuordnung der Codes zu den ausgewählten Emojis (Skript B.6, S. 93). Abb. 5 erläutert die Vorgehensweise. Zunächst werden die 66 ausgewählten Emojis gemeinsam mit ihrer offiziellen Benennung in einem hierfür angelegten Twitter-Account getwittert. Die Tweets dieses Accounts können nach der Authentifizierung bei Twitter mithilfe der Fkt. userTimeline abgerufen und der Emoji-Code, der sich an Textposition 1 und 2 befindet, sowie die Emoji-Benennung, die bei Textposition 3 beginnt, extrahiert werden. Da die Software R Emojis nicht als einzelne Symbole, sondern fälschlicherweise als zwei, zudem ungültige Zeichen codiert, ist eine Transformation der Textnachrichten in native Encodierung für die eindeutige Identifizierung sowie Zuordnung der Emoji-Codes erforderlich. Die unten aufgeführte Tab. 10 stellt einen exemplarischen Auszug des erstellten Dataframes dar. 196 Unter Anwendung dieses Dataframes lassen sich die zum Suchwort Starbucks extrahierten

Tweets gezielt nach den relevanten Emojis durchsuchen und diesen zuordnen.

### 4.2. Auswahl und Durchführung eines stimmungsanalytischen Verfahrens

Für die vorliegende Arbeit ist aus den kostenlosen, frei zugänglichen und mit der Software R kompatiblen automatischen Stimmungsanalysen ein geeignetes Verfahren auszuwählen. Die Eignung wird auf Basis der Übereinstimmungsquote zwischen manueller und automatischer Stimmungsanalyse ermittelt. Es wird deshalb ein Pretest anhand einer kleineren, zufällig gebildeten Stichprobe konzipiert, für die sowohl eine manuelle als auch eine automatische Stimmungsanalyse durchgeführt wird. Zur Bewertung der Übereinstimmung und Eignung der Verfahren werden Kennziffern definiert und ermittelt.

Durchführung eines Pretests zur Auswahl eines geeigneten stimmungsanalytischen Verfahrens

Für die Software R stehen im Rahmen der vorliegenden Arbeit zwei Stimmungsanalyseverfahren, Sentiment140 und Syuzhet, zur Auswahl zur Verfügung. 197 Sentiment 140 wurde an der Stanford Universität (USA) speziell für Tweets entwickelt und ist ein Verfahren, das auf dem Ansatz des maschinellen Lernens beruht. 198 Es ermöglicht eine Zuordnung von englischen sowie spanischen Tweets in die Klassen Negativ, Positiv und Neutral. 199 Hierfür wird ein, in der Sprachverarbeitung häufig eingesetzter, Maximum-Entropie-Klassifikator verwendet.<sup>200</sup> Dieser ermittelt klassenspezifische Gewichtungsfaktoren für das Vorkommen einzelner Features, wie Wörter (Unigram) und Zwei-Wort-Sequenzen (Bigram). Mit der Höhe des Gewichtungsfaktors steigt die Eignung eines Features als spezifischer Klassenindikator an. Die Wahrscheinlichkeit für eine Klassenzugehörigkeit wird gemäß folgender Formel berechnet:

$$P_{ME}(c|d,\lambda) = \frac{exp[\sum_{i} \lambda_{i} f_{i}(c,d)]}{\sum_{c'} exp[\sum_{i} \lambda_{i} f_{i}(c',d)]}$$
(1)

csteht für die entsprechende Klasse, d für den jeweiligen Tweet und  $\lambda$  gibt den ermittelten Gewichtungsfaktor an. Für das Trainieren und Testen des Maximum-Entropie-Klassifikators wurden ausschließlich positive und negative Tweets berücksichtigt, sodass die Zuordnung in die Klasse Neutral zu hinterfragen ist.  $^{201}$ 

Das Modul Syuzhet beruht auf einem lexikonbasierten Ansatz. Dem Nutzer stehen drei verschiedene Emotionslexika zur Auswahl, die nicht spezifisch für eine Klassifizierung von Twitter-Daten entwickelt wurden. Hiervon wird in der vorliegenden Arbeit das NRC-Emotionslexikon<sup>202</sup> verwendet. Es

 $<sup>^{194}\</sup>mathrm{Diese}$  Vorgehensweise führt dazu, dass einige Tweets doppelt abgespeichert werden. Um diese Doppelungen zu beheben, wird einer der entsprechenden Tweets im Nachhinein gelöscht.

<sup>&</sup>lt;sup>195</sup>Sollte es aufgrund der Zugriffsbegrenzung zu einem Abbruch kommen, wird der Zugriff durch die Twitter-API für 15 Minuten gesperrt. Danach kann das Skript erneut ausgeführt werden.

<sup>&</sup>lt;sup>196</sup>Die vollständige Tab. kann im Anhang A (S. 85) eingesehen werden. Die dort abgebildeten Piktogramme wurden händisch hinzugefügt.

 $<sup>^{197}\</sup>mathrm{Das}$  Modul Syuzhet ist zum Zeitpunkt der Arbeit als offizielles R-Softwaremodul gelistet. Das Modul Sentiment140 hingegen wird von dem Online-Dienst Github zur Verfügung gestellt.

<sup>&</sup>lt;sup>198</sup>Vgl. Go et al. (2013a).

<sup>&</sup>lt;sup>199</sup>Vgl. Go et al. (2013c).

<sup>&</sup>lt;sup>200</sup>Vgl. Go et al. (2013b); Nigam et al. (1999): 1.

<sup>&</sup>lt;sup>201</sup>Vgl. Go et al. (2009): 1-4.

<sup>&</sup>lt;sup>202</sup>Das NRC Emotionslexikon ist dem Zusatzmaterial (USB-Stick) beigefügt.

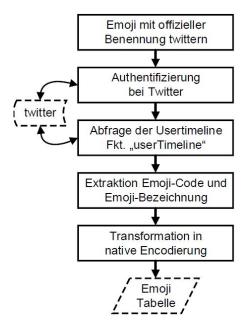


Abbildung 5: Extraktion des Emoji-Codes

Tabelle 10: Auszug aus der Emoji-Tab.

Emoji-Code	Emoji-Benennung
<ed><u+00a0><u+00bd><ed><u+00b8><u+0080></u+0080></u+00b8></ed></u+00bd></u+00a0></ed>	Grinning face
<ed><u+00a0><u+00bd><ed><u+00b8><u+00a2></u+00a2></u+00b8></ed></u+00bd></u+00a0></ed>	Crying face
<u+2764><u+fe0f></u+fe0f></u+2764>	Heavy black heart
<ed><u+00a0><u+00bd><ed><u+00b9><u+0088></u+0088></u+00b9></ed></u+00bd></u+00a0></ed>	See-no-evil monkey
<ed><u+00a0><u+00bd><ed><u+00b1><u+008d></u+008d></u+00b1></ed></u+00bd></u+00a0></ed>	Thumbs up sign

listet 14.182 englischsprachige Wörter auf und ordnet diese im Sinne des Emotion Minings den acht Emotionsklassen Angst, Erwartung, Ärger, Ekel, Freude, Traurigkeit, Überraschung und Vertrauen, aber auch den Polaritätsklassen Negativ und Positiv zu. 203 Enthält ein zu analysierender Tweet ein im Emotionslexikon gelistetes Wort, gibt das Verfahren der zu diesem Wort zugehörigen Emotions- bzw. Polaritätsklasse einen Punkt. Folglich ist der jeweilige Emotions- bzw. Polaritätswert eines Tweets von der Anzahl der gelisteten Wörter abhängig. 204 Tab. 11 verdeutlicht das Vorgehen. Der hier aufgeführte Tweet enthält die im NRC-Lexikon gelisteten Wörter "happy" und "friend". Beide Wörter werden den Emotionen Freude und Vertrauen sowie der Polaritätsklasse Positiv zugeordnet. Das Wort "happy" wird zudem auch der Emotion Erwartung zugeschrieben. So ergeben sich in der Summe der Zuordnungen die unten dargestellten Emotions- bzw. Polaritätswerte.

Für die Vergleichbarkeit der Ergebnisse beider stimmungsanalytischer Verfahren ist es erforderlich, auf Basis der Polaritätswerte analog zu Sentiment140 eine Zuordnung der

Tweets in die Klassen Negativ, Positiv und Neutral zu ermöglichen (Skript B.8, S. 94). Die vorliegende Arbeit ermittelt dafür einen Wert, der die Klassenzuordnung ermöglicht, und hier Gesamtpolarität  $G_{Polaritaet}$  genannt wird.

$$G_{Polaritaet} = Pos_{Polaritaet} - Neg_{Polaritaet}$$
 (2)

Bei der Berechnung der Gesamtpolarität  $G_{Polaritaet}$  wird der negative Polaritätswert  $Neq_{Polaritaet}$  von dem positiven Polaritätswert  $Pos_{Polaritaet}$  subtrahiert. Ist die Gesamtpolarität kleiner als null, wird der Tweet der Klasse Negativ, im Falle eines Wertes von null der Klasse Neutral und bei einem Wert größer als null der Klasse Positiv zugeschrieben. Die Ergebnisse zu der Emotionsklassifizierung bleiben entsprechend unberücksichtigt.

Die Ergebnisse beider Verfahren können nun mit den Stimmungszuordnungen einer manuellen Analyse im Rahmen des Pretests (s. Abb. 6) verglichen werden.

Die zufällig gebildete Stichprobe<sup>205</sup> besteht aus 100

<sup>&</sup>lt;sup>203</sup>Vgl. Jockers (2016): 1; Mohammad (2013b).

<sup>&</sup>lt;sup>204</sup>Vgl. Mohammad (2012): 588.

 $<sup>^{205}\</sup>mbox{Die}$  Stichprobe wurde aus der Grundgesamtheit von Tweets generiert, die im Zeitraum vom 18. bis zum 25. April abgerufen wurden. Die vollständige Stichprobe ist im Anhang C (S. 104) beigefügt. Zur Generierung einer zufälligen Stichprobe wird das Skript B.7 (S. 94) verwendet.

Tabelle 11: Ergebnis der Stimmungsanalyse Syuzhet an einem Beispiel-Tweet

Tweet	Angst	Erwartung	Ekel	Ärger	Freude	Traurigkeit	Überraschung	Vertrauen	Negativ	Positiv
It Is Starbucks date!! 😜 Happy Friday my friend!	0	1	0	0	2	0	0	2	0	2

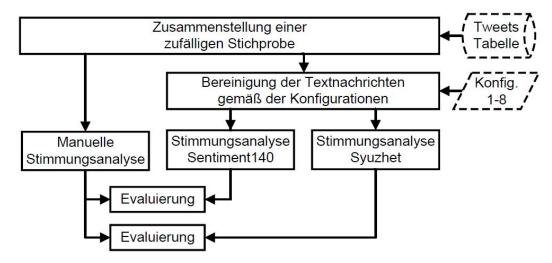


Abbildung 6: Durchführung des Pretests

Tweets der zuvor beschriebenen Datenbasis. Zur Qualitätssicherung analysieren zwei unabhängige Beurteiler diese Tweets manuell und ordnen sie den drei genannten Stimmungsklassen zu. Es ergibt sich eine Beurteiler-Übereinstimmung von 80 Prozent. Bei den übereinstimmenden Tweets werden 33 in die Klasse Negativ, drei in die Klasse Neutral und 44 in die Klasse Positiv eingeordnet. Für den Vergleich mit den Analyseverfahren Sentiment140 und Syuzhet werden lediglich diese 80 Tweets herangezogen. An dieser Stelle ist ergänzend zu erwähnen, dass die Beurteiler bei durchschnittlich rund 45 Prozent der Tweets eine Veränderung der Klassifizierung vornahmen, nachdem zusätzlich die im Tweet verwendeten Emojis in die Beurteilung einbezogen worden waren.

Aufgrund der in Kap. 3.2 beschriebenen syntaktischen und semantischen Besonderheiten von Tweets empfiehlt es sich für die Durchführung einer automatischen Stimmungsanalyse, die Kurznachrichten zunächst einer Textbereinigung zu unterziehen. Den Einfluss dieser Textbereinigung auf die Zuordnung der Tweets gilt es nun in einem weiteren qualitätssichernden Schritt zu analysieren und zu optimieren. Mithilfe des Text-Mining-Moduls tm werden verschiedene Textbereinigungs-Konfigurationen vor der Durchführung der

Stimmungsanalyse angewendet (Skript B.9, S. 95).<sup>207</sup>

Für die Evaluierung von Klassifikatoren werden, wie in Kap. 3.1 erwähnt, häufig die Korrektklassifikationsrate r und die Genauigkeit g als Kennzahlen herangezogen. Die Korrektklassifikationsrate r gibt das über alle Klassen gesehene Verhältnis zwischen der Anzahl der "richtig" klassifizierten Tweets zur Anzahl aller betrachteten Tweets n an und wird gemäß Formel (3) berechnet.  $^{209}$   $K_i$  und  $B_i$  sind die Menge der durch den Klassifikator bzw. den Beurteilern der Klasse i zugeordneten Tweets.

$$r = \frac{1}{n} \sum_{i} |K_i \cap B_i| \tag{3}$$

mit  $i \in \{neg, neutr, pos\}$ .

Im Gegensatz dazu misst die Kennziffer Genauigkeit g jeweils klassenspezifisch die Übereinstimmung zwischen Beurteiler und Klassifikator. Diese ergibt sich somit z. B. für die Klasse Neutral gemäß Formel (4).

$$g_{neutr} = \frac{|K_{neutr} \cap B_{neutr}|}{|K_{neutr}|} \tag{4}$$

 $<sup>^{206}</sup>$ Die Ergebnisse der manuellen Stimmungsanalyse können im Anhang C (S. 104) eingesehen werden.

 $<sup>^{207}</sup>$ Die verschiedenen Text-Bereinigungskonfigurationen können im Anhang C.1 (S. 110) nachgelesen werden.

 $<sup>^{208}\</sup>mathrm{Ergebnisse},$  die mit denen der manuellen Analyse übereinstimmen werden als "richtig" bezeichnet.

<sup>&</sup>lt;sup>209</sup>Vgl. Han et al. (2012): 366.

Sie ist das Verhältnis der Anzahl der Tweets, die, sowohl durch den Klassifikator als auch durch die Beurteiler, als neutral klassifiziert werden zur Anzahl aller durch den Klassifikator als neutral klassifizierten Tweets.<sup>210</sup>

Im Folgenden werden die entscheidenden Evaluierungsergebnisse vorgestellt und in den Tabellen 12 und 13 aufgeführt.  $^{211}$ 

Für beide Analyseverfahren ergeben sich über alle Konfigurationen der Textbereinigung niedrige Korrektklassifikationsraten. Bei der Methode nach Syuzhet ist eine maximale Korrektklassifikationsrate von 31,25 Prozent, bei Sentiment140 eine leicht bessere maximale Korrektklassifikationsrate von 38,75 Prozent zu beobachten. Auffällig ist zudem die große Anzahl an neutral klassifizierten Tweets. Die Genauigkeit für diese Klasse beträgt über alle Konfigurationen der Textbereinigung jeweils nur ca. vier Prozent. Somit liegt eine große Anzahl an Fehlklassifizierungen in der Klasse Neutral vor. Dies erklärt die generelle schwache Übereinstimmung zwischen manueller und automatischer Stimmungsanalyse.

Zudem fällt auf, dass bei dem Verfahren Sentiment140 die Anzahl der neutralen Tweets von 48 (60%) in der Konfig. 6 auf bis zu 68 (85%) in der Konfig. 8 ansteigt. Insbesondere durch das Entfernen der Interpunktion und Zahlen ab Konfig. 7 scheinen für die Klassen Positiv und Negativ entscheidende und präzisierende Textelemente zu fehlen. Dieses Ergebnis lässt den Schluss zu, dass die Klasse Neutral, anders als in Kap. 3 beschrieben, weniger Tweets mit einer neutralen bzw. objektiven Stimmung repräsentiert, sondern eher als Restklasse für nicht identifizierbare bzw. nicht eindeutig zuzuordnende Tweets zu verstehen ist. Da die Klasse Neutral somit ihre Aussagekraft verliert und für die Entwicklung einer Emoji-Stimmungskategorisierung eher klar polarisierende Tweets von Bedeutung sind, beziehen die weiteren Analysen und Auswertungen ausschließlich die positiv und negativ klassifizierten Tweets ein. Die nun gemäß Formel (5) ermittelte Korrektklassifikationsrate für die Klassen Negativ und Positiv  $r_{neg,pos}$  erzielt für beide Klassifikatoren deutlich

$$r_{neg,pos} = \frac{1}{n_{neg,pos}} (|K_{neg} \cap B_{neg}| + |K_{pos} \cap B_{pos}|)$$
 (5)

Sentiment140 weist in Konfig. 3 einen Spitzenwert von 85,29 Prozent auf und erzielt damit einen um knapp 20 Prozentpunkte besseren Wert als das Verfahren Syuzhet. Eine mögliche Erklärung hierfür ist, dass Sentiment140 die in Kap. 3.2 beschriebenen Besonderheiten der Twitter-Kommunikation besser bewerten kann. Es verwendet einen Klassifikator, der gezielt für Twitter-Daten trainiert wurde. Dies ist bei dem Verfahren Syuzhet und dem hier zum Einsatz kommenden NRC-Emotionslexikon nicht der Fall. 212 Mit Abschluss des

Pretests empfiehlt sich somit die Anwendung des Analyseverfahrens Sentiment140 für die im Folgenden durchzuführende Stimmungsanalyse. Ergänzend ist zu erwähnen, dass Syuzhet auch im Hinblick auf die mit diesem Verfahren mögliche Emotionsklassifizierung keine aussagekräftigen Ergebnisse erzielt. Dies verdeutlicht die im Anhang D (S. 111) aufgeführte Tab.. <sup>213</sup>

Durchführung der stimmungsanalytischen Betrachtung mit Sentiment140

Abb. 7 erläutert die Vorgehensweise der Stimmungsanalyse für die in Kap. 4.1 gebildete Datenbasis (Skript B.10/B.11/ B.12, S. 97-101). Da ReTweets die Ergebnisse verzerren würden, sind diese zunächst zu entfernen.<sup>214</sup> Hierfür steht das Modul twitteR mit der bereitgestellten Fkt. strip retweets zur Verfügung. Diese Fkt. ist in ihrer ursprünglichen Form nur für Daten im Listenformat kompatibel. Weil der zu untersuchende Datensatz jedoch im Format eines Dataframes vorliegt, müssen zuvor leichte Anpassungen im Code der Fkt. vorgenommen werden. Durch das Entfernen der ReTweets reduziert sich die Datenbasis von 996.197 Tweets um ca. 34 Prozent auf 657.502 Tweets. Analog zur Extraktion der Emoji-Codes werden auch hier die Textnachrichten in native Encodierung transformiert und so sichergestellt, dass die in den Tweets enthaltenden Emojis im späteren Verlauf identifiziert werden können. Da für die Stimmungsanalyse nur Tweets berücksichtigt werden sollen, die mindestens einen der in Kap. 4.1 ausgewählten Emojis vorweisen, werden die Tweets anhand der Emoji-Code-Liste (s. Anhang A, S. 85) nach den entsprechenden Codes durchsucht, diesen zugeordnet und abgespeichert. Tweets, die keinen der 66 Emoji-Codes enthalten, werden aus der Datengrundlage entfernt. Damit reduziert sich die Datenbasis auf 86.205 Tweets. Laut Vidal et al. besteht ein deutlicher Trend, nur ein Emoji pro Tweet einzusetzen (s. Kap. 3.3). Die vorliegende Arbeit verzichtet deshalb auf eine nach Mehrfachnennungen differenzierte Analyse. Ein möglicher Einfluss auf die zu entwickelnde Emoji-Stimmungskategorisierung wird als gering

Unter Berücksichtigung der Ergebnisse aus dem Pretest werden im Rahmen der Textbereinigung sämtliche Textnachrichten in Kleinbuchstaben transformiert und bestehende URLs gelöscht. Im Anschluss erfolgt die Durchführung der Stimmungsanalyse. Hierbei werden die den Emojis zugeordneten Tweets durch den Klassifikator in die Klassen Negativ, Positiv und Neutral eingeordnet. Die Klasse Neutral bleibt jedoch, wie oben beschrieben, unberücksichtigt. Tab. 14 zeigt einen Auszug der Ergebnisse der Stimmungsanalyse auf.<sup>215</sup>

<sup>&</sup>lt;sup>210</sup>Vgl. Han et al. (2012): 368.

 $<sup>^{211}\</sup>bar{\rm Die}$  vollständigen Evaluierungsergebnisse können dem Zusatzmaterial (USB-Stick) entnommen werden. Dateiname: Ergebnisvergleich manuelle und automatische Stimmungsanalyse.

 $<sup>^{212}\</sup>mbox{Es}$ sei angemerkt, dass bereits Twitter-spezifische Emotionslexika wie z.

B. das NRC Hashtag Emotion Lexicon entwickelt wurden. Vgl. Mohammad (2013a). Dieses konnte jedoch in die vorliegende Arbeit nicht einbezogen werden.

<sup>&</sup>lt;sup>213</sup>Da nicht ausgeschlossen werden konnte, dass die mit Syuzhet mögliche Emotionsklassifizierung einen Beitrag für die Emoji-Stimmungskategorisierung leisten kann, wurde parallel zur folgenden Vorgehensweise für Sentiment140 auch das Verfahren Syuzhet durchgeführt. Das hierzu erforderliche Skript B.13 (S. 102) kann im Anhang eingesehen werden.

 $<sup>^{214}</sup>$ ReTweets sind Duplikate und als solche aus der Datenbasis zu entfernen.  $^{215}$ Die vollständige Tab. kann im Anhang A (S. 85) eingesehen werden.

Tabelle 12: Ergebnisse und Kennziffern zur Stimmungsanalyse mit Sentiment140

Sentiment140 Konfig.								
	1	2	3	4	5	6	7	8
Anzahl neg	12	12	13	13	13	13	14	7
Anzahl neutr	47	47	46	46	48	48	61	68
Anzahl pos	21	21	21	21	19	19	5	5
g <sub>neutr</sub> in %	4,26	4,26	4,35	4,35	4,17	4,17	3,28	2,94
r in %	37,50	37,50	38,75	38,75	36,25	36,25	22,50	15,00
$r_{neg,pos}$ in %	84,85	84,85	85,29	85,29	84,38	84,38	84,21	83,33

Tabelle 13: Ergebnisse und Kennziffern zur Stimmungsanalyse mit Syuzhet

Syuzhet Konfig.								
	1	2	3	4	5	6	7	8
Anzahl neg	11	11	11	11	11	11	11	11
Anzahl neutr	44	44	44	44	46	46	45	45
Anzahl pos	25	25	25	25	23	23	24	24
g <sub>neutr</sub> in %	4,55	4,55	4,55	4,55	4,35	4,35	4,44	4,44
r in %	31,25	31,25	31,25	31,25	28,75	28,75	30,00	30,00
$r_{neg,pos}$ in %	63,89	63,89	63,89	63,89	61,76	61,76	62,86	62,86

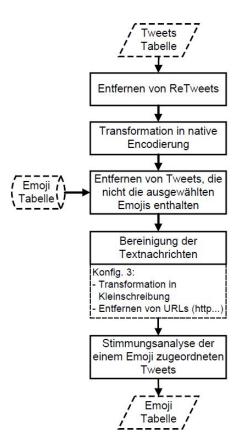


Abbildung 7: Vorgehensweise der Stimmungsanalyse

Dem Zusatzmaterial (USB-Stick) können zudem Emoji-spezifische Dateien, in denen sämtliche Tweets mit ihren ermittelten Stimmungen aufgeführt sind, entnommen werden. Das zur Erstellung dieser Dateien erforderliche Skript ist ebenfalls dem Zusatzmaterial zu entnehmen. Dateiname: 06a Tweets Text with Sentiment per Emoji for Starbucks-Tweets.

So tritt bspw. das Emoji *Grinning face*  $\stackrel{\mbox{\scriptsize $\omega$}}{=}$  in 572 Tweets auf, von denen 34 als negativ und 229 als positiv klassifiziert werden.

Aufbauend auf diesen Ergebnissen kann nun mithilfe der negativ und positiv klassifizierten Tweets ein Emojispezifischer Stimmungswert ermittelt und eine Emoji-Stimmungskategorisierung entwickelt werden. Hierfür wird gemäß Formel (6) die Differenz aus der Anzahl der positiven Tweets  $n_{pos}$  und der Anzahl der negativen Tweets  $n_{neg}$  zur Summe der negativen und positiven Tweets ins Verhältnis gesetzt.  $^{216}$ 

$$W_{Stimmung} = \frac{n_{pos} - n_{neq}}{n_{pos} + n_{neq}}$$

$$Wobei \ W_{Stimmung} \in [-1, 1]$$
(6)

Emojis mit einem Stimmungswert  $W_{Stimmung}$  größer null sind der Stimmungskategorie Positiv zugeordnet. Ein Wert kleiner null hingegen symbolisiert die Zugehörigkeit zu der Stimmungskategorie Negativ. Auf Basis der möglichen Stimmungswerte von minus eins bis plus eins ergibt sich zudem eine intensitätsmessende Emoji-Polaritätsskala. Die ermittelten Stimmungswerte sowie weitere Ergebnisse werden im folgenden Kap. ausführlich dargestellt und näher erläutert.

### 4.3. Vorstellung und der Bewertung der Ergebnisse

Die Ergebnisvorstellung gliedert sich in zwei Abschnitte. Der Fokus liegt auf den Ergebnissen zur Entwicklung einer Emoji-Stimmungskategorisierung. Begleitend werden jedoch zunächst Ergebnisse skizziert, die den allgemeinen Hintergrund der Verwendung von Emojis in der Konsumentenkommunikation beleuchten.

Zuerst wird die Nutzungsintensität der Emojis untersucht. Wie zuvor dargestellt, enthalten 86.205 Tweets der für die Auswertung zu berücksichtigten Datenbasis (657.502 Tweets) mindestens eines der 66 ausgewählten Emojis. Dies entspricht einem prozentualen Anteil von 8,65 Prozent. Die Nutzungsintensität von Emojis liegt damit im Rahmen der Ergebnisse der in Kap. 3.3 vorgestellten Studien, die Werte von vier bis 14 Prozent berichten. Angesichts der wachsenden Bedeutung von Emojis erscheint dieser Wert dennoch als niedrig. Die Social-Media-Plattform Instagram, bei der Nutzer Bilder hochladen und teilen können, berichtet hingegen, dass bis zu 50 Prozent der Beiträge Emojis aufweisen.<sup>217</sup> Twitter-Nutzer scheinen Emojis derzeit eher selten in ihre Kommunikation einzubinden. Dies könnte historisch bedingt sein, da Twitter lange Zeit als ein klassischer, auf Textnachrichten fokussierter Nachrichtendienst galt, der erst im Jahr 2014 die Möglichkeit der Verwendung von Emojis einführte.<sup>218</sup>

Aufgrund der möglichen Mehrfachnennung von Emojis pro Tweet weisen die 86.205 Textnachrichten insgesamt 101.341 Emojis auf, wobei identische Emojis bei Mehrfachnennung im selben Tweet nur einfach gezählt wurden. Die in Tab. 15 aufgeführten Ergebnisse bestätigen den von Vidal et al. identifizierten Trend, pro Tweet nur ein Emoji einzusetzen. Lediglich 14,87 Prozent der Tweets beinhalten Mehrfachnennungen. Hierbei dominiert der Einsatz von zwei Emojis mit 12,76 Prozent.

Die Analyse der Daten zeigt auf, dass nur wenige der 66 gelisteten Emojis häufig genutzt werden. In Abb. 8 werden die zehn meist verwendeten Emojis im Zusammenhang mit der Marke Starbucks dargestellt.

Im Vergleich mit den Top 10 der Emoji-Ranking-Liste der Webseite Emojitracker<sup>219</sup> ergeben sich deutliche Überschneidungen. So sind sechs der Emojis auch dort gelistet. Die Platzierungen eins und sechs sind sogar identisch.<sup>220</sup> Die Top 10 der bei Starbucks am häufigsten genutzten Emojis entsprechen in Bezug auf die Anzahl aller getwitterten Emojis einem Anteil von 63,73 Prozent, wobei allein rund 36 Prozent auf die Top 3 der Emojis entfallen. Demgegenüber werden 42 der 66 Emojis weniger als 1.000-mal, also unter einem Prozent, in der Datenbasis verzeichnet. Das Emoji Kissing face ist mit 26 Zählungen das am wenigsten genutzte Emoji. Anders als bei der Studie von Vidal et al. zeigt sich somit in der Verwendung eine deutliche Dominanz von einzelnen Emojis.

Von der Vielfalt des zur Verfügung stehenden Angebots an Emojis wird somit nur wenig Gebrauch gemacht. Eine mögliche Ursache hierfür ist, dass Konsumenten durch die große Auswahlmöglichkeit überfordert sind oder aus Zeitgründen gerne auf bewährte Muster bzw. Emojis zurückgreifen. Zudem ist es möglich, dass einige Emojis aufgrund fehlender prägnanter Merkmale kaum emotionale Aussagekraft besitzen. Ihre Verwendung wird vermieden, um Fehlinterpretationen bzw. Missverständnisse auszuschließen.

Analog zu den in Kap. 3.3 vorgestellten Studien wurde im Rahmen des Pretests (s. Kap. 4.2) festgestellt, dass Emojis das Textverständnis beeinflussen. Im Rahmen der manuellen Klassifizierung nahmen die Beurteiler Anpassungen hinsichtlich der Klassenzuordnung vor, nachdem sie die im Tweet verwendeten Emojis in ihre Beurteilung einbezogen hatten. Dies bestätigt die emotionale Ausdruckskraft der Emojis.

Im Folgenden werden nun die Ergebnisse mit Blick auf die zu entwickelnde Stimmungskategorisierung vorgestellt. Die Verteilung der 66 berücksichtigten Emojis auf die als positiv sowie negativ klassifizierten Tweets wird untersucht und die sich daraus ergebenden Stimmungswerte analysiert. <sup>221</sup> Zudem werden Auffälligkeiten hinsichtlich der ermittelten Stimmungswerte aufgezeigt, die für die Emoji-Stimmungskategorisierung von Bedeutung sind.

Insgesamt weisen 32 der Emojis einen negativen und 34

<sup>&</sup>lt;sup>216</sup>Berechnung in Anlehnung Novak et al. (2015): 16.

<sup>&</sup>lt;sup>217</sup>Vgl. Dimson (2015). Da es derzeit kein für die Software R kompatibles Modul gibt, mit dem Text nachrichten auf Instagram abgerufen werden können, kam diese Social-Web-Plattform im Rahmen der vorliegenden Arbeit nicht als Datenquelle in Frage.

<sup>&</sup>lt;sup>218</sup>Vgl. Cruse (2015).

<sup>&</sup>lt;sup>219</sup>Die Webseite Emojitracker verfolgt in Echtzeit die Verwendung sämtlicher Emojis bei Twitter und ermittelt die Nutzungshäufigkeit pro Emoji. Vgl. Emojitracker (2016).

<sup>&</sup>lt;sup>220</sup>Vgl. Emojitracker (2016).

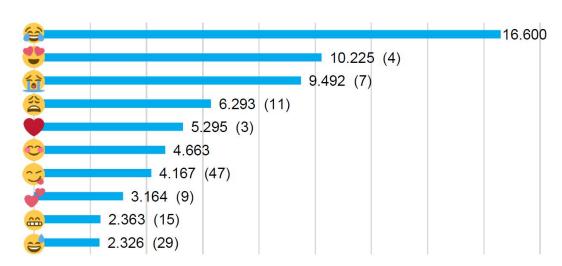
<sup>&</sup>lt;sup>221</sup>Die Stimmungswerte aller Emojis sind im Anhang A (S. 85) aufgelistet.

**Tabelle 14:** Auszug der Ergebnisse der Stimmungsanalyse. Die Häufigkeit umfasst neben der Anzahl der negativen und positiven Tweets auch die Anzahl der neutralen Tweets.

Emoji	Benennung	Häufigkeit	Negativ	Positiv
	Grinning face	572	34	229
	Crying face	1.198	295	103
	Heavy black heart	5.295	209	1.971
	See-no-evil monkey	982	105	227
	Thumbs up sign	1.706	65	661

Tabelle 15: Anzahl der Tweets, die Einfach- und Mehrfachnennung verschiedenartiger Emojis aufweisen

Emoji-Nennung	1	2	3	4	5	6	7	8	9	15	16
Anzahl Tweets	73.388	11.001	1.479	251	53	16	5	7	3	1	1



**Abbildung 8:** Top 10 der am häufigsten verwendeten Emojis (Starbucks). In Klammern sind die abweichenden Platzierungen zur Emoji-Ranking-Liste der Webseite Emojitracker aufgeführt.

der Emojis einen positiven Stimmungswert auf. Hervorzuheben ist, dass unter den zehn meist verwendeten Emojis überwiegend Emojis mit stark positiven Stimmungswerten vertreten sind (s. Tab. 16).

Die folgenden Abbildungen 9 und 10 bilden alle betrachteten Emojis mit negativen und positiven Stimmungswerten in Abhängigkeit von ihrer Nutzungshäufigkeit ab.

Über alle betrachteten Emojis weisen die Emojis Persevering face with look of triumph den niedrigsten Stimmungswert von -0,63 auf. Das Emoji Kissing face with smiling eyes hingegen erzielt mit 0,91 den höchsten Stimmungswert. Keines der betrachteten 66 Emojis erzielt einen Stimmungswert in Höhe von eins bzw. minus eins, d. h. kein Emoji wird ausschließlich in negativ bzw. positiv klassifizierten Tweets verwendet.

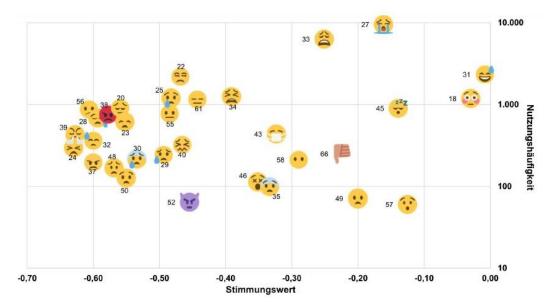
Der Mittelwert für die negativ bzw. positiv assoziierten Emojis liegt bei -0,41 bzw. 0,68. Emojis mit negativem Stimmungswert werden somit deutlich häufiger (zu rund 35 Prozent) in positiv klassifizierten Tweets eingesetzt, als dies umgekehrt bei den Emojis mit positiven Stimmungswerten der Fall ist (zu rund 17%). Ein möglicher Grund hierfür könnte sein, dass negative Emojis auch zum Ausdruck von Ironie verwendet werden. Da Ironie durch die Stimmungsanalyse Sentiment140 derzeit noch nicht erkannt werden kann, kommt es zu Fehlinterpretationen der Tweets. Diesen Aspekt verdeutlichen die in Tab. 17 dargestellten Textnachrichten.

Der Emoji-Report von SwiftKey ordnet Emojis mit einem spezifischen Gesichtsausdruck in die Gruppen Happy faces und Sad faces ein. Alle lächelnden, grinsenden, küssenden und zwinkernden Emojis zählen zur Gruppe Happy faces. Sämtliche Emojis, die diese Kriterien nicht erfüllen oder traurig bzw. verärgert wirken, werden hingegen in die Gruppe der Sad faces eingeordnet.<sup>222</sup> Für die vorliegende Arbeit wurde auf Basis dieser Beschreibung eine Zuordnung der aus-

<sup>&</sup>lt;sup>222</sup>Vgl. SwiftKey (2015): 6.

Tabelle 16: Stimmungwerte der am häufigsten verwendeten Emojis

Rang	Emoji	Bezeichnung	Stimmungswert	Häufigkeit in %
1		Face with tears of joy	0,33	16,38
2		Smiling face with heart-shaped eyes	0,81	10,09
3	Tot	Loudly crying face	-0,16	9,37
4	330	Weary face	-0,25	6,21
5		Heavy Black Heart	0,81	5,22
6	0	Smiling face with smiling eyes	0,78	4,6
7	~	Face savouring delicious food	0,81	4,11
8	•	Two Hearts	0,84	3,12
9	<u></u>	Grinning face with smiling eyes	0,71	2,33
10	<b>\(\text{\tin}\text{\tetx{\text{\tetx{\text{\text{\texi}\text{\text{\texi}\text{\text{\text{\text{\ti}\text{\text{\text{\text{\text{\texi}\texit{\text{\texit{\text{\texi}\titt{\text{\ti}\tintt{\text{\texi}\texit{\text{\tet</b>	Smiling face with open mouth and cold sweat	-0,01	2,3



**Abbildung 9:** Negative Stimmungskategorie (Starbucks): Emojis mit negativem Stimmungswert in Abhängigkeit von ihrer Nutzungshäufigkeit. Die Häufigkeit ist aus Gründen der Übersichtlichkeit logarithmisch skaliert.

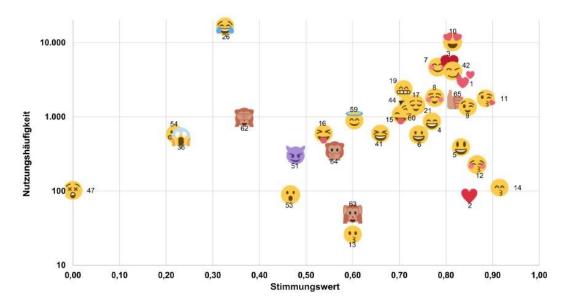
Tabelle 17: Darstellung von Ironie und Sarkasmus anhand von zwei Beispiel-Tweets

Nr.	Textnachricht	Klasse
1	When the Starbucks barista is cute but it's the female barista who calls out "next!" 텯 #rude	Positiv
2	You know it's a real friendship when she remembers your drink at Starbucks 😭 😭 @aliana	Positiv

gewählten 66 Emojis vorgenommen. Ein Abgleich mit den ermittelten Stimmungswerten bestätigt die Vermutung, dass überwiegend positive Stimmungswerte für die Happy faces und überwiegend negative Stimmungswerte für die Sad faces vorliegen.<sup>223</sup> Ausnahmen hierbei bilden lediglich die in Tab. 18 aufgeführten Emojis.

So ist das Emoji Sleeping face der Gruppe Happy faces zugeordnet, weist aber einen leicht negativen Stimmungswert auf. Bei einer stichprobenartigen Überprüfung der entsprechenden Tweets wird deutlich, dass das Emoji häufig eingesetzt wird, um dem Gefühl der Übermüdung Ausdruck zu verleihen. Die Konsumenten berichten über ihren Schlafmangel und sind in einer quengeligen Verfassung. Dies erklärt den negativen Stimmungswert des Emojis. Gleichzeitig zei-

<sup>&</sup>lt;sup>223</sup>Die Zuordnung aller Emojis ist im Anhang A (S. 85) dargestellt.



**Abbildung 10:** Positive Stimmungskategorie (Starbucks): Emojis mit positivem Stimmungswert in Abhängigkeit von ihrer Nutzungshäufigkeit. Die Häufigkeit ist aus Gründen der Übersichtlichkeit logarithmisch skaliert.

**Tabelle 18:** Emojis, bei denen die Zuordnung in die Emoji-Gruppen Happy faces und Sad faces nicht mit dem ermittelten Stimmungswert übereinstimmt. Die letzten beiden Spalten führen die Anzahl der dem Emoji zugeordneten negativen und positiven Tweets auf.

Emoji	Benennung	Emoji-Gruppe	Stimmungs- wert	Häufigkeit in %	Anz. Negativ	Anz. Positiv
, zzz	Sleeping face	Нарру	-0,14	0,89	127	96
<b>e</b>	Smiling face with open mouth and cold sweat	Нарру	-0,01	2,3	376	370
(xx)	Astonished face	Sad	0	0,1	15	15
	Face screaming in fear	Sad	0,23	0,53	63	101
••	Face with open mouth	Sad	0,47	0,09	8	22

gen die Textnachrichten in Tab. 19, dass die Verwendung eines Emojis mit negativem Stimmungswert nicht zwangsläufig auf ein negatives Markenimage schließen lässt. Im konkreten Fall scheinen die Konsumenten vielmehr von der Marke Starbucks überzeugt zu sein und zu glauben, dass nur ein Starbucks-Kaffee ihnen helfen kann, den negativ empfundenen Gefühlszustand der Übermüdung zu überwinden.

Das Emoji Smiling face with open mouth and cold sweat weist einen Stimmungswert von knapp unter null auf und wird somit gleichermaßen in negativ sowie positiv klassifizierten Tweets eingesetzt. Anhand der in Tab. 20 exemplarisch ausgewählten Tweets wird erkennbar, dass dieses Emoji eine Art Verlegenheit ausdrückt. Das Gefühl der Verlegenheit kann sowohl bei negativ wie auch positiv empfundenen Erlebnissen auftreten. Dies erklärt die schwache Polarität und den Stimmungswert nahe null.

Ähnlich verhält es sich für das Emoji Astonished face mit einem ermittelten Stimmungswert von null. Die in Tab. 21 dargestellten Tweets zeigen, dass das Emoji zum Einsatz kommt, um dem Gefühl der Überraschung bzw. dem Erstaunen Ausdruck zu verleihen. Diese Gefühlsregungen können gleichermaßen in positiven sowie negativen Kontexten auftreten. Allerdings wird dieses Emoji in den Ergebnissen insgesamt nur 30-mal aufgeführt. Es ist somit fraglich, ob der neutrale Stimmungwert aussagekräftig ist. Die geringe Nutzung des Emojis deutet zudem darauf hin, dass es vielen Konsumenten nicht vertraut ist bzw. eine eindeutige Interpretation schwer fällt.

Dies gilt auch für das Emoji Face with open mouth •• , das ebenfalls nur 30 Nen- nungen aufweist. In den stichprobenartig überprüften Tweets drücken Konsumenten vorwiegend ihre freudige Überraschung bzw. Vorfreude aus (s. Tab. 22). Dies erklärt den stark positiven Stimmungswert von 0,47.

Das Emoji Face screaming in fear in wird wider Erwarten vorherrschend in positiv klassifizierten Tweets verwendet. Anhand der Tweets in Tab. 23 wird deutlich, dass das Emoji weniger in angsterfüllten, für den Konsumenten als unangenehm wahrgenommenen Situationen Anwendung findet. Vielmehr wird es genutzt, um die angenehme, freudige

Tabelle 19: Exemplarische Tweets zu dem Emoji Sleeping face

Nr.	Textnachricht	Klasse
1	I got up way too early or stayed up way too late when my grande @Starbucks latte can't wake me up! $\center{\mathfrak{S}}$	Negativ
2	Need a Starbucks in my life 😌 4 hours sleep is not good	Negativ

Tabelle 20: Exemplarische Tweets zu dem Emoji Smiling face with open mouth and cold sweat

Nr.	Textnachricht	Klasse
1	Some days you go for a run, some days you go to Starbucks. It's called € https://t.co/zXvv01l3jx	Positiv
2	The lady at Starbucks not only knew my order but commented that I came in later than usual. I should probably stop going so much 😅	Negativ

Tabelle 21: Exemplarische Tweets zu dem Emoji Astonished face

Nr.	Textnachricht	Klasse
1	Why is there no starbucks in Stephen's city?! = 😵 ?!	Negativ
2	Starbucks is fueling our s'mores obsession with their new S'mores Frappuccino with marshmallow infused whipped cream 😵	Positiv

Tabelle 22: Exemplarische Tweets zu dem Emoji Face with open mouth

Nr.	Textnachricht	Klasse
1	Can't believe a stranger in Starbucks just bought my drink for me as an act of kindness! There are nice people in the world 😯 #happymonday	Positiv
2	Nice! #starbucks in #suttoncoldfield 13th-22nd May. Frappuccinos are half price between 3 and 5pm 😯 I'll be back! 😎 https://t.co/Vo24RTrHA5	Positiv

Verwunderung auszudrücken. Zur Bekräftigung dieser Freude, evtl. auch um Missverständnisse zu vermeiden, wird in den aufgezeigten Beispielen ein zweites Emoji aus der Gruppe Happy faces verwendet.

An dieser Stelle bietet es sich somit an, zusätzlich einen Blick auf Emoji-Kombinationen zu richten.<sup>224</sup> Tab. 24 zeigt die am häufigsten auftretenden Kombinationen.

Wie die in Tab. 25 exemplarisch aufgeführten Tweets verdeutlichen, bestätigt die stichprobenartige Überprüfung der Tweets die Vermutung, dass die Verwendung von jeweils gleichgerichteten polaren Emojis die Aussage des Tweets verstärkt (s. Textnachricht Nr. 1). Bei Emojis mit gegensätzlicher Polarität spielt einerseits der Aspekt der Ironie eine Rolle (s. Textnachricht Nr. 2). Andererseits wird, wie zuvor aufgezeigt, durch das zweite Emoji die Bipolarität der Emotion, d. h. ob sie als angenehm oder unangenehm empfunden wird, zum Ausdruck gebracht (s. Textnachricht Nr. 3).

Obwohl Auffälligkeiten und Interpretationsbedarf zu den ermittelten Stimmungswerten vorliegen, lässt sich dennoch über alle Stimmungswerte eine ausgeprägte Polarität erkennen, die eine Kategorisierung der Emojis in eine positive oder negative Klasse rechtfertigt. Diese Polarität wird im dargestellten Histogramm verdeutlicht (s. Abb. 11). Hinsichtlich der Häufigkeitsverteilung der Emojis ergibt sich eine Lastigkeit im positiven Polaritätsbereich und damit ein erster Hinweis bzgl. des wahrgenommenen Markenimages durch die Konsumenten.

Der Mittelwert der im Rahmen der Stimmungsanalyse ermittelten Emoji-Stimmungswerte kann zusätzlich als Indikator für die Bewertung des Images einer Marke herangezogen werden. Da sich die Emojis in ihrer Nutzungshäufigkeit jedoch stark unterscheiden, ist ein um diese Komponente gewichteter Mittelwert erforderlich. Der gewichtete Mittelwert für die Stimmung zur Marke Starbucks beträgt 0,37. Somit besteht eine überwiegend positive Stimmung. Die in Abb. 12 dargestellte Wordcloud stützt diese Aussage. Unter den Wörtern, die mindestens 500-mal in den Textnachrichten der

<sup>&</sup>lt;sup>224</sup>Das Skript zur Analyse der Mehrfachnennungen kann dem Zusatzmaterial (USB-Stick) entnommen werden. Dateiname: 05d Emoji-Co-Occurrence.

Tabelle 23: Exemplarische Tweets zu dem Emoji Face screaming in fear

Nr.	Textnachricht	Klasse
1	I just got 2 frappes from @Starbucks for \$5 for happy hour 🕡 😌 #LifeIsGood	Positiv
2	#starbucks green tea frap with hazelnut syrup ₩ so bombdigity good https://t.co/O3FLw5XPHo	Positiv

Tabelle 24: Top 5 der am häufigsten verwendeten Emoji-Kombinationen (Starbucks)

Kombination	Auftreten	Emoji	Stimmungswert	Emoji	Stimmungswert
1	624	To	-0,16	<b>(3)</b>	0,33
2	469	ToT	-0,16		0,81
3	343	ToT	-0,16		0,81
4	315	ToT	-0,16	33	-0,25
5	261	33	-0,25		0,81

Tabelle 25: Beispiel-Tweets mit mehreren verschiedenen Emojis

Nr.	Textnachricht	Klasse
1	If only Starbucks wasn't 20 minutes away 😩 😭	Negativ
2	Now the line at Starbucks is forever long. What did I do to deserve such bad karma 😂 😭	Negativ
3	The lady at Starbucks just made me so happy 釜 🙂	Positiv

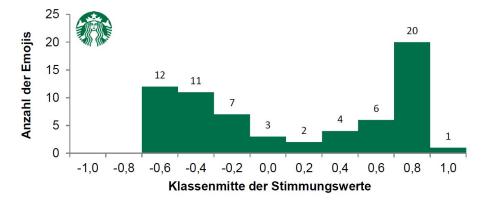


Abbildung 11: Histogramm (Starbucks)

Stichprobe verzeichnet wurden, sind vor allem stark positive Emotionswörter wie love, happy, like, good und free zu verzeichnen. Negativ assoziierte Wörter kommen hingegen kaum vor.<sup>225</sup> Auch die stichprobenartige Überprüfung sowie Interpretation der Tweets unterstreichen eine positive Einschätzung der Konsumenten und deuten auf ein gutes Image der Marke Starbucks hin.

### 4.4. Validierung der Ergebnisse

In diesem Kapitelabschnitt sollen die zuvor vorgestellten Ergebnisse nun mithilfe eines zweiten Datensatzes validiert werden. Hierfür werden Tweets zu dem Unternehmen McDonald's herangezogen. Die Marke McDonald's ist gemäß des Brandz Top 100 Report die wertvollste Marke im Fast-Food-Bereich. Bei Twitter belegt das Profil McDonald's in der Kategorie Lebensmittelhandel gemessen an der Anzahl der

 $<sup>^{225}\</sup>mathrm{Das}$  Skript zur Erstellung der Wordclouds ist dem Zusatzmaterial (USB-Stick) zu entnehmen. Dateiname: 08 Word frequency and word cloud.

<sup>&</sup>lt;sup>226</sup>Vgl. Schept (2015): 123.



Abbildung 12: Wordcloud (Starbucks)

Follower den zweiten Platz, s. Tab. 9 in Kap. 4.1. Wie dort beschrieben, wird vermutet, dass eine reziproke Beziehung zwischen Emotionen und dem Konsum von Lebensmitteln besteht. Somit kann analog zur Marke Starbucks davon ausgegangen werden, dass Tweets, die das Suchwort McDonald's enthalten, einen starken emotionalen Charakter aufweisen und sich für eine Stimmungsanalyse und somit für die Validierung der Ergebnisse besonders gut eignen. Um eine gleiche Ausgangsbasis zu sichern und den Einfluss unterschiedlicher allgemeiner Umweltbedingungen auszuschließen, erfolgte die Extraktion der Tweets zur Marke McDonald's zeitgleich mit der Erhebung der Starbucks-Tweets.

Insgesamt konnten im Beobachtungszeitraum mit 566.597 Tweets nur rund die Hälfte der Anzahl an Starbucks-Tweets abgerufen werden. Im Gegensatz zu McDonald's scheint es Starbucks gelungen zu sein, eine große Markengemeinschaft im Social Web aufzubauen, die aktiv Content generiert und verbreitet. 227

Nach Löschung der ReTweets und Selektion der Tweets, die mindestens einen der 66 zu untersuchenden Emojis enthalten, verbleiben 29.127 Tweets als Vergleichsdatenbasis. Da auch hier Kombinationen von Emojis in einem Tweet vorkommen, liegen insgesamt 34.232 Emoji-Tweet-Zuordnungen vor. Die Ergebnisse zu beiden Marken stimmen in vielen Aspekten überein. So ist auch bei McDonald's die Nutzungshäufigkeit von Emojis gering und liegt mit 5,14 Prozent sogar noch um 3,51 Prozentpunkte unter der für die Starbucks-Tweets. Damit bestätigt sich die Aussage, dass Emojis in der Twitter-Kommunikation zu bekannten Marken aus dem Lebensmittelbereich aktuell nur wenig eingesetzt werden.

Nahezu identisch für beide Marken ist die Tendenz, pro Tweet nur ein Emoji einzusetzen. Auch für McDonald's liegt diese Quote bei 85 Prozent der Textnachrichten. Bei Mehrfachanwendung dominiert mit rund 12 Prozent der Tweets ebenfalls die Kombination von zwei Emojis. Die geringe Verwendung von Emoji-Kombinationen könnte auf die von Twitter diktierte Zeichenbegrenzung zurückzuführen sein.

Auch im Rahmen der Analyse der zehn am häufigsten genutzten Emojis (s. Abb. 13) zeigen sich hinsichtlich des Vorkommens deutliche Parallelen zu den Starbucks-Ergebnissen. Zum einen haben sich im täglichen Gebrauch ebenfalls nur wenige Emojis durchgesetzt. Der Anteil der getwitterten Top 10 Emojis in Bezug auf alle 66 untersuchten Emojis beträgt 63,76 Prozent und entspricht damit nahezu exakt dem Ergebnis der Starbucks-Daten.

Zum anderen zeigt sich im Rankingvergleich der Top 10, dass acht der 66 Emojis jeweils in beiden Aufstellungen gelistet sind, wenn auch mit leicht unterschiedlichen Listenplätzen. Tab. 27 stellt die berechneten Stimmungswerte, die Platzierungen und die Häufigkeiten der Top 10 Emojis für beide Marken gegenüber.

Deutlich erkennbar ist, dass das Emoji Face with tears of joy 😂 im Hinblick auf die Nutzungshäufigkeit das Ranking beider Marken dominiert. Während bei Starbucks jedoch ein positiver Stimmungswert ermittelt werden kann, ergibt sich bei McDonald's ein leicht negativer Wert für das Emoji. Zudem fällt auf, dass unter den Top 10 Emojis der McDonald's-Tweets die beiden, jeweils mit einem hohen negativen Stimmungswert bewerteten Emojis Unamused face 💆 und Face with medical mask  $\bigcirc$  gelistet sind. Insgesamt sind in den Top 10 der McDonald's-Emojis sechs mit z. T. stark ausgeprägtem negativen Stimmungswert vertreten. Bei den Top 10 der Starbucks-Emojis sind es lediglich drei Emojis, für die ein negativer Stimmungswert verzeichnet wurde. Des Weiteren ist zu erkennen, dass die Stimmungswerte der Emojis der negativen Kategorie hinsichtlich ihrer Polarität deutlich ausgeprägter für McDonald's und in der positiven Kategorie deutlich ausgeprägter für Starbucks ausfallen. Dies wird als Hinweis darauf gewertet, dass McDonald's von den Konsumenten negativer wahrgenommen wird als Starbucks. In der Presse wird McDonald's u. a. als "[...] fettiges, fleischiges, bil-

Die fünf meist verwendeten Emoji-Kombinationen sind in Tab. 26 abgebildet. Wie auch bei Starbucks, treten die Emojis Face with tears of joy am häufigsten zusammen auf.

<sup>&</sup>lt;sup>227</sup>Vgl. Postinett (2011).

Tabelle 26: Top 5 der am häufigsten verwendeten Emoji-Kombinationen (McDonald's)

Kombination	Häufigkeit	Emoji	Stimmungswert	Emoji	Stimmungswert
1	454		-0,04	Tot	-0,58
2	154		-0,04	33	-0,57
3	103	Tot	-0,58		-0,57
4	91		-0,04		0,66
5	77		-0,04	•	0,75

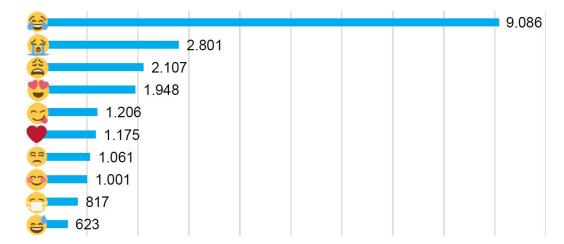


Abbildung 13: Top 10 der am häufigsten verwendeten Emojis (McDonald's)

liges Fast Food"<sup>228</sup> betitelt und kann damit den Trend eines gesunden, fleischfreien und individuellen Lebensmittelkonsums regionaler Produkte nicht bedienen. Neben sinkenden Umsätzen und schwindenden Marktanteilen gerät das Unternehmen bereits seit einiger Zeit in Kritik, genetisch verändertes und mit Chemikalien versetztes Fleisch zu vertreiben. Zudem führen ständige Preiserhöhungen und die angebliche Ausbeutung von Mitarbeitern zu negativen Schlagzeilen. <sup>229</sup>

Die folgenden beiden Abbildungen 14 und 15 zeigen die Stimmungswerte der 66 Emojis in Abhängigkeit ihrer Nutzungshäufigkeit. Das Emoji Face with cold sweat weist mit -0,85 den niedrigsten Stimmungswert auf. Mit einem Wert von 1,0 erzielen die Emojis Kissing face und Black Heart Suit den höchsten Stimmungswert. Allerdings sei angemerkt, dass für die Berechnung des Stimmungswertes dieser beiden Emojis nur wenige Tweets herangezogen werden konnten. Der Wert ergibt sich auf Basis von einem (Kissing face ) bzw. auf Basis von fünf (Black Heart Suit ) als positiv klassifizierten Tweet/s. Die Aussagekraft des Stimmungswertes ist daher als gering anzusehen. Der über alle Emojis gewichtete Mittelwert beträgt -0,04.

Im Folgenden liegt der Fokus auf der Übereinstimmung der Stimmungskategorien bzw. -werte im Hinblick auf alle 66 in der Arbeit berücksichtigten Emojis. Der Nachweis einer hohen Übereinstimmung deutet darauf hin, dass ein überwiegend einheitliches Verständnis in Hinblick auf die Stimmungskategorisierung der einzelnen Emojis vorliegt und sie zur Messung von Stimmungen markenübergreifend geeignet sind

Die Ergebnisse zeigen, dass 62 von 66 Emojis derselben Stimmungskategorie zugewiesen sind. Bei lediglich vier Emojis werden unterschiedliche Stimmungskategorien ermittelt. Sie sind mit ihren jeweiligen Stimmungswerten und Nutzungshäufigkeiten in der Reihenfolge ihrer Abweichung in Tab. 28 aufgeführt.

Um die Unterschiedlichkeit und die sich dahinter abzeichnenden Tendenzen zu untersuchen, wurden die den Emojis zugeordneten Tweets stichprobenartig überprüft und hypothetische Erklärungsansätze für die abweichende Stimmungskategorie ausgearbeitet. Vorweg ist festzuhalten, dass für alle diese Emojis gilt, dass sie in ihrem emotionalen Ausdruck sowohl in positiven als auch in negativen Kontexten Anwendung finden können.

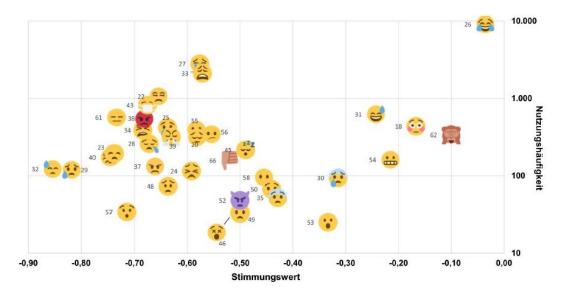
Das Emoji Face with open mouth weist die deutlichste Abweichung auf und wird als Ausdruck des Erschreckens bzw. Erstaunens genutzt, der sowohl positiv als auch negativ empfunden werden kann. Das schockierende Erlebnis scheint im Zusammenhang mit McDonald's, das freudige Ereignis hingegen im Zusammenhang mit Starbucks im Vordergrund

<sup>&</sup>lt;sup>228</sup>Zdrzalek (2015).

<sup>&</sup>lt;sup>229</sup>Vgl. Zdrzalek (2015); Esser (2014); Szarek (2012).

Tabelle 27: Gegenüberstellung der Top 10 Emojis bei McDonald's (McD) und Starbucks (Stb).

	, ,					
Bezeichnung			,		,	,
	Stb		McD		Stb	McD
Face with tears of joy	$\vdash$	16,38	1	26,54	0,33	-0,04
Smiling face with heartshaped eyes	2	10,09	4	2,69	0,81	99,0
Loudly crying face	က	9,37	2	8,18	-0,16	-0,58
Weary face	4	6,21	3	6,16	-0,25	-0,57
Heavy Black Heart	2	5,22	9	3,43	0,81	0,75
Smiling face with smiling eyes	9	4,6	8	2,92	0,78	0,53
Face savouring delicious food	7	4,11	2	3,52	0,81	0,65
Two hearts	∞	3,12	1	ı	0,84	ı
Grinning face with smiling eyes	6	2,33	1	ı	0,71	ı
Smiling face with open mouth and cold sweat	10	2,3	10	1,82	-0,01	-0,24
Unamused face	i	•	7	3,1	1	-0,65
Face with medical mask	'	'	6	2,39	ı	-0,68
	Bezeichnung Face with tears of joy Smiling face with heartshaped eyes Loudly crying face Weary face Heavy Black Heart Smiling face with smiling eyes Face savouring delicious food Two hearts Grinning face with smiling eyes Smiling face with open mouth and cold sweat Unamused face Face with medical mask	rs of joy with heartshaped eyes g face Heart with smiling eyes with smiling eyes with smiling eyes with open mouth and cold sweat ce dical mask	Stb rs of joy with heartshaped eyes g face 3 Heart heart 5 with smiling eyes e with smiling eyes with smiling eyes with open mouth and cold sweat 10 ce ce - dical mask	Stb       M         rs of joy       1       16,38         with heartshaped eyes       2       10,09         g face       3       9,37         Heart       4       6,21         with smiling eyes       6       4,6         with smiling eyes       6       4,11         with smiling eyes       9       2,33         with open mouth and cold sweat       10       2,3         ce       -       -         dical mask       -       -         r       -       -	sto joy       1       16,38       1       2         with heartshaped eyes       2       10,09       4         g face       3       9,37       2         g face       4       6,21       3         Heart       5       5,22       6         with smiling eyes       6       4,6       8         with smiling eyes       7       4,11       5         e with smiling eyes       9       2,33       -         with open mouth and cold sweat       10       2,3       10         ce       -       -       7         dical mask       -       -       9	so fjoy       1       McD         with heartshaped eyes       2       10,09       4       5,69         yface       3       9,37       2       8,18         yface       4       6,21       3       8,18         Heart       5       5,22       6       3,43         with smiling eyes       6       4,6       8       2,92         with smiling eyes       7       4,11       5       3,52         with open mouth and cold sweat       10       2,33       -       -         ce       -       7       3,1         ce       -       9       2,39



**Abbildung 14:** Negative Stimmungskategorie (McDonald's): Emojis mit negativem Stimmungswert in Abhängigkeit von ihrer Nutzungshäufigkeit. Die Häufigkeit ist aus Gründen der Übersichtlichkeit logarithmisch skaliert. Das Emoji *Dizzy face* hat denselben Koordinatenwert wie das Emoji*Frowning face with open mouth* und wird daher linksversetzt angezeigt.



**Abbildung 15:** Positive Stimmungskategorie (McDonald's): Emojis mit positivem Stimmungswert in Abhängigkeit von ihrer Nutzungshäufigkeit. Die Häufigkeit ist aus Gründen der Übersichtlichkeit logarithmisch skaliert. Das Emoji *Kissing face* hat denselben Koordinatenwert wie das Emoji *Black Heart Suit* und wird daher linksversetzt angezeigt.

Tabelle 28: Emojis in unterschiedlichen Stimmungskategorien

Emoji	Bezeichnung	Abweichung	Stimmungsw	ert (l) ι	ınd Häufig	Häufigkeit in % (r)	
			Starl	oucks	McDo	nald's	
••	Face with open mouth	0,80	0,47	0,1	-0,33	0,1	
9	See-no-evil monkey	0,47	0,37	1	-0,1	1	
	Grimacing face	0,44	0,22	0,6	-0,22	0,5	
	Face with tears of joy	0,37	0,33	16,4	-0,04	26,5	

zu stehen. Ähnlich verhält es sich bei dem Emoji-Affen Seeno-evil monkey 🔲 . Ein Erklärungsansatz ist, dass es zum Ausdruck von Peinlichkeit oder Scham verwendet wird. Im negativen Kontext genutzt, unterstreicht das Emoji das Erlebnis, während im positiven Kontext genutzt das peinliche Erlebnis humorvoll aufgenommen und als weniger schlimm empfunden wird. Auch hier überwiegt bei McDonald's die als negativ empfundene Peinlichkeit. Das Grimacing face 📛 fällt durch seine deutlich betonte Mundpartie auf. Diese weist Ähnlichkeit mit dem von Tieren gezeigten Zähnefletschen auf, das Ärger ausdrückt. Der negative Stimmungswert erklärt sich dadurch, dass dieses Emoji tatsächlich von den Kunden verwendet wird, um ihre Verärgerung deutlich zu machen. Gleichzeitig kann die Mundpartie auch als grinsendes Lachen gedeutet werden. Von Starbucks-Kunden wird das Emoji daher zum Ausdruck von Freude verwendet. Dies erklärt die Einordnung in die positive Stimmungskategorie.

Die drei oben beschriebenen Emojis werden mit einer maximalen Nutzungshäufigkeit von einem Prozent nur selten verwendet. Damit ist sowohl die Aussagekraft des jeweiligen Stimmungswertes eingeschränkt als auch die Bedeutung dieser Emojis gering.

Die Einordnung eines Emojis in unterschiedliche Stimmungskategorien zeigt, dass es zu Interpretationsschwierigkeiten kommen kann. Die emotionale Bedeutung eines Emojis scheint nicht eindeutig zu sein und kann im schlimmsten Fall zu Missverständnissen in der computergestützten Konsumentenkommunikation führen. Auch das bei beiden Marken am häufigsten genutzte Emoji Face with tears of joy befindet sich unter den Emojis, die hinsichtlich ihrer Stimmungskategorie nicht übereinstimmen. Wie im vorangegangenen Kap. dargelegt, kann es zum Ausdruck von Ironie verwendet werden und findet sich deshalb sowohl in positiven als auch in negativen Textzusammenhängen wieder. Aufgrund des leicht negativen Stimmungswertes bei McDonald's scheint hier die bittere Ironie zu überwiegen.

Zusammenfassend zeichnet sich ab, dass Emojis in Tweets zu McDonald's einen geringeren positiven bzw. höheren negativen Stimmungswert aufweisen. Der Vergleich der gewichteten Stimmungsmittelwerte beider Marken zeigt einen deutlich niedrigeren Wert für McDonald's (-0,04) als für Starbucks (0,37) und deutet darauf hin, dass die Marke McDonald's von den Konsumenten negativer wahrgenommen wird. Der parallele Blick auf die Wordcloud (s. Abb. 16) sowie auf das Histogramm (s. Abb. 17) bestätigen diesen Eindruck. Zum einen treten in der Wordcloud zu McDonald's deutlich weniger positiv assoziierte Emotionswörter auf als in der Wordcloud zur Marke Starbucks. Zum anderen zeigt die Häufigkeitsverteilung der Emojis eine deutliche Lastigkeit im negativen Polaritätsbereich.

Für die vorliegende Arbeit von entscheidender Bedeutung ist jedoch die Feststellung, dass die Emoji-Stimmungskategorien beider Marken zu 94 Prozent übereinstimmen. 62 der 66 untersuchten Emojis werden nach Durchführung der Tweetbasierten Stimmungsanalyse sowohl für McDonald's als auch für Starbucks in dieselbe Stimmungskategorie (Positiv bzw.

Negativ) eingeordnet. Die durchschnittliche absolute Messabweichung der Stimmungswerte beträgt lediglich 0,21.<sup>230</sup> Abschließend lässt sich somit festhalten, dass im Lebensmittelbereich vorherrschend ein markenübergreifendes einheitliches Verständnis zu den Emojis vorliegt und damit auch die Entwicklung einer Emoji-Stimmungskategorisierung möglich ist.

Auf Basis der ermittelten Stimmungswerte von minus eins bis plus eins ergibt sich zudem eine intensitätsmessende Emoji-Polaritätsskala. Hierfür wurde für jedes Emoji ein über beide Marken gemittelter Stimmungswert berechnet. <sup>231</sup> Eine sich ergebende Zuordnung der Emojis auf der Polaritätsskala zeigt Abb. 18. Die betrachteten Emojis sind in Abhängigkeit ihres Stimmungswertes (von -0,73 bis +1) angeordnet. Die Färbung der Balken (Polaritätsabstufung pro Farbe von 0,2) zeigt die Zugehörigkeit eines Emojis zu einer bestimmten Polaritätsstufe. Die Höhe der Balken gibt die Nutzungshäufigkeit der jeweiligen Emojis an und spiegelt damit deren Relevanz für die Konsumentenkommunikation wider.

### 5. Fazit

Obwohl Emojis als die am schnellsten wachsende Sprachform der Geschichte gelten, ist ihre emotionale Bedeutung empirisch bisher kaum untersucht. Auf Basis der theoretischen Grundlagen zur Emotionsforschung wurde deutlich, dass Emojis in der computergestützten Kommunikation bei der Übermittlung von Emotionen eine bedeutende Rolle einnehmen und das Textverständnis erhöhen. Ziel der vorliegenden Arbeit war es zu klären, ob in der Konsumentenkommunikation im Social Web ein einheitliches Verständnis für das jeweilig genutzte Emoji vorliegt und Emojis konkreten Emotionen und Stimmungen zuzuordnen sind. Hieraus würden sich neue Ansätze in der Emotionsund Marketingforschung ergeben. Neben den traditionellen Verfahren zur Emotionsklassifizierung und -messung intensivieren sich mit der Entwicklung des Social Webs die Forschungsbemühungen, Emotionen bzw. Stimmungen über textbasierte automatische Verfahren zu erfassen. In der Wissenschaft und Praxis haben sich automatische Stimmungsanalysen bewährt, die Texte hinsichtlich ihrer Polarität zu negativen, positiven und neutralen Stimmungsklassen zuordnen. Verfahren der Emotionsklassifizierung, die Texte auf zugrundeliegende konkrete Emotionskategorien analysieren, erzielen derzeit keine zufriedenstellenden Ergebnisse. Die vorliegende Arbeit fokussierte daher den Ansatz, Emojis über eine automatische stimmungsanalytische Betrachtung nach Polaritätsklassen zu kategorisieren. Untersucht wurden Emoji-enthaltende Kurznachrichten der Plattform Twitter.

<sup>&</sup>lt;sup>230</sup>Für die Berechnung wurden zunächst die absoluten Differenzen der jeweiligen Emoji-Stimmungswerte zu beiden Marken berechnet. Nachfolgend wurde die Summe aller errechneten Differenzen gebildet und durch die Gesamtanzahl der Emojis (66) dividiert.

<sup>&</sup>lt;sup>231</sup>Die über beide Marken gemittelten Stimmungswerte der Emojis können dem Zusatzmaterial (USB-Stick) entnommen werden. Dateiname: Berechnung Polaritätsskala.



Abbildung 16: Wordcloud (McDonald's)

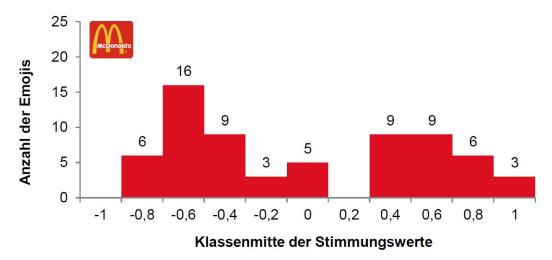


Abbildung 17: Histogramm (McDonald's)

Über die Klassifizierung der Texte in die Stimmungsklassen Negativ, Positiv und Neutral sollte eine entsprechende Emoji-Stimmungskategorisierung entwickelt werden.

Mithilfe der open-source Software R wurden die für die Datenerhebung, die Durchführung der automatischen Stimmungsanalyse sowie für die ersten statistischen Auswertungen erforderlichen Skripte eigens durch die Autoren programmiert. Innerhalb eines einmonatigen Datenerhebungszeitraumes konnten so 996.197 Tweets zur Marke Starbucks und zur Validierung der Ergebnisse 566.597 Tweets zur Marke McDonald's extrahiert werden. Die Wahl der beiden Marken begründet sich u. a. mit der reziproken Beziehung zwischen Emotionen und dem Konsum von Lebensmitteln sowie der Bedeutung beider Marken im Fast-Food-Bereich. Für die Entwicklung einer Stimmungskategorisierung wurden insgesamt 66 Emojis berücksichtigt. Auf Basis eines mit einer kleinen Stichprobe durchgeführten Pretests wurde ein Verfahren für die automatische Stimmungsanalyse ausgewählt.

Nach Durchführung der Stimmungsanalyse sowie der Emoji-basierten Auswertung der Starbucks-Tweets wurden die Ergebnisse mithilfe der Tweets zu McDonald's validiert und diskutiert. Dabei zeichnen sich die folgenden Kernergebnisse ab: Die Nutzungshäufigkeit der ausgewählten Emojis in Tweets zu bekannten Marken aus dem Lebensmittelsektor ist mit durchschnittlich knapp sieben Prozent gering. Werden Emojis in Tweets verwendet, verzichten rund 85 Prozent der Nutzer auf eine Mehrfachnennung und drücken ihre Emotion lediglich mit einem einzigen Emoji aus. Des Weiteren zeigt sich eine deutliche Dominanz vereinzelter Emojis in der Konsumentenkommunikation. Der Anteil der zehn meist getwitterten Emojis in Bezug auf alle 66 untersuchten Emojis beträgt rund 60 Prozent. Ferner weisen die am häufigsten verwendeten Emojis überwiegend positive Stimmungswerte auf. Diese Ergebnisse decken sich größtenteils mit den in Kap. 3.3 vorgestellten Erkenntnissen. Zudem belegt die Arbeit, dass Emojis im Lebensmittelsektor markenübergreifend überwiegend einheitlich verstanden werden. Insgesamt konnten 94 Prozent der untersuchten Emojis bei beiden Marken in dieselbe Stimmungskategorie (Positiv bzw. Negativ) eingeordnet werden.

Emojis erlangen somit für die Emotions- und Marketingforschung Bedeutung. Die Aussagekraft von Stimmungsanalysen kann Emoji-basiert erhöht werden. Auf Basis der ermittelten Stimmungswerte ließ sich zudem eine Emoji-

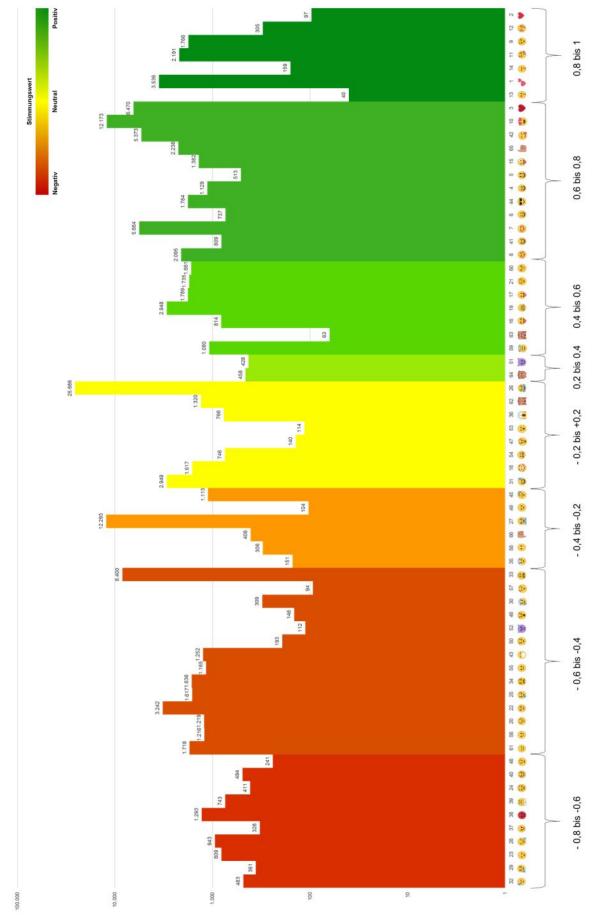


Abbildung 18: Emoji-Polaritätsskala: Der Stimmungswert der Emojis entspricht dem über beide Marken gemittelten Stimmungswert. Die Häufigkeit ist aus Gründen der Übersichtlichkeit logarithmisch skaliert und entspricht dem Vorkommen der Emojis in Tweets beider Marken. Insgesamt treten die Emojis bei beiden Marken in 135.573 Tweets auf.

Polaritätsskala von minus eins bis plus eins ableiten. Der jeweils für die Marke berechnete gewichtete Stimmungsmittelwert bildet einen Indikator für das von den Konsumenten wahrgenommene Image einer Marke. In der Gesamtbetrachtung verzeichnet Starbucks einen deutlich positiveren gewichteten Stimmungsmittelwert (0,37) als McDonald's (-0,04).

#### 5.1. Limitationen

Die identifizierten Limitationen beziehen sich auf drei Themenbereiche: die Eignung der Datenquelle Twitter, die Schwächen der für die Stimmungsanalyse zur Anwendung kommenden Verfahren sowie die vereinfachenden Schritte bzgl. des methodischen Vorgehens der Arbeit.

Trotz der in Kap. 3.2 genannten Vorzüge der Verwendung von Twitter als Datenquelle ergeben sich auch mit der Plattform verbundene Einschränkungen. So stellt Twitter keine nutzerbezogenen Informationen, wie bspw. Angaben zum Geschlecht und Alter, zur Verfügung. Der zu den Tweets angegebene Standort des Nutzers ist lückenhaft. Zum einen könnte es für das bessere Verständnis der Parasprache Emoji interessant sein, Unterschiede in der Nutzungshäufigkeit und der Stimmungskategorisierung eines Emojis in Bezug auf soziodemografische bzw. geografische Merkmale zu untersuchen. Vor dem Hintergrund markenpolitischer Aktionen und Reaktionen erschweren die fehlenden Daten zum anderen eine zielgruppenspezifische Ansprache. Ferner ist auf die Möglichkeit hinzuweisen, dass Tweets von sog. Fake-Profilen erstellt worden sein könnten, um damit die gegenüber einer Marke identifizierte Stimmung positiv oder negativ zu manipulieren. 232 Dies muss jedoch mit dem Argument des kostenlosen Zugangs zu einer großen und Vielfalt abbildenden Datenquelle abgewogen werden. Trotz der wachsenden Bedeutung von Twitter-Analysen in der Markt- und Verhaltensforschung ist zu beachten, dass die Daten nicht repräsentativ sind, da nicht jeder Konsument der beiden untersuchten Marken Twitter nutzt bzw. Tweets zu den Marken Starbucks oder McDonald's verfasst. Es wird jedoch angenommen, dass die in der Analyse nicht erfassten Konsumenten in Bezug auf die Emoji-Verwendung kein anderes Verhalten zeigen als die in der Analyse einbezogenen Konsumenten und somit eine Verallgemeinerung der Ergebnisse möglich ist.

Für die Durchführung der Stimmungsanalyse greift die vorliegende Arbeit auf das mit der Software R kompatible Softwaremodul Sentiment140 zurück. Der Quellcode dieses Moduls ist nicht frei zugänglich. Es bleibt somit teilweise intransparent, welche Features Sentiment140 für die Klassifizierung der Tweets berücksichtigt. Die Vorgehensweise des Klassifikators kann damit nicht vollständig nachvollzogen werden. So bleibt bspw. offen, ob der genauen Position eines für die Klassifizierung relevanten Wortes Bedeutung zugemessen wird. Einer Studie von Pang et al. zufolge nimmt auch die Wortposition Einfluss auf die Polarität eines Satzes<sup>233</sup> und damit auch auf die Zuordnung eines Tweets in

eine Stimmungskategorie. Auch die in Kap. 4.2 aufgezeigte starke Einflussnahme der durchgeführten Textbereinigung auf die Klassifikation kann nicht näher untersucht werden. Die Intransparenz erschwert eine qualitätssichernde und optimierende Klassifizierung der Tweets und kann folglich die Qualität der Emoji-Stimmungskategorisierung beeinflussen. Eine klärende Kontaktaufnahme der Autoren mit den Entwicklern Go et al. blieb erfolglos. 234 Generell ist festzuhalten, dass Stimmungen durch unzählige sprachliche Möglichkeiten zum Ausdruck kommen können. Technisch ist es sehr anspruchsvoll und derzeit noch nicht möglich, sämtliche Wörter, Regeln und Merkmale des Stimmungsausdrucks vollständig zu berücksichtigen.<sup>235</sup> Daher kann es zu Fehlinterpretationen und Fehlklassifizierungen kommen. Zudem sind Stimmungsanalyseverfahren, trotz ihrer stetigen Weiterentwicklung, derzeit kaum in der Lage, Sprachphänomene wie Ironie und Sarkasmus zufriedenstellend zu identifizieren.<sup>236</sup> Nicht korrekt zugeordnete Tweets verzerren die Emoji-Stimmungswerte.

Mit Blick auf die für die Arbeit zur Verfügung stehende Zeit sind die Limitationen bzgl. des methodischen Vorgehens der Arbeit größtenteils vereinfachenden Schritten geschuldet. Da die manuelle Stimmungsanalyse i. d. R. sehr zeitaufwendig ist, konnte im Rahmen des Pretests lediglich eine Stichprobe von 100 Tweets betrachtet werden. Diese ist nicht ausreichend, um qualitätssichernd ein optimales Stimmungsanalyseverfahren auszuwählen. Zudem sind die Beurteiler nicht geschult und keine englischen Muttersprachler, sodass sprachliche Besonderheiten in den Tweets, insbesondere Slang und Ironie, möglicherweise nicht richtig interpretiert wurden. Es besteht somit prinzipiell die Möglichkeit der falschen manuellen Zuordnung und einer Verzerrung der für die Auswahl des Verfahrens entscheidenden Kennziffern  $r_{neg,pos}$  und  $g_{neutr}$ . Auch konnten hinsichtlich der Textbereinigung nicht alle 28 Konfigurationsmöglichkeiten getestet werden. Um mögliche Qualitätseinbußen einzugrenzen, wurde deshalb lediglich die Komplexität in sinnvoller Weise erhöht.

Ferner ist darauf hinzuweisen, dass die Arbeit Emoji-Mehrfachnennungen nicht differenziert betrachtet hat. Ein in eine Stimmungskategorie eingeordneter Tweet wurde somit mehreren Emojis zugeordnet und kann damit zu Verzerrungen der Ergebnisse führen. Da jedoch nur rund 15 Prozent der klassifizierten Tweets Mehrfachnennungen enthalten, wird der Einfluss als gering eingeschätzt. Dennoch empfiehlt es sich für nachfolgende Studien, eine nach Mehrfachnennungen differenzierte Stimmungsanalyse durchzuführen und zu prüfen, ob sich hierdurch evtl. Veränderungen in den Emoji-Stimmungskategorien ergeben. Zudem wird die Aussagekraft des Stimmungswertes von Emojis mit einer nur geringen Nutzungshäufigkeit in den untersuchten

<sup>&</sup>lt;sup>232</sup>Vgl. Mostafa (2013): 4248.

<sup>&</sup>lt;sup>233</sup>Vgl. Pang et al. (2002): 83f.

 $<sup>^{234} \</sup>rm In$ ihrem Paper führen die Entwickler Go et al. auf, dass ihr Verfahren Unigrams und Bigrams berücksichtigt. Auf der Homepage ist jedoch vermerkt, dass der Klassifikator kontinuierlich weiterentwickelt wird. Welche Features bis dato verwendet werden, ist somit nicht nachvollziehbar.

<sup>&</sup>lt;sup>235</sup>Vgl. Liu (2015): 135f.

<sup>&</sup>lt;sup>236</sup>Vgl. Rosenthal et al. (2015): 459; Pang und Lee (2008): 38.

Tweets angezweifelt. Diese Werte sollten in nachfolgenden Studien überprüft werden.

Die Aussage von Novak et al., die nach einer manuellen Inhaltsanalyse von Tweets zu dem Ergebnis kommen, dass die Positionierung eines Emojis am Satzende mit einer höheren Polarität des Emojis einhergeht, <sup>237</sup> bleibt in der vorliegenden Arbeit unberücksichtigt. Der Grund hierfür ist die Komplexität der dafür notwendigen Programmierschritte in RStudio, die den vorgegebenen Rahmen der Arbeit überstiegen hätte.

Auch wenn die einmonatige Datenerhebung für beide in der Arbeit untersuchten Marken gleichzeitig erfolgte und sie damit den gleichen Umweltbedingungen unterliegt, ist es möglich, dass der Zeitraum der Erhebung in den Frühlingsmonaten Einfluss auf den Inhalt der Tweets hat. So könnte es sein, dass die ansteigende Temperatur zu einem höheren Konsum von to-go-Produkten, wie bspw. Starbucks-Kaffee, führt und sich gleichzeitig positiv auf die Grundstimmung der Menschen auswirkt. Folglich würden mehr Tweets in die positive Stimmungskategorie zugeordnet, als es in den Wintermonaten der Fall ist. Infolgedessen wäre es möglich, dass die Emoji-Polaritätsskala positiv verzerrt ist. Es empfiehlt sich für nachfolgende Studien deshalb, einen möglichen Einfluss des Erhebungszeitraumes auf die Emoji-Stimmungswerte empirisch zu überprüfen.

### 5.2. Implikationen für Wissenschaft und Praxis

Zunächst werden Implikationen für weitere Forschungsbemühungen aufgezeigt. Zum einen soll auf Verbesserungen der automatischen Stimmungsanalyseverfahren, zum anderen auf die Klärung der Allgemeingültigkeit des emotionalen Ausdrucks der Emojis eingegangen werden. Zum Abschluss werden Implikationen für die Praxis gegeben. Die vorliegende Arbeit hat gezeigt, dass eine Klassifizierung der Emojis zu den Stimmungskategorien Negativ und Positiv möglich ist. Die eindeutige Zuordnung der Emojis zu Emotionskategorien, bspw. Freude, Vertrauen oder Enttäuschung, scheint hingegen derzeit auf Basis von textbasierten Stimmungsanalvsen kaum möglich. Dies zeigen auch der Auszug der Ergebnisse der Emotionsklassifizierung (s. Anhang D, S. 111). Eine aussagekräftige und differenzierte Messung von Emotionen ist aber sowohl für die Emotions- wie die Marketingforschung von Bedeutung, z. B. für die gezielte Analyse der emotionalen Wirkung von Marketingaktivitäten. Die Forschungsbemühungen im Bereich des automatischen Emotion Minings sollten daher intensiviert werden. Ein weiterer Aspekt bei stimmungsanalytischen Betrachtungen ist die Identifizierung von Ironie. Im Rahmen der Arbeit wurde deutlich, dass Emojis für das Textverständnis von Bedeutung sind und u. a. als Ausdruck von Ironie Anwendung finden. Daher empfiehlt es sich, Stimmungsanalysen bzw. Klassifikatoren zu entwickeln, die sowohl lexikalische und syntaktische Features als auch den Emoji-Code als Merkmal für eine entsprechende Klassenzuordnung berücksichtigen. Solch ein Verfahren sollte für

Das Vorliegen eines größtenteils einheitlichen Verständnisses für Emojis bei den Konsumenten begründet die wachsende Bedeutung von Emojis für die Emotions- und Marketingforschung. Dennoch deuten die Ergebnisse des Pretests an, dass es bezüglich der Emojis zu Fehlinterpretationen kommen kann. Nachdem die Beurteiler die im Tweet verwendeten Emojis in ihre Beurteilung einbezogen hatten, nahm die Beurteiler-Übereinstimmung leicht ab. Des Weiteren zeigte sich im Rahmen der Ergebnisvorstellung, dass viele der betrachteten Emojis in der Social-Web-Kommunikation der Konsumenten kaum Anwendung finden. Hierfür könnten der mehrdeutige emotionale Ausdruck dieser Emojis und die mit ihrer Verwendung auftretenden Missverständnisse mögliche Erklärungsansätze sein. Für zukünftige Forschungsarbeiten empfiehlt es sich daher, den emotionalen Ausdruck und das Verständnis von Emojis differenzierter zu untersuchen. Da die vorliegende Arbeit sich lediglich auf Marken aus dem Lebensmittelsektor konzentriert, gilt es zum einen zu prüfen, ob sich die entwickelte Emoji-Stimmungskategorisierung mit Daten anderer, möglicherweise weniger emotionalen Sektoren bestätigen lässt. Zum anderen könnten die Ergebnisse mithilfe von anderssprachigen Tweets validiert werden. Dies bietet sich vor allem deshalb an, weil die Mimik der Emojis in verschiedenen Kulturkreisen aufgrund der sozialen Prägung möglicherweise unterschiedlich wahrgenommen wird. Allgemein sind die Erkenntnisse über kulturelle Gemeinsamkeiten und Unterschiede im Verständnis von Emojis noch unzureichend.<sup>238</sup> Da jedoch vermutet wird, dass der Einfluss interkultureller Unterschiede hinsichtlich der Interpretation der Emojis deutlich geringer ist als der Einfluss kultureller Normen auf die sprachliche Äußerung von erlebten Emotionen, bietet eine Emoji-Skala dennoch einen vielversprechenden Fortschritt für die Emotionsmessung. 239 In Kap. 2.2 wurden die Vorteile einer bildbasierten Skala zur Messung von Emotionen bereits dargelegt. Unter der Annahme, dass, zumindest für die meist verwendeten Emojis, ein kulturübergreifendes, universelles Verständnis bestätigt werden kann, könnte die Messung von Emotionen international und sprachenunabhängig Emoji-basiert durchgeführt werden.

Abschließend werden Implikationen für die Praxis vorgestellt. Neben der Durchführung Emoji-basierter Stimmungsanalysen können Unternehmen durch die Einbindung von Emojis in ihre Markenkommunikation profitieren. Emojis bieten als bildbetonte Kommunikationsmittel die Möglichkeit, Konsumenten in ihrer Wahrnehmung zu beeinflussen und den emotionalen Zusatznutzen einer Marke hervorzuheben. Insbesondere positive Emotionen begünstigten den Erfolg einer Marke und stellen den Unterschied zwischen starken und schwachen Marken dar. Dem Beispiel von McDonald's folgend (s. Abb. ??), können Unternehmen

die zuverlässige Identifizierung von Ironie durch weitere Forschungsarbeiten erprobt werden.

<sup>&</sup>lt;sup>237</sup>Vgl. Novak et al. (2015): 1.

<sup>&</sup>lt;sup>238</sup>Vgl. Vidal et al. (2016): 127.

<sup>&</sup>lt;sup>239</sup>Vgl. van Zyl und Meiselman (2015): 212; Vidal et al. (2016): 127.

<sup>&</sup>lt;sup>240</sup>Vgl. Kroeber-Riel und Esch (2011): 119, 216.

<sup>&</sup>lt;sup>241</sup>Vgl. Esch (2010): 70.

mit Emoji-geprägten Kommunikationsmaßnahmen außergewöhnliche und einzigartige Erlebnisse kreieren und sich damit von konkurrieren- den, funktional gleichwertigen Marken differenzieren.

Auch die von Twitter seit kurzer Zeit angebotene Emojibasierte Zielgruppenansprache unterstreicht die steigende Relevanz von Emojis in der Konsumentenkommunikation. Das Geschäftsmodell bietet Unternehmen die Möglichkeit, Twitter-Nutzer, gemäß ihrer jeweiligen durch Emojis ausgedrückten Stimmungen und Emotionen, gezielt durch Werbung anzusprechen. Dafür ist jedoch ein fundiertes Emoji-Verständnis unabdingbar, wofür die vorliegende Arbeit mit den ermittelten Emoji-Stimmungskategorien sowie der Emoji-Polaritätsskala eine gute Ausgangsbasis darstellt. Da Emojis neben Emotionen auch Interessen, Ideen und Konzepte ausdrücken, können Werbeanzeigen auch zu themenbasierten Emojis, wie bspw. Lebensmitteln, Kleidung, Schmuck oder Dienstleistungen, eingeblendet werden. Vor allem in Verbindung mit dem vom Konsumenten auf Twitter teilweise angegebenen Standort bietet sich personalisierte Werbung an. 242 So zeigte z. B. die detaillierte Analyse des Emojis *Sleeping face* <sup>©</sup> in Kap. 4.3, dass es meistens genutzt wird, um das als negativ empfundene Gefühl der Übermüdung ausdrücken. Starbucks könnte denjenigen Konsumenten, die dieses Emoji twittern, bspw. einen Rabattgutschein für einen Kaffee in der nächstgelegenen Starbucks-Filiale einblenden lassen und so die Aufmerksamkeit auf die Marke lenken. Eine Werbedominanz in der Produktkategorie Kaffee kann zudem dazu führen, dass die Marke Starbucks beim Kunden top of mind angesiedelt wird. Dies impliziert wettbewerbsbezogene Vorteile und stellt eine Möglichkeit dar, die Marktposition langfristig zu festigen.<sup>243</sup>

<sup>&</sup>lt;sup>242</sup>Vgl. Shah (2016); Bauer (2016).

<sup>&</sup>lt;sup>243</sup>Vgl. Gabler Wirtschaftslexikon (2016).

#### Literatur

- Aman, S. und Szpakowicz, S. Identifying expressions of emotion in text. In Matoušek, V. und Mautner, P., editors, Text, Speech and Dialogue, pages 196–205. Springer, 2007.
- Aßmann, S. und Röbbeln, S. Social Media für Unternehmen. Das Praxisbuch für KMU. Bonn: Galileo Computing, 2013.
- Averill, J. R. The emotions. An integrative approach. In: Robert Hogan (Hg.): Handbook of personality psychology. San Diego u. a.: Academic Press: 513-541, 1997.
- Barash, V. und Golder, S. Twitter. Conversation, entertainment, and information, all in one network! In: Derek Hansen (Hg.): Analyzing Social Media Networks with NodeXL. Insights from a Connected World. Unter Mitarbeit von Ben Shneiderman und Marc A. Smith. Amsterdam u. a.: Morgan Kaufmann: 143-164, 2011.
- Bauer, T. Stimmungsbarometer für Werber: Twitter erfindet Emoji Targeting, 2016. URL http://onlinemarketing.de/news/twitter-emoji-tar geting. Abruf am 20.08.2016.
- Becker, L., Erhart, G., Skiba, D., und Matula, V. AVAYA: Sentiment Analysis on Twitter with Self-Training and Polarity Lexicon Expansion. 2nd Joint Conference on Lexical and Computational Semantics, Vol. 2: Proceedings of the 7th International Workshop on Semantic Evaluation. Atlanta, GA: 333-340, 2013. URL http://www.aclweb.org/anthology/S13-2055. Abruf am 20.08.2016.
- Behrens, G. Konsumentenverhalten. Entwicklung, Abhängigkeiten, Möglichkeiten. 2. Aufl. Heidelberg: Physica-Verlag (Konsum und Verhalten, 18), 1991.
- Berger, J. und Heath, C. Where consumers diverge from others: Identity signaling and product domains. *Journal of Consumer Research*, 34(2): 121–134, 2007.
- Beutelsbacher, S. Wenn ein Shitstorm das Konzern-Image zerstört. In: Die Welt, 15.07.2011, 2011. URL http://www.welt.de/wirtschaft/webwelt/article13488539/Wenn-ein-Shitstorm-das-Konzern-Image-zerstoert.html. Abruf am 20.08.2016.
- Boiy, E. und Moens, M.-F. A machine learning approach to sentiment analysis in multilingual web texts. *Information Retrieval*, 12(5):526–558, 2009.
- Bollen, J., Mao, H., und Zeng, X.-J. Twitter mood as a stock market predictor. *Computer*, 44(10):91–94, 2011.
- Bradley, M. M. und Vrana, S. R. The startle probe in the study of emotion and emotional disasters. In: Niels Birbaumer und Arne Öhman (Hg.): The Structure of Emotion. Psychophysiological, cognitive and clinical aspects. Seattle u. a.: Hogrefe & Huber: 270-287, 1993.
- Bradley, M. M. und Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.
- Bruhn, M. Integrierte Unternehmens- und Markenkommunikation. Strategische Planung und operative Umsetzung. 6. Aufl. Stuttgart: Schäffer- Poeschel 2014
- Byrne, Z. S., Masterson, S. S., und Hurd, B. M. Fairness in a virtual world: The implications of communication media on employees' justice and trust perceptions. In: Chester Schriesheim und Linda L. Neider (Hg.): Perspectives on justice and trust in organizations. Charlotte, NC: IAP (Research in management, 9): 45-75, 2012.
- Canetti, L., Bachar, E., und Berry, E. M. Food and emotion. *Behavioural Processes*, 60(2):157–164, 2002.
- Cha, M., Haddadi, H., Benevenuto, F., und Gummadi, K. P. Measuring user influence in Twitter: The million follower fallacy. Proceedings of the 4th international AAAI conference on Weblogs and Social Media. Washington, DC: 10-17, 2010. URL https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/download/1538/1826. Abruf am 20.08.2016.
- Chaovalit, P. und Zhou, L. Movie review mining: a comparison between supervised and unsupervised classification approaches. 38th Hawaii International Conference on System Sciences: 1-9, 2005. URL https://www.computer.org/csdl/proceedings/hicss/2005/2268/04/22680112c.pdf. Abruf am 20.08.2016.
- Ciuk, D., Troy, A., und Jones, M. Measuring emotion: Self-reports vs. physiological indicators. Annual Meeting of the Midwest Political Science Association: 1-28, 2015. URL http://papers.srn.com/sol3/papers.cfm?abstract id=2595359. Abruf am 20.08.2016.
- Cruse, J. Emoji usage in TV conversation, 2015. URL https://blog

- .twitter.com/2015/emoji-usage-in-tv-conversation. Abruf am 20.08.2016.
- Culotta, A. und Cutler, J. Mining brand perceptions from Twitter social networks. *Marketing Science*, 35(3):343–362, 2016.
- Derks, D., Bos, A. E., und Von Grumbkow, J. Emoticons and social interaction on the internet: the importance of social context. *Computers in Human Behavior*, 23(1):842–849, 2007.
- Dery, M. Flame wars. The South Atlantic Quarterly, 92:559-568, 1993.
- Dieckmann, A., Hupp, O., Gröppel-Klein, A., Broeckelmann, P., und Walter, K. Beyond verbal scales: Measurement of emotions in advertising effectiveness research. *Yearbook of Marketing and Consumer Research*, 6(4): 72–99, 2008.
- Dimson, T. Emojineering part 1: Machine learning for emoji trends, 2015. URL http://instagram-engineering.tumblr.com/post/117889701472/emojineering-part-1-machine-learning-for-emoji. Abruf am 20.08.2016.
- Draft, R. L. und Lengel, R. H. Organizational information requirements, media richness and structural design. *Management Science*, 32(5):554– 571, 1986.
- Duden. Emoticon, 2016. URL http://www.duden.de/rechtschreibung /Emoticon. Abruf am 20.08.2016.
- Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., und Madden, M. Social media update 2014. In: Pew Reseach Center, 09.01.2015, 2015. URL http://www.pewinternet.org/2015/01/09/social-med ia-update-2014/. Abruf am 20.08.2016.
- Ekman, P. Universals and cultural differences in facial expressions of emotions. Hg. v. J. Cole. University of Nebraska Press. Lincoln, 1972.
- Ekman, P. und Friesen, W. V. Facial action coding system. Palo Alto, CA: Consulting Psychologists Press, 1978.
- Ekman, P. und Friesen, W. V. Unmasking the face. A guide to recognizing emotions from facial clues. Cambridge, MA: Malor Books, 2003.
- Ekman, P., Friesen, W. V., und Ellsworth, P. What emotion categories or dimensions can observers judge from facial behavior? In: Paul Ekman (Hg.): Emotion in the human face. 2. Aufl. Cambridge u. a.: Cambridge University Press (Studies in emotion and social interaction): 39-55, 1982.
- Ekman, P., Levenson, R. W., und Friesen, W. V. Autonomic nervous system activity distinguishes among emotions. *Science*, (221):1208–1210, 1983.
- Ekman, P., Friesen, W. V., und Hager, J. C. Facial action coding system. *Journal of Personality and Social Psychology*, 54:414–420, 2002.
- Ekman, P. und Friesen, W. V. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1):49–98, 1969.
- Emoji Research Team. 2015 Emoji Report. Hg. v. Emogi The Emotion Engine, 2015. URL http://emogi.com/documents/Emoji\_Report\_2015.pdf. Abruf am 20.08.2016.
- Emojipedia. How many emoji characters are there?, 2016a. URL http://emojipedia.org/faq/. Abruf am 20.08.2016.
- Emojipedia. Twitter Twemoji 1.0, 2016b. URL http://emojipedia.org/twitter/twemoji-1.0/. Abruf am 20.08.2016.
- Emojipedia. Unicode version 6.0, 2016c. URL http://emojipedia.org/unicode-6.0/. Abruf am 20.08.2016.
- Emojitracker. Realtime emoji use on Twitter, 2016. URL http://emojitracker.com/. Abruf am 20.08.2016.
- Esch, F.-R. Strategie und Technik der Markenführung. 6. Aufl. München: Vahlen, 2010.
- Esser, C. Ammoniumhydroxid. Starkoch Jamie Oliver gewinnt langen Kampf gegen McDonalds, 2014. URL http://www.business-on.de/ammoniumhydroxid-starkoch-jamie-oliver-gewinnt-langen-kampf-gegen-mcdonalds\_id45890.html. Abruf am 20.08.2016.
- Euler, H. A. Evolutionstheoretische ansätze. In: Jürgen H. Otto (Hg.): Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psychologie-Verlags-Union: 45-63, 2000.
- Ewert, O. Ergebnisse und Probleme der Emotionsforschung. In: Hans Thomae (Hg.): Theorien und Formen der Motivation. Göttingen u. a.: Verlag für Psychologie Hogrefe (Enzyklopädie der Psychologie: Themenbereich C, Serie IV, 1): 398-452, 1983.
- Fischer, D. Social Media Marketing und Strategien Facebook, Twitter, Xing & Co. erfolgreich nutzen. Berlin: E. Schmidt, 2015.
- Franke, M.-K. Der Konsument. Homo Emoticus statt Homo Oeconomicus? Hamburg: Springer Gabler, 2014.
- Gabler Wirtschaftslexikon. Stichwort: Top of mind, 2016. URL http://wirtschaftslexikon.gabler.de/Archiv/57506/top-o

- f-mind-v5.html. Abruf am 20.08.2016.
- Gentry, J. Package 'twitteR'. R Based Twitter Client, 2015. URL https://cran.r-project.org/web/packages/twitteR/twitteR.pdf. Abruf am 27.04.2016.
- Gezici, G., Dehkharghani, R., Yanikoglu, B., Tapucu, D., und Saygin, Y. SU-Sentilab: A Classification System for Sentiment Analysis in Twitter. 2nd Joint Conference on Lexical and Computational Semantics. Vol 2: Proceedings of the 7th International Workshop on Semantic Evaluation. Atlanta, GA: 471-477, 2013. URL http://www.aclweb.org/anthology/S13-2078. Abruf am 20.08.2016.
- Ghiassi, M., Skinner, J., und Zimbra, D. Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications*, 40(16):6266–6282, 2013.
- Go, A., Bhayani, R., und Huang, L. Twitter sentiment classification using distant supervision. Technical Report. Stanford University, 2009. URL http://www-cs.stanford.edu/people/alecmgo/papers/TwitterDistantSupervision09.pdf. Abruf am 28.08.2016.
- Go, A., Bhayani, R., und Huang, L. Sentiment 140, 2013a. URL http://help.sentiment140.com/. Abruf am 20.08.2016.
- Go, A., Bhayani, R., und Huang, L. Sentiment 140. For Academics, 2013b. URL http://help.sentiment140.com/for-students. Abruf am 20.08.2016.
- Go, A., Bhayani, R., und Huang, L. Sentiment 140. Site Functionality, 2013c. URL http://help.sentiment140.com/site-functionality. Abruf am 20.08.2016.
- Gründerszene. Application-Programming-Interface (API), 2016. URL http://www.gruenderszene.de/lexikon/begriffe/applicatio n-programming-interface-api. Abruf am 20.08.2016.
- Guenther, T. und Furrer, L. GU-MLT-LT: Sentiment analysis of short messages using linguistic features and stochastic gradient descent. 2nd Joint Conference on Lexical and Computational Semantics, Vol. 2: Proceedings of the 7th International Workshop on Semantic Evaluation. Atlanta, GA: 328-332, 2013. URL http://www.aclweb.org/anthology/S13-2054. Abruf am 20.08.2016.
- Haimerl, E. Emotionale oder rationale Werbung? über einen verhängnisvollen Irrtum in Kommunikation und Werbeforschung. *Jahrbuch der Absatzund Verbrauchsforschung*, (1):4–30, 2007.
- Han, J., Pei, J., und Kamber, M. Data mining: Concepts and techniques. 3. Aufl. Elsevier, Amsterdam, 2012.
- Hassan, A., Abbasi, A., und Zeng, D. Twitter sentiment analysis: A bootstrap ensemble framework. International Conference on Social Computing Alexandria, VA: 357-364, 2013. URL https://www.researchgate.net/publication/261123105\_Twitter\_Sentiment\_Analysis\_A\_Bootstrap\_Ensemble\_Framework. Abruf am 20.08.2016.
- Heymann-Reder, D. Social-Media-Marketing. Erfolgreiche Strategien für Sie und Ihr Unternehmen. München u. a.: Addison-Wesley, 2011.
- Huang, A. H., Yen, D. C., und Zhang, X. Exploring the potential effects of emoticons. *Information & Management*, 45(7):466–473, 2008.
- Isen, A. M. Toward understanding the role of affect in cognition. Robert S. Wyer (Hg.): Handbook of Social Cognition. Hillsdale, NJ u. a.: Erlbaum (3), pages 179–236, 1984.
- Izard, C. E. The psychology of emotions. New York, NY: Plenum Press (Emotions, personality, and psychotherapy), 1991.
- Izard, C. E. Die Emotionen des Menschen: eine Einführung in die Grundlagen der Emotionspsychologie. 2. Aufl. Weinheim u. a.: Psychologie-Verlags-Union, 1994.
- Izard, C. E. The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review*, 2(4):363–370, 2010.
- Jansen, B., Zhang, M., Sobel, K., und Chowdury, A. Micro-blogging as online word of mouth branding. CHI'09 Extended Abstracts on Human Factors in Computing Systems. Boston, MA: 3859-3864, 2009a. URL https://faculty.ist.psu.edu/jjansen/academic/pubs/jans en\_Micro\_blogging\_as\_Online\_Word\_of\_Mouth\_Branding.pdf. Abruf am 28.08.2016.
- Jansen, B. J., Zhang, M., Sobel, K., und Chowdury, A. Twitter power: Tweets as electronic word of mouth. *Journal of the Association for Information Science and Technology*, 60(11):2169–2188, 2009b.
- Jockers, M. Package 'syuzhet'. Extracts sentiment and sentiment derived plot arcs from text, 2016. URL https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf. Abruf am 27.04.2016.

- Kaiser, S. und Wehrle, T. Ausdruckspsychologische Methoden. In: Jürgen H. Otto (Hg.): Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psychologie-Verlags-Union: 419-428, 2000.
- Kaplan, A. M. und Haenlein, M. The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons*, 54(2):105– 113, 2011.
- Keller, K. L. Strategic brand management. Building, measuring, and managing brand equity. 3. Aufl. Upper Saddle River, NJ: Pearson/Prentice Hall, 2008.
- Keltner, D., Oatley, K., und Jenkins, J. M. Understanding emotions. 3. Aufl. Hoboken, NJ: Wiley. 2014.
- Kenning, P. Consumer neuroscience. Ein transdisziplinäres Lehrbuch. Stuttgart: Kohlhammer, 2014.
- Kiesler, S., Siegel, J., und McGuire, T. W. Social psychological aspects of computer-mediated communication. *American Psychologist*, 39(10): 1123–1134, 1984.
- Kim, A. J. und Ko, E. Do social media marketing activities enhance customer equity? an empirical study of luxury fashion brand. *Journal of Business Research*, 65(10):1480–1486, 2012.
- Kiritchenko, S., Zhu, X., und Mohammad, S. M. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, 50:723–762, 2014
- Kleinginna, P. R. und Kleinginna, A. M. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 5(4):345–379, 1981.
- Kontopoulos, E., Berberidis, C., Dergiades, T., und Bassiliades, N. Ontologybased sentiment analysis of Twitter posts. *Expert Systems with Applications*, 40(10):4065–4074, 2013.
- Kroeber-Riel, W. und Esch, F-R. Strategie und Technik der Werbung. Verhaltens- und neurowissenschaftliche Erkenntnisse. 7. Aufl. Stuttgart: Kohlhammer, 2011.
- Kroeber-Riel, W. und Gröppel-Klein, A. Konsumentenverhalten. 10. Aufl. München: Vahlen, 2013.
- Kroeber-Riel, W. und Weinberg, P. Konsumentenverhalten. 6. Aufl. München: Vahlen. 1996.
- Kroeber-Riel, W. und Weinberg, P. Konsumentenverhalten. 7. Aufl. München: Vahlen, 1999.
- Leahu, L., Schwenk, S., und Sengers, P. Subjective objectivity: Negotiating emotional meaning. Proceedings of the 7th ACM Conference on Designing Interactive Systems. Cape Town, South Africa: 425-434, 2008. URL http://dl.acm.org/citation.cfm?id=1394491&dl=ACM&coll=DL&CFID=661693401&CFTD-KEN=19043322. Abruf am 30.08.2016.
- Liu, B. Sentiment analysis. Mining Opinions, Sentiments, and Emotions. New York u. a.: Cambridge University Press, 2015.
- Macht, M. How emotions affect eating: a five-way model. *Appetite*, 50(1): 1–11, 2008.
- MacMillan, G. Foodie tweets. 10 facts about Twitter, restaurants and food, 2014. URL https://blog.twitter.com/en-gb/2014/foodie-tweets-10-facts-about-twitter-restaurants-and-food. Abruf am 20.08.2016.
- Marchand, M., Ginsca, A. L., Besancon, R., und Mesnard, O. [IVIC-LIMSI]: Using syntactic features and multi-polarity words for sentiment analysis in Twitter. 2nd Joint Conference on Lexical and Computational Semantics, Vol. 2: Proceedings of the 7th International Workshop on Semantic Evaluation. Atlanta, GA: 418-424, 2013. URL https://aclweb.org/anthology/S/S13/S13-2069.pdf. Abruf am 30.08.2016.
- Mattscheck, M. Definition Microblogs/Microblogging, 2016. URL http://www.onlinemarketing-praxis.de/glossar/microblogs-microblogging. Abruf am 20.08.2016.
- Mau, G. Die Bedeutung der Emotionen beim Besuch von Online-Shops. Messung, Determinanten und Wirkungen. Wiesbaden: Gabler, 2009.
- McDougall, W. Introduction to social psychology. London: Meuthen & Co, 1924.
- Mehrabian, A. und Russell, J. A. An approach to environmental psychology. Cambridge, Mass. u. a.: MIT Press, 1974.
- Merten, J. Einführung in die Emotionspsychologie. Stuttgart: Kohlhammer, 2003.
- Meyer, W.-U., Schützwohl, A., und Reisenzein, R. Einführung in die Emotionspsychologie. Band 1. Bern u. a.: Hans Huber, 1993.
- Michaeli, R. Competitive intelligence. Strategische Wettbewerbsvorteile erzielen durch systematische Konkurrenz-, Markt- und Technologieanaly-

- sen. Berlin, Heidelberg: Springer, 2006.
- Mohammad, S. M. Portable features for classifying emotional text. 2012 Conference of the North American Chapter of the Association for Computational Linguistics. Montréal, Canada: 587-591, 2012. URL http://www.aclweb.org/anthology/N12-1071. Abruf am 20.08.2016.
- Mohammad, S. M. NRC Hashtag Emotion Lexicon, 2013a. URL http://sa ifmohammad.com/WebPages/lexicons.html. Abruf am 05.08.2016.
- Mohammad, S. M. NRC Word-Emotion Association Lexicon, 2013b. URL http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon. htm. Abruf am 20.08.2016.
- Mohammad, S. M., Kiritchenko, S., und Zhu, X. NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. 2nd Joint Conference on Lexical and Computational Semantics, Vol. 2: Proceedings of the 7th International Workshop on Semantic Evaluation. Atlanta, GA: 321-327, 2013. URL http://www.aclweb.org/anthology/S13-2053. Abruf am 30.08.2016.
- Möll, T. und Esch, F.-R. Emotionen machen den Unterschied. Absatzwirtschaft, 7:34–37, 2008.
- Moore, A., Steiner, C. M., und Conlan, O. Design and development of an empirical smiley-based affective instrument. 1st Workshop on Emotions and Personality in Personalized Services will be organized in conjunction with UMAP. Rome, Italy, 2013. URL http://ceur-ws.org/Vol-997/empire2013\_paper\_4.pdf. Abruf am 20.08.2016.
- Mori, K. und Mori, H. Another test of the passive facial feedback hypothesis: When your face smiles, you feel happy. *Perceptual and Motor Skills*, 109 (1):76–78, 2009.
- Mostafa, M. M. More than words: Social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, 40(10):4241–4251, 2013.
- Naylor, R. W., Lamberton, C. P., und West, P. M. Beyond the "like" button: The impact of mere virtual presence on brand evaluations and purchase intentions in social media settings. *Journal of Marketing*, 76(6):105–120, 2012.
- Neibecker, B. Konsumentenemotionen: Messung durch computergestützte Verfahren. Eine empirsche Validierung nicht-verbaler Methoden. Würzburg u. a.: Physica-Verlag (Konsum und Verhalten, 8), 1985.
- Nigam, K., Lafferty, J., und McCallum, A. Using maximum entropy for text classification, 1999. URL http://www.kamalnigam.com/papers/ma-xent-ijcaiws99.pdf. Abruf am 20.08.2016.
- Novak, P. K., Smailović, J., Sluban, B., und Mozetič, I. Sentiment of emojis. *PloS One*, 10(12):1–22, 2015.
- o. V. Emoji is Britain's fastest growing language as most popular symbol revealed. The Telegraph, 19.05.2015, 2015. URL http://www.telegraph.co.uk/news/newstopics/howaboutthat/11614804/Emoji-is-Britains-fastest-growing-language-as-most-popular-symbol-revealed.html. Abruf am 20.08.2016.
- o. V. FACS Facial Action Coding System, 2016a. URL http://www.cs.cmu.edu/~face/facs.htm. Abruf am 20.08.2016.
- o. V. Digitalisierung. Eine Minute im Internet. Wirtschaftswoche, (19):10–11, 2016b.
- Ortega, R., Fonseca, A., und Montoyo, A. SSA-UO: Unsupervised Twitter sentiment analysis. In Second joint conference on lexical and computational semantics, Vol. 2: Proceedings of the 7th International Workshop on Semantic Evaluation. Atlanta, GA, volume 2, pages 501–507, 2013. URL http://www.aclweb.org/anthology/S13-2083. Abruf am 20.08.2016.
- Ortigosa, A., Martín, J. M., und Carro, R. M. Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31:527–541, 2014.
- Otto, J., Euler, H. A., und Mand, H. Begriffsbestimmungen. Jürgen H. Otto (Hg.): Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psychologie-Verlags-Union, pages 11–18, 2000.
- Oxford Dictionaries. Emoji, 2016. URL http://www.oxforddictionaries.com/de/definition/englisch/emoji. Abruf am 20.08.2016.
- Pak, A. und Paroubek, P. Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of the 7th Conference on International Language Resources and Evaluation. Malta: 1320-1326, 2010. URL http://www.lrecconf.org/proceedings/lrec2010/pdf/385\_Paper.pdf. Abruf am 20.08.2016.
- Pang, B., Lee, L., und Vaithyanathan, S. Thumbs up? Sentiment classification using machine learning techniques. EMNLP '02 Proceedings of the

- ACL-02 conference on Empirical methods in natural language processing. Philadelphia, PA: 79-86., 2002. URL http://www.cs.cornell.edu/home/llee/papers/sentiment.pdf. Abruf am 30.08.2016.
- Pang, B. und Lee, L. Opinion mining and sentiment analysis. *Foundations* and *Trends in Information Retrieval*, 2(1–2):1–135, 2008.
- Parrott, W. G. Emotions in social psychology: Essential readings. Psychology Press, 2001.
- Pauli, P. und Birbaumer, N. Psychophysiologische Ansätze. Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psychologie-Verlags-Union, pages 75–84. 2000.
- Plutchik, R. *Emotion: A psychoevolutionary synthesis*. New York u. a.: Harper & Row, 1980.
- Plutchik, R. Emotions and life: Perspectives from Psychology, Biology, and Evolution. American Psychological Association, Washington, DC, 2003.
- Postinett, A. Starbucks als Paradebeispiel für das Einbinden der Kunden. Handelsblatt, 14.06.2011., 2011. URL http://www.handelsblatt.com/unternehmen/management/soziale-netze-starbucks-als-paradebeispiel-fuer-das-einbinden-der-kunden/4282544.html. Abruf am 20.08.2016.
- R Foundation. The R Project for Statistical Computing, 2016. URL https://www.r-project.org/. Abruf am 20.08.2016.
- Rezabek, L. L. und Cochenour, J. J. Visual cues in computer-mediated communication: Supplementing text with emoticons. *Journal of Visual Literacy*, 18(2):201–215, 1998.
- Rice, R. E. und Love, G. Electronic emotion: Socioemotional content in a computer-mediated communication network. *Communication Research*, 14(1):85–108, 1987.
- Richins, M. L. Measuring emotions in the consumption experience. *Journal of Consumer Research*, 24(2):127–146, 1997.
- Rosenthal, S., Nakov, P., Kiritchenko, S., Mohammad, S., Ritter, A., und Stoyanov, V. SemEval-2015 task 10: Sentiment analysis in Twitter. In *Proceedings of the 9th International Workshop on Semantic Eval-uation.*Denver, CO, pages 451–463, 2015. URL http://www.anthology.acl-web.org/S/S15/S15-2078.pdf. Abruf am 20.08.2016.
- RStudio. Take control of your R code, 2016. URL https://www.rstudio.com/products/rstudio/. Abruf am 20.08.2016.
- Rutledge, L. L. und Hupka, R. B. The facial feedback hypothesis: Methodological concerns and new supporting evidence. *Motivation and Emotion*, 9(3):219–240, 1985.
- Rutter, D. R. Communicating by telephone. Oxford: Pergamon Press, 1987.
  Schept, K. BRANDZ Top 100. Most Valuable Global Brands 2015. Hg.
  v. Millward Brown., 2015. URL https://www.millwardbrown.com/
  BrandZ/2015/Global/2015\_BrandZ\_Top100\_Report.pdf. Abruf am 20.08.2016.
- Scherer, K. R. Theorien und aktuelle Probleme der Emotionspsychologie. K. R. Scherer (Hg.): Psychologie der Emotion. Göttingen u. a.: Verlag für Psychologie Hogrefe (Enzyklopädie der Psychologie, Themenbereich C, Serie IV, 3), pages 1–38, 1990.
- Schiel, S. Entwicklung einer Bilderskala zur Messung markenrelevanter Emotionen. Christian Bosch, Stefan Schiel und Thomas Winder (Hg.): Emotionen im Marketing. Verstehen Messen Nutzen. Nachdruck. Wiesbaden: Deutscher Universtitäts-Verlag (Gabler Edition Wissenschaft: Werbe- und Markenforschung): 161-321, 2007.
- Schmidt-Atzert, L. Emotionspsychologie. Stuttgart u. a.: Kohlhammer, 1981. Schmidt-Atzert, L. Struktur der Emotionen. Jürgen H. Otto (Hg.): Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psychologie-Verlags-Union, pages 30–44, 2000.
- Schönpflug, W. Geschichte der Emotionskonzepte. J. H. Otto (Hg.): Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psycholo-gie-Verlags-Union, pages 19–29, 2000.
- Schweiger, G. Image und Imagetransfer. B. Tietz (Hg.): Handwörterbuch des Marketing. 2. Aufl. Stuttgart: Schäffer-Poeschel (4), pages 915–928, 1995.
- Schweiger, G. und Schrattenecker, G. Werbung: Eine Einführung. 8. Aufl. Konstanz: UVK Verlagsgesellschaft., 2013.
- Shah, N. Introducing emoji targeting, 2016. URL https://blog.twitter.com/2016/introducing-emoji-targeting. Abruf am 20.08.2016.
- Sharma, A. und Dey, S. A comparative study of feature selection and machine learning techniques for sentiment analysis. In *Proceedings of the 2012 ACM research in applied computation symposium*, pages 1–7, San Antonio, TX. 2012.
- Short, J., Williams, E., und Christie, B. The social psychology of telecommu-

- nications. London u. a.: Wiley, 1976.
- Socialbakers Twitter Statistics. Twitter profiles stats. Largest audience, 2016a. URL http://www.socialbakers.com/statistics/twitter/profiles/brands/. Abruf am 20.08.2016.
- Socialbakers Twitter Statistics. Twitter profiles stats. Largest audience retail food, 2016b. URL http://www.socialbakers.com/statistics/twitter/profiles/brands/retail-food/. Abruf am 20.08.2016.
- Statista. Starbucks Statista-Dossier. Hamburg, 2015.
- Statista. Twitter Statista-Dossier. Hamburg, 2016.
- Strack, F., Martin, L. L., und Stepper, S. Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, 54(5):768–777, 1988.
- Stürmer, R. und Schmidt, J. Erfolgreiches Marketing durch Emotionsforschung. Messung, Analyse, Best Practice. Freiburg u. a.: Haufe, 2014.
- SwiftKey. SwiftKey Emoji Report, 2015. URL http://www.aargauer zeitung.ch/asset\_document/i/129067827/download. Abruf am 20.08.2016.
- Szarek, D. ARD-"Markencheck". Die unappetitlichen Geheimnisse von McDonald's. In: Focus, 17.01.2012., 2012. URL http://www.focus.de/finan-zen/news/markencheck-in-der-ard-die-unappetitlichen-geheimnisse-von-mcdonalds\_aid\_703056.html. Abruf am 20.08.2016.
- Thayer, R. E. *The biopsychology of mood and arousal*. Oxford University Press, New York, 1989.
- Thompsen, P. A. und Foulger, D. A. Effects of pictographs and quoting on flaming in electronic mail. *Computers in Human Behavior*, 12(2):225–243, 1996.
- Tomkins, S. S. Affect, imagery, consciousness. Vol. 1: The positive effects. New York: Springer, 1962.
- Traunmüller, H. Paralinguale Phänomene/Paralinguistic phenomena. Ulrich Ammon (Hg.): Sociolinguistics. An international handbook of the science of language and society. 2. Aufl. Berlin u. a.: Gruyter (Handbücher zur Sprachund Kommunikationswissenschaft, 3,1): 653-665, 2004.
- Twitter. Twitter usage: Company facts, 2016. URL https://about.twitter.com/company. Abruf am 20.08.2016.
- Twitter Leitfaden. Was ist ein ReTweet?, 2015a. URL http://www.twitter-leitfaden.de/fragen/was-ist-ein-retweet. Abruf am 20.08.2016.
- Twitter Leitfaden. Was ist ein Tweet?, 2015b. URL http://www.twitter-leitfaden.de/fragen/was-ist-ein-tweet. Abruf am 20.08.2016.
- van Zyl, H. und Meiselman, H. L. The roles of culture and language in designing emotion lists: Comparing the same language in different English and Spanish speaking countries. Food Quality and Preference, 41:201–213, 2015.
- Vidal, L., Ares, G., Machín, L., und Jaeger, S. R. Using Twitter data for food-related consumer research: A case study on "what people say when tweeting about different eating situations". Food Quality and Preference, 45: 58–69. 2015.
- Vidal, L., Ares, G., und Jaeger, S. R. Use of emoticon and emoji in tweets for food-related emotional expression. *Food Quality and Preference*, 49: 119–128, 2016.
- Walther, J. B. und D'Addario, K. P. The impacts of emoticons on message interpretation in computer-mediated communication. Social Science Computer Review, 19(3):324–347, 2001.
- Watson, L. und Spence, M. T. Causes and consequences of emotions on consumer behaviour: A review and integrative cognitive appraisal theory. *European Journal of Marketing*, 41(5/6):487–511, 2007.
- Web Analytics Tools. Stimmungsanalyse, 2014. URL https://web-analytics-tools.com/glossar/155-stimmungsanalyse.html. Abruf am 20.08.2016.
- Weinberg, P. Nonverbale Marktkommunikation. Heidelberg: Physica-Verlag (Konsum und Verhalten, 11), 1986.
- Weinberg, T. Social Media Marketing. Strategien für Twitter, Facebook & co. Unter Mitarbeit von Wibke Ladwig und Corina Pahrmann. 4. Aufl. Peking u. a.: O'Reilly, 2014.
- Wikipedia. Starbucks, 2016. URL https://de.wikipedia.org/wiki/Starbucks#/media/File:Starbucks-Logo\_2011.svg. Abruf am 20.08.2016.
- Wildner, R. und Jäncke, L. Validierung von Messinstrumenten für die Markenstärke mit bildgebenden Verfahren. In: Manfred Bruhn (Hg.): Wie Marken wirken. Impulse aus der Neuroökonomie für die Markenführung.

- München: Vahlen: 93-107, 2010.
- Winder, T. Emotionen im Marketingkontext. In: Christian Bosch, Stefan Schiel und Thomas Winder (Hg.): Emotionen im Marketing. Verstehen Messen Nutzen. Nachdruck. Wiesbaden: Deutscher Universtitäts-Verlag (Gabler Edition Wissenschaft: Werbe- und Markenforschung): 1-157, 2007.
- Zdrzalek, L. Ich liebe es (nicht mehr). In: Zeit Online, 17.03.2015., 2015. URL http://www.zeit.de/wirtschaft/unternehmen/2015-03/mc-donalds-deutschland-krise/komplettansicht. Abruf am 20.08.2016.
- Zentner, M. R. und Scherer, K. R. Partikuläre und integrative Ansätze. J. H. Otto (Hg.): Emotionspsychologie. Ein Handbuch. Weinheim: Beltz, Psychologie Verlags Union, pages 151–164, 2000.
- Zhang, H., Gan, W., und Jiang, B. Machine learning and lexicon based methods for sentiment classification: A survey. 11th Web Information System and Application Conference. Tianjin, China: 262-265., 2014. URL http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7058024. Abruf am 20.08.2016.



# **Junior Management Science**

journal homepage: www.jums.academy



# Measuring Corporate Tax Avoidance - An Analysis of Different Measures

Markus Sebastian Gebhart

Ludwig-Maximilians-Universität München

### Abstract

This study (1) gives an overview of and analyzes the different existing measures of tax avoidance in the empirical tax research literature and (2) aims at answering the question of how closely related they are in measuring corporate tax avoidance. Starting with a theoretical comparison of different Effective Tax Rate based measures, Book-Tax-Differences based measures, the measure developed by Henry and Sansing (2014), Tax Shelter Scores, and Unrecognized Tax Benefits an empirical comparison of the first three mentioned measures is conducted. The purely descriptive analysis reveals that although there are differences between the single measures and those differences persist over time, especially annual proxies exhibit considerable correlation, which increases with similarity in computation and inputs used. A comparison of annual with long-run measures implies that the reliability of annual measures in depicting long-run tax avoidance is not sufficiently high.

Keywords: tax avoidance, tax sheltering, Effective Tax Rate, Book-Tax-Difference

## 1. Introduction

Corporate tax avoidance has been and still is an important issue discussed in public media and politics. The U.S. Treasury Department Secretary, Lawrence H. Summer, characterized tax avoidance in 2000 as "what may be the most serious compliance issue threatening the American tax system today" (U.S. Department of the Treasury (2000)). Nowadays, sixteen years later, data leak scandals such as the "Luxembourg Leaks" in 2014 (see The Guardian (2015); Houlder (2014)) or most recently the "Panama Papers" in 2016 (see Lipton and Creswell (2016)) still reveal massive tax avoidance by corporations as well as individuals. Every year, the loss in tax revenue in the EU due to corporate tax avoidance is estimated from 50 to 70 billion Euros (Dover et al. (2015), p. 16). The U.S. Treasury estimates the loss in tax revenue due to tax sheltering (an aggressive form of tax avoidance) to 10 billion dollars annually (U.S. Department of the Treasury (1999), p. 31). To tackle the problem of corporate tax avoidance there have been actions made by supranational organizations such as the Base Erosion and Profit Shifting Project (BEPS) by the OECD/G20 or the Anti Tax Avoidance Package by the EU, which try to combat corporate tax avoidance on a global scale.1

Besides the social and political interest in, and the development of, actions against corporate tax avoidance, there has also been increasing theoretical and empirical research concerning corporate tax avoidance. Prevalent interests of research in this context are the reasons, drivers, determinants, and implications of tax avoidance for firms and their shareholders. One of the main challenges these empirical studies face is the measurement of corporate tax avoidance since firms have strong incentives not to disclose their tax avoidance activities (Desai and Dharmapala (2009a), p. 539; Wilson (2009), p. 970). Therefore, over the decades, several different proxies for tax avoidance have been created, all of which are argued to have different properties and face different limitations.

Facing those numerous proxies, uncertainty may arise about which measure to choose. To give guidance for researchers in picking the right measure(s) for their research, this bachelor thesis aims to (1) give an overview of and analyze the different measures of tax avoidance existing in the literature and (2) answer the question of how closely related they are in measuring corporate tax avoidance. Because most international tax research refers to the U.S., the focus in this

<sup>&</sup>lt;sup>1</sup>For information about the BEPS project see the webpages of the Organisation for Economic Co-operation and Development (OECD) (2016). For

information on the Anti Tax Avoidance Package see the webpages of the European Comission (2016).

<sup>&</sup>lt;sup>2</sup>For an overview of the empirical tax research see Shackelford and Shevlin (2001) and Hanlon and Heitzman (2010).

thesis is therefore on the U.S. GAAP accounting environment.

The structure proceeds as follows: Section 2 and 3 define the term "tax avoidance" and give guidance on the choice of data sources used as inputs for the measures. In Section 4 the different measures are introduced and their properties and limitations are discussed theoretically. The focus lies thereby on Effective Tax Rate based measures, a measure of Henry and Sansing (2014), Book-Tax-Difference measures, Tax Shelter Scores, and the idea of Unrecognized Tax Benefits. Next, Section 5 analyzes and compares the different measures empirically. The analysis is purely descriptive in nature and compares annual measures among one another and furthermore examines their relationship with long-run measures. Section 6 concludes. Overall, it is found that, although the different measures of corporate tax avoidance do exhibit differences and those differences persist over time, measures based on the same rationale (such as Effective Tax Rate measures) are subject to large correlation among one another. The comparison of annual and long-run measures shows that annual measures may not be sufficient predictors of long-run tax avoidance.

# 2. Corporate Tax Avoidance

Before making attempts to measure corporate tax avoidance one needs to define the term first. According to many textbooks tax avoidance refers to the minimization of tax liability within the framework of the law, in contrast to tax evasion, which refers to minimizing tax liabilities including fraud (Miller and Oats (2014), pp. 15–16; Dover et al. (2015), p. 18).

Unfortunately, there is neither such a precise nor one single prevalent definition of tax avoidance in the field of tax research. This means that several different definitions of the term are used, depending on the research question and the means to measure avoidance. Hanlon and Heitzman, for instance, define "tax avoidance broadly as the reduction of explicit taxes" (2010, p. 137). According to Chen et al. (2010) tax avoidance or tax aggressiveness includes "tax planning activities that are legal, or that may fall into the gray area, as well as activities that are illegal. Thus, tax aggressive activities do not necessarily indicate that the firm has done anything improper" (2010, pp. 41-42). And Tang and Firth define "tax management, tax planning, tax avoidance, or tax shelter as taxpayers exploiting uncertainty in tax law to choose an advantageous method in tax reporting that influences their tax liabilities" (2011, p. 176).

As one can see, tax avoidance cannot be defined as precisely for tax research purposes as some of the textbooks suggest. In this context Hanlon and Heitzman (Hanlon and Heitzman (2010), p. 137) speak of a broad continuum of actions that can be classified as tax avoidance, which ranges from tax favored investments on one end to evasion and sheltering on the more aggressive end. This bachelor thesis follows the definition of Hanlon and Heitzman (2010); defining tax avoidance very broadly reflects that the following measures do not differentiate between legal or illegal tax avoid-

ance (not even intentionally or unintentionally) and thus measure a whole range of activities that reduce the tax burden as avoidance. In the following the terms "tax avoidance", "tax aggressiveness", "tax management" and "tax planning" are all used simultaneously and "tax sheltering" is defined as a particularly aggressive form of tax avoidance.

# 3. Data Sources for Measuring Corporate Tax Avoidance

For measuring corporate tax avoidance one needs information about taxable income and tax liabilities (Salihu et al. (2013), p. 415) as well as financial statement data. The latter one is readily available (at least for publicly traded firms), while the two former ones are less available. Potential information sources for the required inputs are either financial statements of the company or its tax returns. While tax returns should contain the most accurate inputs of taxable income and tax liabilities, access to such information is limited since tax returns are confidential and thus can only be used by a few people (Salihu et al. (2013), p. 415). Alternatively, taxable income and tax liabilities can be estimated from financial statements by grossing-up the current tax expense (as developed by Manzon Jr and Plesko (2002), p. 188). To estimate taxable income, current tax expense is therefore divided by the statutory tax rate<sup>3</sup>:

$$taxable\ income = \frac{current\ tax\ expense}{statutory\ tax\ rate} \tag{1}$$

(Manzon Jr and Plesko (2002), p. 188)

However, this could lead to a poor estimate of the actual taxable income. As Hanlon's answer to the title of her paper "What Can We Infer about a Firm's Taxable Income from Its Financial Statements?" "The short answer: Usually not much" (Hanlon (2003), p. 831) suggests, there are several problems with estimating taxable income from financial statement data. She identifies three main reasons, why simply grossing-up current tax expense may lead to wrong estimates: (1) Current tax expense is likely to be over- or understated in comparison to the actual tax liability.<sup>4</sup> (2) Grossing-up using the top statutory tax rate will lead to wrong estimates of worldwide taxable income for multinational corporations doing business in several countries and therefore facing different tax rates. (3) In case of tax losses, current tax expense will usually be truncated at zero (Hanlon (2003)).

<sup>&</sup>lt;sup>3</sup>Note that the statutory tax rate is assumed to be the top federal corporate income tax rate for the U.S. of 35 percent (see KPMG (2016); PWC (2016)), which is widely used in the tax research literature (Hanlon (2003), p. 844).

<sup>&</sup>lt;sup>4</sup>According to Hanlon, the main reasons for an overstated current tax expense are the differential treatment of nonqualified employee stock options in financial accounting and tax accounting and the "tax cushion" which is incorporated into the current tax expense. Further, current tax expense could be understated because it only captures taxes on continuing operations but not on discontinued operations and extraordinary expenses (benefits), and thus does not reflect the tax on all types of earnings (Hanlon (2003), pp. 838–843). For the effect of employee stock options on the current tax expense see also Hanlon and Shevlin (2002).

However, since tax returns can only be accessed by a few people, and additional disclosures would be necessary to develop a better measure of taxable income, estimating it by grossing-up is a widely used practice (Hanlon (2003), p. 832, 843). Nonetheless one needs to keep these limitations in mind when using such estimates as inputs in measures of corporate tax avoidance. Especially measures based on Book-Tax-Differences may suffer from incorrect estimates as described below.

### 4. Measures of Corporate Tax Avoidance

In the tax research literature several different measures of tax avoidance have been developed and used. These measures differ with respect to the underlying rationale of measuring tax avoidance.

The purpose of this section is to give an overview of the proposed measures in the literature as well as analyze them in terms of validity and performance as measures of corporate tax avoidance.

### 4.1. Effective Tax Rate Based Measures

The first approach relies on the Effective Tax Rate (ETR). The Effective Tax Rate is basically the average tax rate a corporation pays on its pre-tax profits and is calculated by dividing a measure of tax liability by a measure of pre-tax income (Hanlon and Heitzman (2010), p. 139; Gupta and Newberry (1997), p. 1).

$$ETR = \frac{measure\ of\ tax\ liability}{measure\ of\ pre-tax\ income} \tag{2}$$

ETR based measures can be compared with the statutory tax rate. If an ETR measure is below the statutory tax rate, this could signal tax avoidance.

The ETR can be calculated based on different measures of tax liability (total tax expense, current tax expense, cash tax expense)<sup>5</sup> and pre-tax income, and can vary in terms of periods included in the measure. Broadly, one can differentiate between annual ETR measures and long-run ETR measures.

# 4.1.1. Annual ETR Measures

GAAP ETR

The basic form of annual ETR measures is the annual GAAP Effective Tax Rate (GAAP ETR), which is disclosed by firms in their financial statements (Dyreng et al. (2008), p. 65). It is defined as:

$$GAAP \ ETR = \frac{total \ tax \ expense}{pre - tax \ income}$$
 (3)

(Dyreng et al. (2008), p. 65; Chen et al. (2010), p. 46; Dyreng et al. (2010), pp. 1168–1170)<sup>6</sup>

This basic form of ETR suffers from several severe limitations. Dyreng et al. (Dyreng et al. (2008), p. 65) criticize the GAAP ETR for not measuring tax deferral strategies (so do Hanlon and Heitzman (2010), p. 139 and Salihu et al. (2013), p. 416). They argue since total tax expense incorporates both current and deferred tax expense, but tax deferral strategies will reduce current tax expense and simultaneously increase deferred tax expense, these tax planning activities will not have any influence on GAAP ETR. Furthermore, the reliance on pre-tax income in the denominator limits GAAP ETR to non-conforming tax avoidance<sup>7</sup> (Badertscher et al. (2015), p. 6; Hanlon and Heitzman (2010), pp. 139-140; Salihu et al. (2013), p. 416). Thus, GAAP ETR would, for instance, not capture the tax effects of interest deductibility, because this reduces both taxable and financial income (Hanlon and Heitzman (2010), p. 141). Due to these limitations it becomes clear, that the GAAP ETR does not measure a considerable portion of tax avoidance.

Moreover, GAAP ETR does not distinguish between reductions in tax liabilities due to actual tax planning strategies and other "accidental" reductions (e.g. changes in valuation allowances) not associated with tax considerations. In fact every reduction of the explicit tax liability will alter GAAP ETR (Hanlon and Heitzman (2010), pp. 139–141).

Furthermore, the high volatility in GAAP ETR due to the use of annual data (Dyreng et al. (2008), p. 65) makes long-term predictions difficult. There is also a significant truncation bias of GAAP ETR rates (Henry and Sansing (2014)). This truncation bias results when pre-tax income is negative, in which case researchers usually drop loss years (Henry and Sansing (2014), p. 2; Wilkie and Limberg (1993), p. 49) since the interpretation of negative ETR is difficult. According to Henry and Sansing (Henry and Sansing (2014), p. 3) this practice also leads to an asymmetrical treatment of income and loss years, which may distort the results. Because of these shortcomings, there have been attempts to modify the GAAP ETR in order to increase the power of Effective Tax Rate measures.

Current ETR

To overcome the limitation of the GAAP ETR only measuring permanent tax avoidance, one could use current tax expense in the numerator instead which yields the Current Effective Tax Rate (Current ETR) measure (Salihu et al. (2013), p. 416):

$$Current \ ETR = \frac{current \ tax \ expense}{pre - tax \ income} \tag{4}$$

<sup>&</sup>lt;sup>5</sup> To guarantee consistency with the empirical part of this thesis the definitions for these items are in line with those provided by the Compustat database. Therefore, total tax expense includes all income taxes imposed on the firm, current tax expense represents the current amount of taxes payable by the firm and cash taxes paid is the sum of all income taxes actually paid by the firm (Wharton Research Data Services (2016b)).

 $<sup>^6</sup>$ The annual GAAP ETR is used, for instance, in studies by Chen et al. (2010), Dyreng et al. (2010), Armstrong et al. (2012), or .

<sup>&</sup>lt;sup>7</sup>Non-conforming tax avoidance refers to tax planning strategies that result in a decrease in taxable income, while leaving financial income unaffected. Conforming tax avoidance reduces both taxable and financial income (Badertscher et al. (2015), p. 1).

(Salihu et al. (2013), p. 416)<sup>8</sup>

This variation allows measurement of tax deferral strategies (Salihu et al. (2013), pp. 416–417), because a reduction in current tax expense will not get compensated by an increase in the deferred tax expense, as this was the case using total tax expense (as described above). However, besides the enhancement in measuring tax deferral strategies, the Current ETR measure suffers from the same problems as the GAAP ETR. Additionally, as already described in Section 3 current tax expense may be over- or understated in comparison to the actual tax expense (Hanlon (2003), p. 831), which may distort the Current ETR. *Cash ETR* 

A further variation of annual ETR is the Cash Effective Tax Rate (Cash ETR), defined as:

$$Cash \ ETR = \frac{cash \ taxes \ paid}{pre - tax \ income} \tag{5}$$

(Chen et al. (2010), p. 46; Dyreng et al. (2010), p. 1169)<sup>9</sup>

Using cash taxes paid in the numerator instead of total or current tax expense makes the measure robust to changes in estimates as, for example, in the valuation allowances (Dyreng et al. (2008), p. 66). It also takes into account the tax effects of employee stock options (Dyreng et al. (2008), p. 66; Chen et al. (2010), p. 48), one main factor identified by Hanlon (2003) resulting in an overstatement of current tax expense, and thus is not subject to overstatement in this matter

On the other hand, cash taxes paid could also include tax payments of former periods as it includes all taxes paid in one year regardless of which periods they actually arose in (e.g. tax payment related to an IRS audit of former years), which could lead to a mismatch of numerator and denominator and thus distort the Cash ETR (Hanlon and Heitzman (2010), p. 139; Dyreng et al. (2008), pp. 66–67). As in the case of Current ETR, Cash ETR also suffers from the remaining problems of GAAP ETR, not fixed by using cash taxes paid.

Cash flow based ETRs Another though less frequently used alternative to GAAP ETR (using pre-tax income) is the application of operating cash flows in the denominator. This yields measures such as total tax expense over operating cash flow (further called CFM1) (as used in Zimmerman (1983); Salihu et al. (2013)) or cash taxes paid over operating cash flow (further called CFM2) (as used in Salihu et al. (2013))

$$CFM_1 = \frac{total\ tax\ expense}{operating\ cash\ flow} \tag{6}$$

(Zimmerman (1983), pp. 122–123; Salihu et al. (2013), p. 417)

$$CFM_2 = \frac{cash\ taxes\ paid}{operating\ cash\ flow} \tag{7}$$

(Salihu et al. (2013), p. 418)

Although  $CFM_1$  is not impacted by accrual accounting in the denominator, it still is subject to accrual effects in the numerator by using total tax expense and thus only measures non-conforming tax avoidance (Salihu et al. (2013), p. 417).  $CFM_2$  solves this problem by using cash taxes paid. Therefore, it is argued that cash taxes paid over operating cash flow is able to measure conforming tax avoidance (Salihu et al. (2013), p. 418). The remaining problems associated with the total/cash tax expense are also applicable to the  $CFM_1$  and  $CFM_2$ .

# 4.1.2. Long-Run ETR Measure

To overcome some of the previously discussed limitations of the annual ETR measures, Dyreng et al. (2008) develop a measure called Long-run Cash Effective Tax Rate (Long-Run Cash ETR / lr. Cash ETR). They define their measure as the sum of cash taxes paid over a long period of time (e.g. ten years) divided by the sum of pre-tax income (excluding special items) over the same time period:

$$Long - Run \ Cash \ ETR = \frac{\sum_{t=1}^{N} \ cash \ taxes \ paid_{t}}{\sum_{t=1}^{N} \ (pre - tax \ income_{t} - special \ items_{t})}$$
(8)

(Dyreng et al. (2008), p. 67)<sup>10</sup>

This enables them to (1) use and benefit from cash taxes paid in the numerator, without suffering from a potential mismatch of periods<sup>11</sup> (Dyreng et al. (2008), pp. 66–67), (2) measure corporate tax avoidance over a long period of time, reducing volatility present in annual ETR measures (Hanlon and Heitzman (2010), p. 141; Salihu et al. (2013), p. 417), and (3) reduce (though not eliminate) data truncation bias due to loss years (Henry and Sansing (2014), p. 4).

Although Long-Run Cash ETR solves many problems associated with annual ETR measures, it still suffers certain limitations: It is still limited to non-conforming tax avoidance, measures all influences on explicit tax liabilities and does not differentiate between aggressive and non-aggressive tax avoidance (Hanlon and Heitzman (2010), p. 141). Although reduced, there is still the risk of a truncation bias due to the elimination of loss firms (Henry and Sansing (2014), p. 4).

# 4.2. Henry and Sansing's HS Measure

Henry and Sansing (Henry and Sansing (2014), pp. 9–14) develop a measure (further referred to as HS) based

<sup>&</sup>lt;sup>8</sup>This measure is used as a proxy for tax avoidance in Hope et al. (2013), or Lennox et al. (2013).

<sup>&</sup>lt;sup>9</sup>The Cash ETR is, for instance, applied in studies from Chen et al. (2010), Dyreng et al. (2010) and Lennox et al. (2013).

<sup>&</sup>lt;sup>10</sup>Long-Run Cash ETR has been widely used in tax research literature as a proxy for corporate tax avoidance as in Chen et al. (2010), Dyreng et al. (2010), Armstrong et al. (2012), Hope et al. (2013), Lennox et al. (2013), and Lisowsky (2010).

<sup>&</sup>lt;sup>11</sup>Summing up cash taxes paid reduces the risk that a large portion of the numerator refers to a time period not captured in the denominator (Dyreng et al. (2008), p. 67).

on the Cash ETR which does not suffer a truncation bias.  $^{12}$  They argue that the omission of loss firms leads to a substantial truncation bias for annual Cash ETR as well as for Long-Run Cash ETR.  $^{13}$ 

To avoid the omission of loss firms, they decompose the Cash ETR components. Cash taxes paid can be viewed as the combination of the expected amount of tax payments based on the pre-tax income times the statutory tax rate  $\tau$  (the hypothetical case of a firm with no book-tax differences facing only one tax rate) and a combination of tax preferences,  $\Delta$ . Those preferences explain the differences between the hypothetical case and the actual tax payments (Henry and Sansing (2014), p. 9):

$$cash\ taxes\ paid = \tau * (pre - tax\ income) + \Delta \qquad (9)$$

This can be solved for  $\Delta$ , yielding the tax preferences of a firm:

$$\Delta = cash \ taxes \ paid - \tau * (pre - tax \ income)$$
 (10)

Henry and Sansing (2014) argue that pre-tax income equals size (Market Value of Assets, MVA) times accounting profitability (Return on Assets, ROA):

$$pre-tax\ income = MVA*ROA$$
 (11)

In their view, using pre-tax income may lead to biased results, since ETR measures therefore are also driven by the pre-tax profitability (see decomposition of pre-tax income in eq. 10) and not solely due to differences in tax preferences. They propose to use Market Value of Assets (MVA) in the denominator to account for this problem, whereby MVA is defined as:

$$MVA = book \ value \ of \ assets + (markt \ value \ of \ equity) - book \ value \ of \ equity) = BVA + (MVE - BVE)$$

$$(12)$$

(Henry and Sansing (2014), p. 10)

They measure corporate tax avoidance as the tax preferences,  $\Delta$ , scaled by size (MVA):

$$HS = \frac{\Delta}{MVA} = \frac{cash\ taxes\ paid - \tau * (pre - tax\ income)}{MVA}$$
(13)

A firm without tax preferences ( $\Delta$ = 0, i.e. cash taxes paid = expected tax payment) will have a Cash ETR equal to the statutory rate and an HS of zero. If the tax preferences are positive ( $\Delta$ > 0, i.e. cash taxes paid > the expected tax payment), Cash ETR will be higher than the statutory rate and the HS will be positive. In the case of negative tax preferences ( $\Delta$ < 0, i.e. cash taxes paid < statutory tax rate), Cash ETR will be lower than the statutory tax rate and HS will be negative.

The HS measure solves two problems associated with ETR measures: First, it eliminates the risk of data truncation bias, because MVA will always be positive and no observations need to be dropped (Henry and Sansing (2014), p. 11). Furthermore, as already mentioned, HS is not driven by the pretax profitability of the firm (Henry and Sansing (2014), p. 10), which may distort ETR measures. Additionally, using MVA in the denominator should enable measuring conforming and non-conforming tax avoidance (Badertscher et al. (2015), p. 3). It is also possible to measure long-run tax avoidance using the HS measure (Henry and Sansing (2014), pp. 15–16). On the other hand, the HS measure lacks the intuitive interpretation of ETR measures, which makes interpretation more complicated (Bonsall et al. (2015, p. 13)).

### 4.3. Book-Tax-Differences Based Measures

The third approach uses differences between the financial income and taxable income, or Book-Tax-Differences (BTD). The BTD is usually defined as the difference between pretax income according to the financial statement (also called "book income") and the taxable income according to the tax return (Manzon Jr and Plesko (2002), p. 192; Guenther (2014), p. 7). Because of the confidentiality of tax return data, taxable income is estimated based on current tax expense as described above. This yields the basic formula for the Total BTD developed by Manzon Jr and Plesko (2002):

$$Total\ BTD = pre-tax\ income - \frac{current\ tax\ expense}{statutory\ tax\ rate}$$
(14)

(Manzon Jr and Plesko (2002), p. 192; Guenther (2014), p. 7)<sup>14</sup>

Prior research suggests that BTD are partly attributable to corporate tax avoidance. Desai (2003) argues that the increase in BTD during the 1990s is to some extent due to an increase in tax sheltering activity. Wilson (2009) also finds a significant effect of tax sheltering on BTD and so do Tang and Firth (2011). On the other hand, BTD are also associated with earnings management (see Tang and Firth (2011) and Hanlon (2005)). Earnings management could influence the

<sup>&</sup>lt;sup>12</sup>This relatively young measure finds application as a proxy of tax avoidance in papers by Bonsall et al. (2015), Cen et al. (2016) and Koester et al. (2016)

<sup>&</sup>lt;sup>13</sup>For instance, Dyreng et al. (2008) drop 18 percent of the one year and 15 percent of the ten year observations in their study of "Long-Run Corporate Tax Avoidance".

<sup>&</sup>lt;sup>14</sup>The Total BTD is used in papers by Wilson (2009), Chen et al. (2010), Lisowsky (2010), Lennox et al. (2013), and Desai and Dharmapala (2009a).

<sup>&</sup>lt;sup>15</sup>Earnings management refers to actions managers take using their own judgement in financial reporting to influence financial outcomes (Healy and Wahlen (1999), p. 368) and thus may have an influence on BTD.

BTD in cases when managers try to inflate financial earnings, which would all else equal lead to an increase in the BTD. Therefore, measures based on Total BTD have been developed to account for the influence of earnings management.

# 4.3.1. Discretionary Total BTD

Based on the basic formula of the Total BTD developed by Manzon Jr and Plesko (2002), Desai and Dharmapala (Desai (2003), pp. 158–160) create a measure of corporate tax avoidance called Discretionary Total BTD (Lee et al. (2015), p. 27) or Abnormal BTD (Hanlon and Heitzman (2010), p. 141). To avoid BTD being driven by earnings management, they regress the Total BTD (eq. 13) scaled by lagged total assets<sup>16</sup> on total accruals (scaled by lagged total assets) which they argue reflect earnings management attempts:

$$\frac{BTD_t}{lagged\ total\ assets_t} = \beta_0 + \beta_1 * \frac{TA_t}{lagged\ total\ assets_t} + \epsilon_t$$
(15)

(Desai and Dharmapala (2006), p. 159)<sup>17</sup>

The residual  $\epsilon_t + \beta_0$  is interpreted as the portion of BTD not associated with earnings management and, thus, can be used as a proxy for tax avoidance in year  $\theta$ :

Discretionary Total 
$$BTD_t = \epsilon_t + \beta_0$$
 (16)

# (Desai and Dharmapala (2006), p. 159)

Although Discretionary Total BTD controls for earnings management it might be biased due to the estimation of taxable income as discussed in Section 3. However, Desai and Dharmapala argue that at least the problem associated with nonqualified employee stock options will not impact their measure (2006, pp. 158-159). Because the measure consists of residuals, it sums up to zero over all firms and years and therefore can neither be interpreted as a dollar amount nor can it be used to calculate economy wide tax avoidance over all firms (Desai and Dharmapala (2006), p. 160). This also aggravates the interpretation of the Discretionary Total BTD, since it cannot be interpreted as easily as the ETR based measures. Per definition, Discretionary Total BTD is limited to non-conforming tax avoidance (Hanlon and Heitzman (2010), p. 141). Additionally, the Discretionary Total BTD does also not distinguish between actions that intentionally create Book-Tax-Differences and actions that accidentally create such differences (except for earnings management) (Frank et al. (2009), p. 472).

### 4.3.2. Discretionary Permanent BTD

To address the problem of Book-Tax-Differences not associated with tax planning, Frank et al. (Frank et al. (2009), pp. 472–474) develop their own measure, called Discretionary Permanent BTD. The measure only captures permanent Book-Tax-Differences following the argument that ideal tax shelter activities result in permanent differences between financial and taxable income and temporary BTD are also influenced by earnings management. They use Total Permanent BTD, defined as:

Total Permanent BTD =

Total BTD - 
$$\frac{\text{deferred tax expense}}{\text{statutory tax rate}}$$
 (17)

(Frank et al. (2009), p. 473)

The Total Permanent BTD can also be computed using an ETR differential, i.e. the difference between statutory tax rate and GAAP ETR multiplied by pre-tax income (Lee et al. (2015), p. 27).

This Total Permanent BTD is regressed on a set of control variables, which are known to create permanent Book-Tax-Differences but are not associated with tax planning, in order to make their measure robust against "accidental" BTD. Their measure is as follows:

Total Permanent 
$$BTD_t = \beta X + \epsilon_t$$
 (18)

(Frank et al. (2009), p. 473)

Where X is a  $m \times n$  matrix of control variables (with m observations for each firm and n control variables)<sup>18</sup>, and  $\beta$  is the corresponding  $1 \times n$  vector of coefficients. The error term  $\epsilon_t$  is the Discretionary Permanent BTD (or DTAX); i.e. the portion of Total Permanent BTD that is driven by tax planning activities:

Discretionary Permanent 
$$BTD_t = DTAX_t = \epsilon_t$$
 (19)

(Frank et al. (2009), p. 473)<sup>19</sup>

By construction, the Discretionary Permanent BTD excludes Temporary BTD, which has led to criticism by Hanlon and Heitzman (Hanlon and Heitzman (2010), p. 142). They argue that tax avoidance creates Permanent and Temporary BTD and thus the measure developed by Frank et al. (2009) may fail to measure a significant proportion of tax avoidance (such as tax deferral strategies). On the other hand, Frank et al. (Frank et al. (2009), p. 472) reason that the benefits

<sup>&</sup>lt;sup>16</sup> The BTD is scaled by lagged total assets in order to control for firm scale (Desai and Dharmapala (2006), p. 158; Desai and Dharmapala (2009a), p. 540)), because the same BTD might be high for a relatively small firm and small for a relatively large firm (Guenther (2014), p. 7). This scaling makes BTD comparable between firms. They use the Ordinary Least Squares Method (OLS) for their regression.

<sup>&</sup>lt;sup>17</sup>This measure is used for instance in studies from Desai and Dharmapala (2009b), and Chen et al. (2010).

<sup>&</sup>lt;sup>18</sup> The control variables are: goodwill and other intangibles, income (loss) reported under the equity method, income (loss) attributable to minority interest, current state income tax expense, change in net operating loss carry forwards, and one-year lagged Total Permanent BTD (Frank et al. (2009), p. 473)
<sup>19</sup> The Discretionary Permanent BTD is, for instance, used in studies from,

<sup>&</sup>lt;sup>19</sup> The Discretionary Permanent BTD is, for instance, used in studies from Lisowsky (2010), Tang and Firth (2011), and Lennox et al. (2013).

of excluding temporary differences is worth the costs of not measuring temporary tax avoidance. In their view, controlling for earnings management justifies this limitation. However, one could argue that it depends on the research question and targets to decide about the usefulness of the Discretionary Permanent BTD. Similarly to the Total BTD and Discretionary Total BTD, the Discretionary Permanent BTD only captures non-conforming tax avoidance (Hanlon and Heitzman (2010), p. 142).

### 4.4. Tax Shelter Scores

Besides ETR and BTD based measures, there have been attempts to measure the likelihood of tax shelter activities (considered to be one of the most aggressive form of tax avoidance; Lisowsky et al. (2013), pp. 590–591; Hanlon and Heitzman (2010), p. 137) based on certain firm characteristics.

Based on a sample of firms being accused of tax sheltering, Wilson (Wilson (2009), pp. 979–984) uses a logistic regression model to regress the probability that a firm is identified as being engaged in tax sheltering (defined as  $ln(\frac{P_{Shelter}}{1-P_{Shelter}})$ ) on certain firm

$$ln(\frac{P_{Shelter}}{1 - P_{Shelter}}) = \beta X + \epsilon_t \tag{20}$$

(Wilson (2009), p. 982)

Where X is the  $m \times n$  matrix of firm characteristics (with m observations and n control variables)<sup>20</sup> and  $\beta$  the corresponding  $1 \times n$  vector of coefficients. He finds that BTD (Total, Permanent, and Temporary), leverage, and discretionary accruals (measure of earnings management) are significantly associated with tax sheltering. For instance, a one percent increase in Total BTD results in a 2.78 percent increase in the probability of tax sheltering and a one percent increase in discretionary accruals results in a 3.45 percent increase in the probability of tax sheltering (Wilson (2009), p. 986).

Lisowsky (2010) extends the framework of Wilson (2009) using a larger sample of firms and more firm characteristics.<sup>21</sup> Overall, his model has a 88 percent chance of identifying a tax shelter firm as one with a higher predicted probability of tax shelter likelihood (Lisowsky (2010), p. 1718).

Using the results of Wilson (2009) or Lisowsky (2010) one could estimate the likelihood of tax sheltering for the firm of interest and thus create a proxy for tax avoidance. While this approach may be a useful approximation for tax avoidance behavior, the results should be used with caution.

Because the sample of firms used only contains firms being accused of tax sheltering it may not be possible to generalize those results for the whole population of firms. Corporations usually use many different practices to reduce their tax burden and only firms that either have already exploited all other techniques or are not able to exploit other techniques would rely on tax shelters. Thus the sample is likely to reflect only extreme cases of tax avoidance (Lee et al. (2015), pp. 29–30; Hanlon and Heitzman (2010), pp. 143–144). Furthermore, firms may be able to engage in tax sheltering without getting caught by the tax authorities. These firms, however, are not included in the sample (Lee et al. (2015), p. 30), making the generalization of the results questionable.

# 4.5. Unrecognized Tax Benefits

Another way to get a proxy for corporate tax avoidance is the usage of Unrecognized Tax Benefits reported in financial statements.<sup>22</sup> For fiscal years after December 15, 2006 all publicly traded U.S. firms are required to disclose so-called "Unrecognized Tax Benefits" in their financial statements according to Financial Accounting Standards Board Interpretation No. 48 (FIN 48) (Financial Accounting Standards Board (FASB) (2006), p. 7). Under this FIN 48 environment, there is a two-step process of evaluating uncertain tax positions (see Financial Accounting Standards Board (FASB) (2006), pp. 1–7). First, a firm recognizes all tax positions for which it is more likely than not (i.e. with a probability > 50 percent) that the tax authorities will challenge those positions. Second, the identified tax positions should be valued at "the largest amount of tax benefit that is greater than 50 percent likely of being realized upon ultimate settlement with a taxing authority that has full knowledge of all relevant information" (Financial Accounting Standards Board (FASB) (2006), p. 3). As a result of this two- step process, there may be differences between the benefit recognized in the financial statement and the tax return. These differences are labeled Unrecognized Tax Benefits (UTB) and they are reflected by a contingent liability in the financial statements.

The rationale behind the argument that these UTB reflect corporate tax avoidance is that firms which are taking a more tax aggressive approach will have higher uncertainty in their tax positions and thus higher UTB compared to firms taking a more conservative approach (Lisowsky et al. (2013), p. 590). Therefore one can infer the degree of tax avoidance a firm is involved in from the UTB. To make the UTB comparable across firms it can be scaled by total assets (Lisowsky et al. (2013), p. 599).

Lisowsky et al. (2013) find a significant, positive relationship between tax shelter use and UTB. They also show that the results of Lisowsky (2010) (see Section 4.4) improve in quality when UTB is included. Furthermore, they find that the relationship between UTB and tax avoidance is not driven by earnings management, which could present a

<sup>&</sup>lt;sup>20</sup>The firm characteristics used by Wilson (Wilson (2009), pp. 982–986) are Total BTD, Temporary BTD (defined as deferred tax expense grossed up by statutory tax rate), Permanent BTD (defined as Total BTD minus Temporary BTD), discretionary accruals, log of total assets, leverage ratio, foreign income, research and development scaled by total assets, and return on assets. These characteristics are associated with increased opportunities for

<sup>&</sup>lt;sup>21</sup>The variables involve those used in Wilson (2009) and add several new variables. For descriptions of those variables, see Lisowsky (Lisowsky (2010), p. 1704).

<sup>&</sup>lt;sup>22</sup>Unrecognized tax benefits are discussed in studies from Lisowsky et al. (2013), Waegenaere et al. (2015), and Lee et al. (2015).

severe shortcoming of UTB as argued in Hanlon & Heitzman (Hanlon and Heitzman (2010), p. 143).

# 5. Empirical Analysis of Corporate Tax Avoidance Measures

The previously discussed measures of corporate tax avoidance are based on similar inputs. The variations of ETR and BTD measures especially are quite similar in computation. This leads to the question: How closely related are these measures among one another in measuring tax avoidance?

This question has already been discussed by other researchers. Lisowsky et al. (Lisowsky et al. (2013), pp. 589–592) argue that different measures capture different degrees of the "continuum" tax avoidance (see Section 2). They place measures such as GAAP ETR and Cash ETR on the less aggressive side of the continuum, and Total BTD, Discretionary Total BTD, Discretionary Permanent BTD and UBT on the more aggressive end of the range. Hanlon & Heitzman on the other hand find BTD based measures to be "closely related" (2010, p. 141) to ETR measures. Guenther (2014) goes even so far as to claim that BTD do not contain any different information compared to ETR measures besides adding measurement error. He shows that BTD (and also the HS) measures are basically just transformations of ETR measures.

To get a better understanding of the relationship between the different measures they will be compared empirically. In a first step the distributional characteristics and development of annual measures over time are analyzed (for the time period of 1996-2015) in order to examine if there are differences and similarities. Second, it will be investigated how good annual measures are at predicting long-term tax avoidance (for the time period of 2006-2015).

In the following, the focus will be on ETR, BTD, and the HS measures of tax avoidance because necessary information for computing tax shelter scores are not publicly available. Further, since UTB are only available for fiscal years after December 15, 2006, but the observation period will be from 1996-2015, UTB will not be part of the investigation because for a substantial part of the observations no UTB data will be available.

# 5.1. Sample Selection

The data retrieved covers all the available publicly traded companies from the Compustat Database for the years 1996 to 2015 for the first analysis and 2006-2015 for the comparison of annual and long-run measures. Reducing the observation period for the long-run comparison is necessary to maximize the sample since the fraction of firms having non-missing data for more than ten consecutive years is very small. For the first analysis, however, it is not important to have data over such a long time horizon without missing values. Thus, in order to maximize the sample, the observation period will cover twenty years of data.

The data items retrieved from the Compustat Database are summarized in Appendix 1. Requirements for the data in the first sample (Sample 1) are that there is no missing data for the items "total assets" (# 6 AT)23, "total liabilities" (# 181 LT), "market value of equity" (# - MKVALT), "pre-tax income" (# 170 PI), "operating cash flow" (# 308 OANCF), "total tax expense" (# 16 TXT), "current tax expense" (# - TXC), "deferred tax expense" (# 50 TXDI), "cash taxes paid" (# 317 TXPD), or "net income" (# 172 NI). These restrictions are necessary to ensure that it is technically possible to compute the different measures of tax avoidance. In order to avoid the problem of firms facing several different tax rates which creates problems when estimating taxable income, only firms without foreign income are included in the sample (following Dyreng et al. (2008), p. 67). Further, certain items used for the computation of BTD are set equal to zero if they are missing. These items are "intangible assets - total" (# 33 INTAN), "Income (loss) reported under the equity method" (# 55 ESUB), "minority interest" (# 49 MII), "income taxes state" (# 173 TXS), and "tax loss carry forward" (# 52 TLCF). This is done based on the assumption that missing values in these categories indicate that they were zero for the observation period. Total accruals needed for the Discretionary Total BTD are calculated by subtracting "operating cash flow" from "net income" (# 172 NI) (Penman (2013), p. 123, 125), and are set equal to zero if missing.

For the second analysis (Sample 2), firms are additionally required to have ten years of non-missing data for the years 2006 to 2015. The five-year measures are computed for the time period 2006 - 2010 and 2011 - 2015, the ten-year measures for the time period 2006 - 2015.

The ETR measures are computed according to the formulas above. Because negative ETRs are difficult to interpret they are excluded from the analysis if either the numerator is negative or the denominator is less or equal to zero. Further, ETR measures above one and below zero are excluded to improve interpretation.

The HS measure is calculated according to the formulas above. In order to control for outliers the smallest and biggest percentage of observations is excluded from the analysis.  $^{24}$ 

The BTD measures are regressed with two modifications. The Total BTD is scaled by the one year lagged value of total assets in order to make the measure comparable across firms (following Desai and Dharmapala (2009a), p. 540; Lisowsky (2010), p. 1704). For the computation of the Discretionary Permanent BTD all variables are scaled by the one-year lagged value of total assets in order to achieve the same goal. Values for BTD are computed unless a firm has only data existing for one year, which makes a linear regression

<sup>&</sup>lt;sup>23</sup>The brackets display the Legacy CST Item Number (retrieved from the Center for Research in Security Prices (2016)) and the mnemonic abbreviation of the Compustat Database (retrieved from Wharton Research Data Services (2016a)) of the data items.

<sup>&</sup>lt;sup>24</sup>For the effects outliers might have on the results see Toutenburg, Heumann, and Schomaker (Toutenburg et al. (2009), p. 187); Wooldridge (Wooldridge (2013), pp. 316–324).

**Table 1:** Distributional Characteristics of Annual Measures

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
GAAP ETR	20,003	0.3009	0.1582	0	0.9989
Current ETR	19,793	0.2365	0.1874	0	1
Cash ETR	19,848	0.2124	0.1900	0	1
CFM1	20,400	0.2358	0.2117	0	1
CFM2	22,597	0.1682	0.2021	0	1
HS	36,018	0.0281	0.0677	-0.0630	0.4994
Total BTD	19,795	0.0324	0.0849	-0.1332	0.8587
Discretionary Total BTD	18,079	0.0232	0.0657	-0.1937	0.4227
Discretionary Permanent BTD	19,223	0.0071	0.0439	-0.1358	0.4445

impossible. Loss years are omitted from the sample (i.e. in cases where pre-tax income or current tax expense < 0) since the BTD is likely to be erroneous because the estimation of taxable income is likely to result in meaningless results. To control for outliers the smallest and biggest percentage of observations is excluded from further analysis.

At this point it should be mentioned that these data requirements can lead to biased results, since a substantial fraction of observations is omitted. However, the requirements are necessary in order to be able to compute the measures and interpret them properly. A summary of the impact of the data requirements is given in Appendix 2. Overall, only approximately 50 percent of the observations are usable for ETR and BTD measures in the first sample. In contrast, the HS measure enables the use of 91.40 percent of the observations (in Sample 1). As already described, this is due to the use of MVA as the numerator instead of pre-tax income, which is always positive (see Section 4.2 for a detailed argumentation).

Also note that the high fraction of used observations for Discretionary Total BTD (0.9801) is only due to the fact that this kind of BTD can solely be computed if the Total BTD is available. Therefore a substantial amount of observations is already lost due to technical reasons (i.e. all observations for which Total BTD is not usable; this includes all observations with pre-tax income or current tax expense below zero). Interestingly, the fraction of usable observations is higher for the second sample. On the other hand, increasing the observation period from one to up to ten years (for long-run measures) reduces the possible truncation bias just slightly.

# 5.2. Comparison of Annual Measures

In a first step of the analysis the annual measures are compared to show differences and similarities between them for the observation period and over time.

### 5.2.1. Distributional Characteristics

Table 1 gives an overview of the main summary statistics of the measures.  $^{25}$ 

The data reveals that the GAAP ETR has the highest mean of all ETR based measures (0.3009) followed by Current ETR (0.2365),  $CFM_1$  (0.2358), Cash ETR (0.2124), and  $CFM_2$ (0.1682). GAAP ETR, having the highest mean, is consistent with the theoretical argumentation that it only captures nonconforming tax avoidance and ignores tax deferral strategies. Therefore, it is consistent for Current ETR to be lower since it reflects tax deferral strategies. One could argue that Cash ETR should lie below Current ETR because cash taxes paid is not subject to overstatement to the same extent as current tax expense is. This argument is supported by Cash ETR having a lower mean than Current ETR. Finally  $CFM_2$ , having the lowest mean, is in line with it being the only measure also capturing conforming tax avoidance. The interpretation of  $CFM_1$  is in some way hard, since it should be less but still vulnerable to accrual accounting making an interpretation of the result difficult. The mean of HS is close to zero because of the differing interpretation of the measure. Because of this it does not make sense to compare the mean of HS with ETR or BTD measures. Surprisingly, the positive value of 0.0281 indicates that firms have, on average, positive tax preferences. The mean of all BTD measures is slightly above zero, which implies that on average the BTD is positive and thus the taxable income is below the book income. The results support the view that the Total BTD may be influenced also by other factors than tax planning (i.e. it has the highest mean of all BTD, 0.0324). The Discretionary Total BTD is lower (0.0232) which is in line with excluding earnings management as a driver of large BTD. Finally, the Discretionary Permanent BTD controls for more drivers of BTD not associated with tax planning and thus yields the lowest average of BTD measures (0.0071).

Looking at the standard deviation (Std. Dev.) of ETR measures one can infer that all measures using either cash taxes paid in the numerator and/or operating cash flow in the denominator have a higher variability than those using pre-tax income or total tax expense. This result is not surprising since accrual accounting (evident in pre-tax income and total tax expense) smoothens fluctuation in cash flows over the years (Dechow (1994), p. 19). Cash taxes paid and operating cash flow on the other hand should be subject to a higher volatility. The volatility of BTD decreases from To-

<sup>&</sup>lt;sup>25</sup>Those are the arithmetic mean (Mean), standard deviation (Std. Dev.) and the minimal (Min.) and maximal (Max.) values. Please note that the values in all tables are rounded to four decimal places.

tal to Discretionary Total BTD and is lowest for Discretionary Permanent BTD.

Graphically, the distributions of ETR and HS measures are shown in Figure 1. Notably, the measures are distributed around their means. Further there is a relatively high fraction of ETR observations in the range [0, 0.1]. This is due to the fact that for these observations the numerator often takes very small values. It also becomes evident that higher values of ETR (values in the range (0.5, 1]) are relatively rare. Not surprisingly, the HS measure is centered on zero, which is in line with the interpretation of the measure and the mean close to zero (0.0281). The distribution also reveals that although the mean is positive, a large fraction of HS is slightly negative, suggesting that a substantial fraction of observations is identified as having negative tax preferences. The distributions of BTD measures in Figure 2 reveal that those measures are centered on zero. Total BTD and Discretionary Total BTD have quite similar distributional patterns with slightly positive BTD. In line with a mean of 0.0071 and a small standard deviation of 0.0439, the Discretionary Permanent BTD is clustered closely around zero. Overall, the distributions of the single measures exhibit considerable vari-

All in all, the comparison of means, standard deviations, and distributions has shown that there are differences between annual measures. Additionally, possible explanations for those differences have been given. However, it cannot be ruled out, that those differences are due to sampling error.

### 5.2.2. Development of Measures over Time

To examine whether the means of the annual measures face changes over time, Figure 3 and 4 display the means of those over the time horizon 1996-2015. In the figures it becomes evident that on average the ETR decreased over time while the HS and BTD measures increased. These findings are in line with prior research. Graham and Tucker (Graham and Tucker (2006), pp. 564–565) find a decrease in Effective Tax Rates for S&P 500 firms over the period of 1988-2004 and reports a similar pattern for all publicly traded firms, while Desai (2003) finds an increase in BTD during the 1990s. A decrease in ETR corresponding with a simultaneous increase in BTD suggests, that both measures could capture an increase in tax avoidance (as argued in Desai (2003), p. 170 for BTD).

It becomes evident, that the differences between the means as described above are constant over time. The order of the means (from highest to lowest) remains the same for the ETR measures (except for  $CFM_1$ ) and BTD. This supports the argumentation that there are differences between the measures. However, all measures seem to follow the same pattern over time, which suggests, that despite their differences they might capture, at least to some point, similar aspects of tax avoidance (alternatively the measures might be driven by other economy wide powers).

## 5.2.3. Correlation between Annual Measures

Although summary statistics and frequency distributions give a first insight about the properties and similarities of the different measures, it is interesting to investigate how closely related the measures are, given that despite differences in means, they seem to behave in some way similarly over time. This can be examined using a correlation matrix. The correlation matrix is based on the Bravais-Pearson correlation coefficient, which can be used for metric data such as the data at hand. The Bravais-Pearson correlation coefficient yields values from -1 to 1, where -1 indicates a perfect negative linear relationship and 1 a perfect positive linear relationship; 0 represents no linear relationship (Toutenburg et al. (2009), pp. 130–137). The correlation matrix of the annual measures is displayed in Table 2.

The correlation matrix shows that there is a high positive linear relationship between ETR measures ranging from 0.3953 between GAAP ETR and  $CFM_2$  to 0.7789 between  $CFM_1$  and  $CFM_2$ . This is not surprising since the measures are based on the same rationale and use similar inputs. Also not surprising is that measures with the same inputs in the numerator or denominator such as  $CFM_1$  and  $CFM_2$  (both using operating cash flow) or Cash and Current ETR (both using pre-tax income) seem to be more highly correlated. The correlation between ETR measures and the HS measure is also positive, yet lower. This may be due to the different interpretations of the measures. As one would expect, the correlation coefficients between ETR and BTD are negative, suggesting that a low ETR is associated with positive BTD. However, the correlation decreases when ETR measures are compared with Discretionary Total BTD or Discretionary Permanent BTD. The BTD measures are positively correlated among one another (ranging from 0.1343 to 0.6507), which should be due to the similarities of computation.

Overall, the correlation analysis suggests that ETR measures behave quite similarly. Also ETR and BTD are correlated, meaning that those measures should measure similar aspects of tax avoidance. BTD measures (especially Discretionary Total and Permanent BTD) face a weaker correlation among themselves, suggesting that those measures might be able to proxy different aspects of tax avoidance. However, since this analysis is only descriptive in nature, one cannot rule out the possibility that those relationships are sample-specific.

### 5.3. Comparison of Annual and Long-Run Measures

In a second step, several measures of long-run tax avoidance (spanning time periods of five to ten years) are compared with one another. These measures are then compared to annual measures in order to investigate how closely related annual and long-run measures are and how good the former are in predicting long-run tax aggressiveness.

# 5.3.1. Distributional Characteristics of Long-Run Measures

Looking at the means in Table 3 (and Table 1), Long-Run Cash ETR seems to increase as the observation period

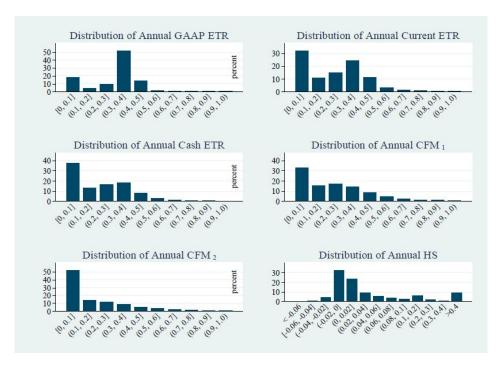


Figure 1: Distribution of Annual ETR and HS

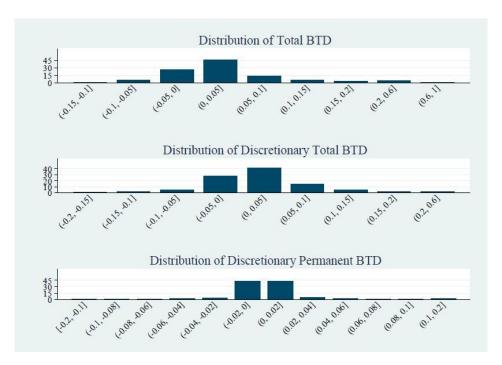


Figure 2: Distribution of BTD

increases (from annual to ten years 0.2124, 0.2415, 0.2456, respectively). This suggests that it becomes more difficult for firms to maintain a low Cash ETR over time. Conversely, the HS measure decreases as the observation period increases (from annual to ten years 0.0281, 0.0048, 0.0042, respectively). So, according to the HS measure, tax preferences decrease over time, meaning that the actual tax payments converge to the hypothetical tax payments. Taken together,

both long-run measures seem to converge to a lower degree of tax avoidance as the observation period increases, suggesting that it is difficult to maintain tax benefits over long periods of time.

The standard deviation reveals, that the variation of the measures decreases as the observation period increases (from annual to ten years 0.1900, 0.1518, 0.1450, respectively, for Cash ETR, and 0.0677, 0.0195, 0.0167, respectively, for HS),

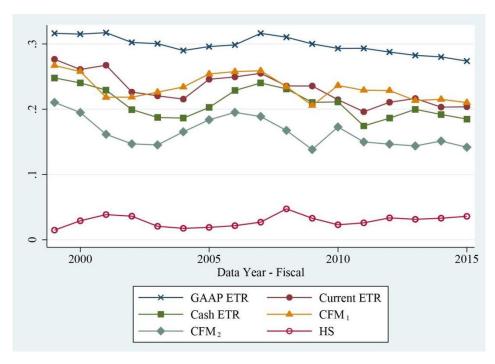


Figure 3: Development of ETR and HS Measures over Time

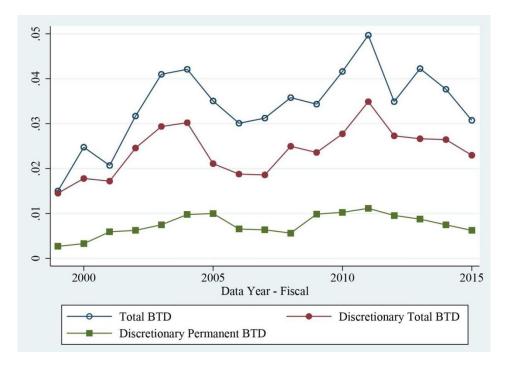


Figure 4: Development of BTD Measures over Time

which is in line with the theoretical argumentation.

Similarly, the distributions depicted in Figure 5 (and Figure 1) show that over time, the distributions tighten as the observation period increases. For Cash ETR, the fractions of observations close to zero and one decrease while more observations fall into the range of (0.3, 0.4] (near the statutory tax rate of 0.35). The HS measure gets clustered even closer around zero.

5.3.2. Correlation between Annual and Long-Run Measures
To compare annual and long-run measures, the BravaisPearson correlation coefficients are displayed in Table 4.

The correlation coefficients between annual ETR and long-run measures of tax avoidance are positive. This positive correlation with long-run measures is strongest for Current and Cash ETR (between 0.5147 and 0.6664 for Current ETR and between 0.4962 and 0.6793 for Cash ETR), The cor-

Table 2: Correlation Matrix of Annual Measures

	GAAP ETR	Current ETR	Cash ETR	$CFM_1$	$CFM_2$	HS	Total BTD	Discr. Total BTD	Discr. Per- manent BTD
GAAP ETR	1.0000								
Current ETR	0.5857	1.0000							
Cash ETR	0.5135	0.7306	1.0000						
$CFM_1$	0.5830	0.4482	0.4015	1.0000					
$CFM_2$	0.3953	0.5636	0.7301	0.7789	1.0000				
HS	0.3298	0.5216	0.7410	0.2169	0.5628	1.0000			
Total BTD	-0.3619	-0.6386	-0.4946	-0.2493	-0.3899	-0.5971	1.0000		
Discr. Total BTD	-0.3343	-0.5069	-0.4037	-0.2399	-0.3385	-0.4545	0.6507	1.0000	
Discr. Permanent BTD	-0.1845	-0.1094	-0.0836	-0.1213	-0.0718	-0.0800	0.1505	0.1343	1.0000

**Table 3:** Distributional Characteristics of Long-Run Measures

Variable	Obs.	Mean	Std. Dev.	Min	Max
lr. Cash ETR <sub>5</sub>	682	0.2415	0.1518	0	0.9348
lr. Cash $ETR_{10}$	354	0.2456	0.1450	0	0.8357
lr. HS5	911	0.0048	0.0195	-0.0262	0.1099
lr. $HS_{10}$	470	0.0042	0.0167	-0.0275	0.0997

Table 4: Correlation between Annual and Long-Run Measures

	lr. Cash ETR <sub>5</sub>	lr. Cash ETR <sub>10</sub>	lr. HS5	lr. <i>HS</i> <sub>10</sub>
	21115	27710		
GAAP ETR	0.3809	0.4051	0.3244	0.3397
Current ETR	0.6664	0.6243	0.5544	0.5147
Cash ETR	0.6793	0.6030	0.5751	0.4962
CFM1	0.4205	0.4559	0.3200	0.3507
CFM2	0.5793	0.5503	0.4859	0.4434
HS	0.5229	0.4467	0.5691	0.4629
Total BTD	-0.4786	-0.4435	-0.5165	-0.4554
Discr. Total BTD	-0.3696	-0.3374	-0.3859	-0.3271
Discr. Permanent BTD	0.0370	0.0324	0.0490	0.0380
Lr Cash ETR <sub>5</sub>	1.0000			
Lr. Cash ETR <sub>10</sub>	0.8448	1.0000		
Lr. HS5	0.8019	0.6894	1.0000	
Lr. $HS_{10}$	0.6776	0.8051	0.7765	1.0000

relation of  $CFM_2$  ranges between 0.4434 and 0.5793 while those of GAAP ETR and  $CFM_1$  are lower, yet still positive. The correlation is stronger between annual and long-run ETR measures than between annual ETR and long-run HS. Also, Annual HS is positively correlated with the long-run measures while the correlation between annual and long-run measures of the same kind is higher than across measures. It appears plausible that the correlation between annual and long-run measures increases with higher similarity in inputs and computation. For all ETR measures except GAAP ETR and  $CFM_1$  the correlation with the five-year measures is higher than with the ten-year measures. Similarly, the

correlation between Annual HS and the long-run measures decreases over time.

The Total and Discretionary Total BTD are negatively correlated to the proxies for long-run tax avoidance. Notably, this negative correlation decreases as the observation period increases. Further, the correlation for Discretionary Total BTD (ranging from -0.3271 to -0.3696) is lower than for Total BTD (ranging from -0.4554 to -0.5165). Surprisingly, the Discretionary Permanent BTD is positively correlated with the long-run measures (ranging from 0.0324 to 0.0490). However, this correlation is close to zero suggesting no linear relationship. Overall, the BTD seem to be higher (negatively)

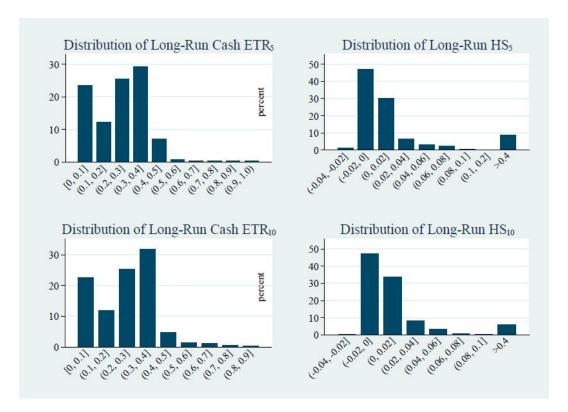


Figure 5: Distribution of Long-Run Measures

correlated with the Long-Run HS measure than with Long-Run Cash ETR. As one would expect, the correlations between the five- and ten-year measure of the same kind are very high (0.8448 between lr. Cash  $ETR_5$  and lr. Cash  $ETR_{10}$  and 0.7765 between lr. HS5 and lr.  $HS_{10}$ ). The correlation between Long-Run Cash ETR and Long-Run HS is high for both five- and ten-year measures. This suggests that both kinds of long-run measures will detect similar aspects of long-run tax avoidance.

Altogether, the correlation between most annual and long-run measures is relatively high. Not surprisingly, annual ETR (HS) measures are more highly correlated with long-run ETR (HS) measures. In most cases the correlation between BTD and Long-run HS is higher than between BTD and Long-Run Cash ETR. The correlation between long-run measures is very high.

# 5.3.3. Reliability of Annual Measures in Predicting Long-Run Tax Avoidance

Although the correlation coefficients show the relationship between the single measures, it does not become clear, how reliable annual measures can predict long-run tax aggressiveness. Therefore, an additional analysis is conducted. The proxies based on annual numbers and over time are sorted into two groups, "low" and "high" depending on the values they take. For ETR measures, an observation is allocated to the group "low" if the Effective Tax Rate is below 0.35 and "high" if the tax rate is above or equal to 0.35. BTD measures are classified as "low" if BTD > 0 and "high" if BTD

 $\leq$  0. The HS measure is defined as "high" if HS > 0 and "low" if HS  $\leq$  0. Subsequently, the observations are counted for which both the annual and long-run measures either take "low" or "high" values, i.e. for which the annual measure correctly identifies the firm to pay a low (high) amount of corporate taxes according to the longrun measure; Table 5 column (A) displays the relative frequencies for which annual measures identify the long-run tax behavior correctly.

As one can see, the reliability of annual measures in predicting long-run tax avoidance is not very high. GAAP ETR is only able to identify 0.5854 of the observations correctly according to Long-Run Cash ETR. The fractions for  $CFM_1$  and  $CFM_2$  are slightly higher (0.6529, 0.6342 (0.6534, 0.6250) for Long-Run Cash ETR (HS), respectively). Only Current and Cash ETR exhibit a greater reliability with fractions of approximately 0.67 to 0.68 for both long-run measures. The Annual HS measure is able to identify 0.6822 of observations correctly according to the Long-Run Cash ETR. According to the Long-Run HS measure, the fraction of correctly identified observations increases to 0.6902. Total BTD are able to identify 0.4512 (0.4776) of observations correctly according to Long-Run Cash  $ETR_{10}$  ( $HS_{10}$ ). This fraction increases for Discretionary Total BTD (0.4848, 0.5172 for Long-Run Cash  $ETR_{10}$  and Long-Run  $HS_{10}$ , respectively) and Discretionary Permanent BTD (0.5616, 0.5888, respectively). Interestingly, according the Long-Run  $HS_{10}$ , the reliability of annual measures is in most cases higher than the Long-Run Cash ETR<sub>10</sub> suggests. When looking separately at the correct identification of "low tax paying firms" and "high tax paying firms" it

Table 5: Reliability of Annual Measures in Predicting Long-Run Tax Avoidance

		(A)		(B)		
		lr. Cash	lr. HS <sub>10</sub>	lr. Cash	lr. HS <sub>10</sub>	
		$ETR_{10}$		$ETR_{10}$		
	low	0.7303	0.6814	0.5214	0.6674	
GAAP ETR	high	0.5212	0.6150	0.7364	0.6366	
	total	0.5854	0.6354	0.6610	0.6474	
	low	0.7545	0.6790	0.6288	0.6871	
Current ETR	high	0.6100	0.6937	0.7832	0.6384	
	total	0.6758	0.6870	0.7318	0.6546	
	low	0.7393	0.6648	0.6089	0.6756	
Cash ETR	high	0.6229	0.7083	0.7750	0.6340	
	total	0.6804	0.6868	0.7192	0.6480	
	low	0.7094	0.6370	0.5399	0.6080	
CFM1	high	0.5901	0.6716	0.7421	0.6013	
	total	0.6529	0.6534	0.6732	0.6036	
	low	0.6681	0.5893	0.5311	0.5610	
CFM2	high	0.5901	0.6716	0.7458	0.5813	
	total	0.6342	0.6250	0.6684	0.5740	
	low	0.7456	0.6713	0.5345	0.6864	
HS	high	0.6225	0.7080	0.7879	0.7000	
	total	0.6822	0.6902	0.6764	0.6940	
	low	0.5166	0.3928	0.2264	0.4485	
Total BTD	high	0.4331	0.5010	0.5825	0.5193	
	total	0.4512	0.4776	0.4642	0.4958	
	low	0.5851	0.4858	0.2928	0.4978	
Discr. Total BTD	high	0.4532	0.5271	0.6167	0.5433	
	total	0.4848	0.5172	0.5118	0.5286	
	low	0.6844	0.6016	0.4112	0.5942	
Discr. Permanent BTD	high	0.5041	0.5828	0.6745	0.5918	
	total	0.5616	0.5888	0.5870	0.5926	

becomes clear that the persistence of "low tax paying firms" is higher looking at the Long-run Cash  $ETR_{10}$  measure, while the reverse is the case according to the Long-Run  $HS_{10}$  measure.

These results may be distorted by economy-wide shocks that could result in large losses for the firms on average in single years. Thus all or at least a majority of the firms would exhibit low tax payments in years around those shocks (due to loss carry backs and loss carry forwards) not associated with tax avoidance. In those years, comparing ETR with the statutory tax rate and HS and BTD with zero to decide about the degree of tax avoidance would lead to wrong inferences. To test whether the results in Table 5, column (A) are driven by such economy-wide shocks, column (B) compares the measures of a year with the corresponding median of all firms in that year.

As one can see, the overall reliability of annual measures in predicting long-run tax avoidance increases in most cases when medians instead of fixed values are used in allocating observations (this is true for both measures). Interestingly, the persistence of observations being identified as "low" decreases while the reverse is the case for "high" observations. Thus, using fixed values for allocation purposes may be mis-

leading.

Overall, the reliability of annual measures in depicting long-run tax avoidance is not very high. Only looking at one year seems not to provide enough information to make solid conclusions about the tax aggressiveness across time since a substantial fraction of observations would be incorrectly classified as "high" or "low". Furthermore, the decisions about the tax aggressiveness solely made on fixed limits may be misleading in single years.

### 6. Conclusion

The aim of this bachelor thesis is to give an overview of and analyze the different measures of tax avoidance used in the literature and to answer the question of how closely related these measures are among one another. In the empirical tax research there are several proxies for tax avoidance, most of which rely on financial accounting data. Most prevalent are measures based on ETR and BTD. Besides these two, measures from Henry and Sansing (2014), Tax Shelter Scores, and the Unrecognized Tax Benefits are used to proxy for tax avoidance. All measures differ with respect to the underlying rationale of measuring tax avoidance and the inputs

used. This leads to the question of how closely related the measures are, which is tested empirically.

The descriptive results suggest that, although there are differences between the single measures and those differences persist over time, especially annual proxies exhibit considerable correlation. This correlation gets stronger as the similarity in computation and inputs increases and is strongest between measures of the same kind. The results suggest that the different measures may in fact be closely related. A comparison between annual and long-run measures implies that the reliability of annual measures in depicting long-run tax avoidance seems not sufficiently high. Since the samples used in the empirical investigation may be subject to a substantial truncation bias due to stringent assumptions, the results may only be sample-specific.

In conclusion it should be mentioned that the right choice for a measure of corporate tax avoidance heavily depends on the data availability and aims of the research question at hand. Since some measures require more inputs than others and may exclude/include specific forms of tax avoidance, which others do not, specific measures may be more suitable than others. Also a common approach of researchers if data availability is no issue is to use several different proxies simultaneously to investigate, how their results are influenced by the proxies for tax avoidance.

Besides the existing literature, the need for research about (measuring) corporate tax avoidance still remains. First, a universally accepted definition of tax avoidance is still missing, which prevents the creation of more powerful measures. Although one single sharp definition may be unrealistic, Hanlon and Heitzman (Hanlon and Heitzman (2010), p. 137) argue that continued research may increase the probability of an accepted definition. Second, since all existing measures exhibit limitations, the search for new forms of measuring tax avoidance should be continued. Third, further research should analyze the relationships between measures beyond purely descriptive statistics to be able to make universal statements. Among others, these are issues that need to be addressed in the future.

#### References

- Armstrong, C. S., Blouin, J. L., and Larcker, D. F. The incentives for tax planning. *Journal of Accounting and Economics*, 53(1):391–411, 2012.
- Badertscher, B. A., Katz, S. P., Rego, S. O., and Wilson, R. J. Conforming tax avoidance and capital market pressure. *SSRN Electronic Journal*, 2015.
- Bonsall, S. B., Koharki, K., and Watson, L. Deciphering tax avoidance: Evidence from credit rating disagreements. Retrieved June 28, 2016, 2015. URL http://papers.ssrn.com/sol3/papers.cfm?abstract\_ id=2511454.
- Cen, L., Maydew, E. L., Zhang, L., and Zuo, L. Customer-supplier relationships and corporate tax avoidance. SSRN Electronic Journal, 2016.
- Center for Research in Security Prices. Annual data industrial. Retrieved July 12, 2016, 2016. URL http://www.crsp.com/products/docume ntation/annual-data-industrial.
- Chen, S., Chen, X., Cheng, Q., and Shevlin, T. Are family firms more tax aggressive than non-family firms? *Journal of Financial Economics*, 95(1): 41–61, 2010.
- Dechow, P. M. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics*, 18(1):3–42, 1994.
- Desai, M. A. The divergence between book income and tax income. *Tax Policy and the Economy*, 17:169–206, 2003.
- Desai, M. A. and Dharmapala, D. Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics*, 79(1):145–179, 2006.
- Desai, M. A. and Dharmapala, D. Earnings management, corporate tax shelters, and book-tax alignment. *National Tax Journal*, pages 169–186, 2009a.
- Desai, M. A. and Dharmapala, D. Corporate tax avoidance and firm value. *The Review of Economics and Statistics*, 91(3):537–546, 2009b.
- Dover, R., Ferrett, B., Gravino, D., Jones, E., and Merler, S. Bringing transparency, coordination and convergence to corporate tax policies in the European Union: Part I: Assessment of the magnitude of aggressive corporate tax planning. European Parliamentary Research Service, Retrieved June 20, 2016, 2015. URL http://www.europarl.europa.eu/RegData/etudes/STUD/2015/558773/EPRS\_STU(2015)558773\_EN.pdf.
- Dyreng, S. D., Hanlon, M., and Maydew, E. L. Long-run corporate tax avoidance. *The Accounting Review*, 83(1):61–82, 2008.
- Dyreng, S. D., Hanlon, M., and Maydew, E. L. The effects of executives on corporate tax avoidance. *The Accounting Review*, 85(4):1163–1189, 2010
- European Comission. Anti tax avoidance package. Retrieved July 14, 2016, 2016. URL http://ec.europa.eu/taxation\_customs/taxation/company\_tax/anti\_tax\_avoidance/index\_en.htm.
- Financial Accounting Standards Board (FASB). FASB Interpretation No. 48: Accounting for Uncertainty in Income Taxes. Retrieved July 04, 2016, from Financial Accounting Standards Board (FASB):, 2006. URL http://www.imanet.org/docs/default-source/advocacy/frc/frc\_200606\_fasb48.pdf?sfvrsn=2.
- Frank, M. M., Lynch, L. J., and Rego, S. O. Tax reporting aggressiveness and its relation to aggressive financial reporting. *The Accounting Review*, 84 (2):467–496, 2009.
- Graham, J. R. and Tucker, A. L. Tax shelters and corporate debt policy. *Journal of Financial Economics*, 81(3):563–594, 2006.
- Guenther, D. A. Measuring corporate tax avoidance: Effective tax rates and book-tax differences. Retrieved June 10, 2016, 2014. URL http://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2478952.
- Gupta, S. and Newberry, K. Determinants of the variability in corporate effective tax rates: Evidence from longitudinal data. *Journal of Accounting* and Public Policy, 16(1):1–34, 1997.
- Hanlon, M. What can we infer about a firm's taxable income from its financial statements? *National Tax Journal*, 56(4):831–863, 2003.
- Hanlon, M. The persistence and pricing of earnings, accruals, and cash flows when firms have large book-tax differences. *The Accounting Review*, 80 (1):137–166, 2005.
- Hanlon, M. and Heitzman, S. A review of tax research. *Journal of Accounting and Economics*, 50(2):127–178, 2010.
- Hanlon, M. and Shevlin, T. Accounting for tax benefits of employee stock options and implications for research. *Accounting Horizons*, 16(1):1–16, 2002.
- Healy, P. M. and Wahlen, J. M. A review of the earnings management liter-

- ature and its implications for standard setting. *Accounting horizons*, 13 (4):365–383, 1999.
- Henry, E. and Sansing, R. C. Data truncation bias and the mismeasurement of corporate tax avoidance. *SSRN Electronic Journal*, 2014.
- Hope, O.-K., Ma, M. S., and Thomas, W. B. Tax avoidance and geographic earnings disclosure. *Journal of Accounting and Economics*, 56(2):170–189, 2013.
- Houlder, V. Leak reveals scale of corporate tax deals with Luxembourg. Retrieved July 19, 2016, from Financial Times:, 2014. URL https://next.ft.com/content/93e75c1a-6545-11e4-91b1-00144feabdc0.
- Koester, A., Shevlin, T. J., and Wangerin, D. The role of managerial ability in corporate tax avoidance. Retrieved June 28, 2016, 2016. URL http: //papers.ssrn.com/sol3/papers.cfm?abstract\_id=2753152.
- KPMG. Corporate tax rates table. Retrieved July 15, 2016, 2016. URL https://home.kpmg.com/xx/en/home/services/tax/tax-tools-and-resources/tax-rates-online/corporate-tax-rates-table.html.
- Lee, B. B., Dobiyanski, A., and Minton, S. Theories and empirical proxies for corporate tax avoidance. *The Journal of Applied Business and Economics*, 17(3):21–34, 2015.
- Lennox, C., Lisowsky, P., and Pittman, J. Tax aggressiveness and accounting fraud. *Journal of Accounting Research*, 51(4):739–778, 2013.
- Lipton, E. and Creswell, J. Panama papers show how rich united states clients hid millions abroad. Retrieved July 19, 2016, from New York Times, 2016. URL http://www.nytimes.com/2016/06/06/us/panam a-papers.html.
- Lisowsky, P. Seeking shelter: Empirically modeling tax shelters using financial statement information. *The Accounting Review*, 85(5):1693–1720, 2010.
- Lisowsky, P., Robinson, L., and Schmidt, A. Do publicly disclosed tax reserves tell us about privately disclosed tax shelter activity? *Journal of Accounting Research*, 51(3):583–629, 2013.
- Manzon Jr, G. B. and Plesko, G. A. The relation between financial and tax reporting measures of income. *Tax Law Review*, 55(2):175–214, 2002.
- Miller, A. and Oats, L. Principles of international taxation (fourth edition). Haywards Heath, West Sussex: Bloomsbury Professional, 2014.
- Organisation for Economic Co-operation and Development (OECD). Base erosion and profit shifting. Retrieved July 26, 2016, 2016. URL http://www.oecd.org/tax/beps/.
- Penman, S. H. Financial statement analysis and security valuation (5. ed., internat. ed.). New York, NY: McGraw-Hill Irwin, 2013.
- PWC. United States Corporate Taxes on corporate income. Retrieved July 15, 2016, 2016. URL http://taxsummaries.pwc.com/uk/taxsummaries/wwts.nsf/ID/United-States-Corporate-Taxes-on-corporate-income.
- Salihu, I. A., Obid, S. N. S., and Annuar, H. A. Measures of corporate tax avoidance: Empirical evidence from an emerging economy. *International Journal of Business and Society*, 14(3):412–427, 2013.
- Shackelford, D. A. and Shevlin, T. Empirical tax research in accounting. Journal of Accounting and Economics, 31(1-3):321–387, 2001.
- Tang, T. and Firth, M. Can book-tax differences capture earnings management and tax management? Empirical evidence from China. *The International Journal of Accounting*, 46(2):175–204, 2011.
- The Guardian. Luxembourg tax files: how tiny state rubber-stamped tax avoidance on an industrial scale. Retrieved July 19, 2016, 2015. URL https://www.theguardian.com/business/2014/nov/05/-sp-lux embourg-tax-files-tax-avoidance-industrial-scale.
- Toutenburg, H., Heumann, C., and Schomaker, M. Deskriptive Statistik: Eine Einführung in Methoden und Anwendungen mit R und SPSS (7., aktualisierte und erw. Aufl.). Springer-Lehrbuch. Berlin: Springer, 2009.
- U.S. Department of the Treasury. The problem of corporate tax shelters: Discussion, analysis and legislative proposals. Retrieved July 14, 2016, 1999. URL https://www.treasury.gov/resource-center/tax-policy/Documents/Report-Corporate-Tax-Shelters-1999.pdf.
- U.S. Department of the Treasury. Tackling the growth of corporate tax shelters. Retrieved July 26, 2016, 2000. URL https://www.treasury.gov/press-center/press-releases/Pages/1s421.aspx.
- Waegenaere, A., Sansing, R., and Wielhouwer, J. L. Financial accounting effects of tax aggressiveness: Contracting and measurement. *Contemporary Accounting Research*, 32(1):223–242, 2015.
- Wharton Research Data Services. Retrieved July 12, 2016, 2016a. URL

- $\verb|https://wrds-web.wharton.upenn.edu/wrds/.|$
- Wharton Research Data Services. Compustat monthly updates fundamentals annual. Retrieved July 22, 2016, 2016b. URL https://wrds-web.wharton.upenn.edu/wrds/ds/compm/funda/index.cfm?navId=84.
- Wilkie, P. J. and Limberg, S. T. Measuring explicit tax (dis)advantage for corporate taxpayers: An alternative to average effective tax rates. *The Journal of the American Taxation Association*, 15(1):46–71, 1993.
- Wilson, R. J. An examination of corporate tax shelter participants. *The Accounting Review*, 84(3):969–999, 2009.
- Wooldridge, J. M. Introductory econometrics: A modern approach (5. ed., internat. ed.). Mason, Ohio: South-Western Cengage Learning, 2013.
- Zimmerman, J. L. Taxes and firm size. *Journal of Accounting and Economics*, 5:119–149, 1983.



# **Junior Management Science**

journal homepage: www.jums.academy



# Der Einfluss von CEO Duality auf die Bewertung von Schweizer Unternehmen

Raphael Uster

Universität Zürich

### **Abstract**

Die Gestaltung der obersten Führungsstruktur stellt börsenkotierte Unternehmen vielfach vor grosse Herausforderungen. Die Wahl zwischen einer Personalunion von CEO und Verwaltungsratspräsident – oft als CEO Duality oder Doppelmandat bezeichnet – oder einer getrennten Doppelspitze führt regelmässig zu starken Reaktionen von Investoren, Politikern und Wirtschaftsverbänden. Die bisherige Forschung liefert für den Zusammenhang von CEO Duality und der Bewertung der Unternehmen am Kapitalmarkt sowohl international als auch in der Schweiz inkonsistente Ergebnisse. Die vorliegende Arbeit untersucht den Einfluss von CEO Duality auf die Bewertung von 194 Schweizer Unternehmen aus dem Swiss Performance Index für die Jahre 2000 bis 2015. Zur Überprüfung werden zwei gegensätzliche Hypothesen aufgestellt, die auf den beiden vorherrschenden Denkschulen, Agency Theory und Stewardship Theory, basieren. Ein handgesammelter Paneldatensatz erlaubt es, multiple Regressionsanalysen im Rahmen einer Panelanalyse durchzuführen. Es resultiert ein statistisch signifikant negativer Zusammenhang zwischen CEO Duality und Tobin's Q, der verwendeten Kennzahl für die Unternehmensbewertung. Unternehmen mit einem Doppelmandat haben einen um 0.1959 tieferen Tobin's Q als Unternehmen mit einer Doppelspitze. Dieses Resultat ergibt eine gewisse Evidenz für die Hypothese der Agency Theory, die eine Trennung der beiden Funktionen vorsieht. Die Ergebnisse dieser Arbeit bestätigen somit die Richtung, die der Swiss Code of Best Practice for Corporate Governance mit der Präferenz für die Trennung der beiden Funktionen verfolgt. Gleichzeitig wird ersichtlich, dass ein striktes Verbot von CEO Duality die Schweizer Unternehmen unnötig einschränken würde und nicht optimal sein kann. Somit kann die momentane Regulierungsgrundlage als vernünftig beurteilt werden.

Keywords: CEO Duality, Corporate Governance, Chief Executive Officer, Personalunion, Unternehmensbewertung

# 1. Einleitung

### 1.1. Ausgangslage und Problemstellung

Unternehmensskandale in den Vereinigten Staaten von Amerika, exzessive Saläre von Top-Managern in der Schweiz und die staatliche Rettung der UBS während der globalen Finanzkrise haben dazu geführt, dass der Begriff Corporate Governance in den letzten Jahren stetig an Bedeutung gewonnen hat. Die Reaktionen der Regulierungsbehörden auf diese Ereignisse im 21. Jahrhundert sind dabei sehr unterschiedlich ausgefallen. Beispielsweise veröffentlichten die USA 2002 den Sarbanes-Oxley Act als Gesetz für börsenkotierte US-Unternehmen (Amstutz (2007), S.79). In der Schweiz wurde im gleichen Jahr die erste Version des Swiss Code of Best Practice for Corporate Governance, nachfolgend Swiss Code genannt, vom Dachverband der Schweizer Wirtschaft Economiesuisse veröffentlicht.

Böckli et al. (2004, S. 76) nennen in ihrem Expertenbericht zur Teilrevision des Aktienrechts die Wahl der Ausge-

staltung der Führungsstruktur auf oberster Ebene in einem Unternehmen als eine der umstrittensten Grundsatzfragen in der Corporate Governance Forschung. Einerseits kann eine Unternehmung eine Struktur wählen, bei welcher dieselbe Person gleichzeitig als Verwaltungsratspräsident und als Chief Executive Officer amtet. Dabei wird von einer sogenannten Personalunion oder CEO Duality gesprochen. Andererseits kann eine Unternehmung durch eine Trennung dieser beiden Funktionen gekennzeichnet sein. In diesem Fall liegt eine Doppelspitze vor; es gibt also je einen Verwaltungsratspräsidenten und einen separaten CEO als Vorsitzenden der Geschäftsleitung.

Diese Thematik wird seit der Jahrtausendwende in der Presse häufig sehr kontrovers diskutiert. Mitverantwortlich ist die Tatsache, dass die wissenschaftliche Forschung bisher keine eindeutige Meinung hervorbrachte. Die Auswirkungen dieser Organisationsstrukturen auf die Performance und die Bewertung von Unternehmen scheinen sehr unterschiedlich und vom jeweiligen Land und ihren Regulierungsbehörden

abhängig zu sein. Dennoch ist es für die Unternehmen essentiell zu verstehen, wie sich die Ausgestaltung und auch Veränderungen ihrer Führungsstruktur auswirken werden.

# 1.2. Zielsetzung

Das übergeordnete Ziel der vorliegenden Bachelorarbeit ist die quantitative Untersuchung der Auswirkung von CEO Duality auf die Unternehmensbewertung von Schweizer Unternehmen. Mittels einer Paneldatenanalyse über mehrere Jahre soll ermittelt werden, ob das Vorhandensein einer solchen Governance-Struktur einen positiven oder negativen Effekt auf die Unternehmen hat. Schmid und Zimmermann (2008) führten in ihrer Studie mit dem Titel "Should Chairman and CEO Be Separated? Leadership Structure and Firm Performance in Switzerland" eine ähnliche Analyse für die Jahre 1993 bis 2002 durch. Sie erwähnen dabei insbesondere, dass sich der 2002 veröffentlichte Swiss Code somit noch nicht auf ihr Sample ausgewirkt hat (Schmid und Zimmermann (2008), S. 190).

Der Swiss Code hat jedoch seither grossen Einfluss auf die Qualität und die Umsetzung der Corporate Governance in der Schweiz (Amstutz (2007), S. 82). Bezüglich CEO Duality gibt die revidierte Version 2014 eine Präferenz für die Trennung der beiden Funktionen vor, erlaubt jedoch die Personalunion weiterhin unter gewissen Voraussetzungen, wenn spezifische Gründe vorliegen (Economiesuisse (2014), S.12). Die vorliegende Arbeit versucht zu erfassen, inwiefern sich diese Thematik in den Jahren 2000 bis 2015 verändert hat und ob die Empfehlungen des Swiss Code präzise genug formuliert sind. Mittels einer ausführlichen Literaturanalyse soll zudem der aktuelle Stand der Forschung aufbereitet werden, um die Vor- und Nachteile der beiden Organisationsstrukturen ableiten und verstehen zu können.

Die Schweiz ist gemäss Schmid und Zimmermann (2008, S. 202) gut für eine solche Analyse geeignet, da sich die Grundvoraussetzungen bezüglich CEO Duality von Ländern wie beispielsweise den USA oder Deutschland erheblich unterscheiden. Die Resultate aus der vorliegenden Arbeit mit Schweizer Unternehmen und aktuellen Daten sind also einerseits für die Frage der Regulierung in der Schweiz von Bedeutung und andererseits als Ergänzung für die bisher vor allem amerikanisch geprägte Forschung auf diesem Gebiet der Corporate Governance.

# 1.3. Vorgehensweise und Methodik

Die Arbeit ist grob in zwei Teile gegliedert. Der erste Teil umfasst eine theoretische Auseinandersetzung zum Thema der CEO Duality als Aufbau für den zweiten Teil. Der Fokus liegt dabei einerseits auf der Corporate Governance in der Schweiz mit den Akteuren Aktionär, Verwaltungsrat und Management. Andererseits wird die internationale Forschung zur CEO Duality anhand von den zwei vorherrschenden Denkschulen, der Agency Theory und der Stewardship Theory, aufbereitet. Daraus resultieren zwei gegensätzliche Hypothesen bezüglich des Zusammenhangs von CEO Duality und der Bewertung von Unternehmen.

Der zweite Teil besteht aus der empirischen Analyse, um die beiden aufgestellten Hypothesen testen zu können. Dabei wird eine Paneldatenanalyse mit 194 an der SIX Swiss Exchange kotierten Unternehmen aus dem Swiss Performance Index für die Jahre 2000 bis 2015 durchgeführt. Die Daten zur CEO Duality basieren auf einer eigenen Erhebung aus den öffentlich verfügbaren Geschäftsberichten der betreffenden Unternehmen. Mittels einer multiplen Regressionsanalyse soll anschliessend der Einfluss auf die Unternehmensbewertung anhand der Kennzahl Tobin's Q gemessen werden. Die erhaltenen Resultate werden im Kontext der bisherigen Studien interpretiert und allfällige Empfehlungen für Politik sowie Unternehmen abgeleitet.

Die Arbeit untersucht primär den Einfluss von CEO Duality auf die Bewertung, also Valuation, von Unternehmen und nicht auf die kurz- oder langfristige relative Performance gegenüber einem Aktienmarktindex. Ebenso wenig wird die Thematik der Unabhängigkeit von Mitgliedern in Verwaltungsräten untersucht oder beachtet. Die vorliegende Arbeit ist in vier Kapitel aufgeteilt. Nach dieser Einleitung behandelt das zweite Kapitel die Theorie in insgesamt drei Unterkapitel. Das Kapitel 2.1 befasst sich mit der Corporate Governance im Allgemeinen und mit den Besonderheiten im Schweizerischen Umfeld. Das Kapitel 2.2 ist der CEO Duality gewidmet und betrachtet diese sowohl im internationalen wie auch im schweizerischen Kontext. Im Kapitel 2.3 werden zwei Hypothesen aufgestellt, die auf den beiden vorgestellten Denkschulen basieren. Das Kapitel 3 behandelt das Sample, die empirische Methodik und bespricht die erhaltenen Resultate. Das abschliessende Kapitel 4 besteht aus einer Zusammenfassung, einem Fazit und einem kurzen Ausblick.

### 2. Theorie

### 2.1. Corporate Governance

### 2.1.1. Einführung und Definition

Der Begriff Corporate Governance wird in der Praxis, in der Forschung und auch in den Medien auf sehr unterschiedliche Art und Weise benutzt. Die stark steigende Bedeutung der Corporate Governance für Unternehmen ist unter anderem auf die Globalisierung, die Konzernbildung und auf fatale Managemententscheidungen in der Vergangenheit zurückzuführen (Volkart und Wagner (2014), S. 1054). Böckli et al. (2004, S. 15) definieren den Begriff Corporate Governance als sogenannte Steuerungsmechanismen, "die ein ausgewogenes Verhältnis von Führung und Kontrolle und Transparenz innerhalb von Aktiengesellschaften anstreben und die darauf abzielen, die Rolle der Aktionäre als Träger des Kapitalrisikos zu stärken".

Der Dachverband Economiesuisse (2014, S. 6) definiert die Corporate Governance in ihrem Swiss Code of Best Practice for Corporate Governance als "die Gesamtheit der auf das nachhaltige Unternehmensinteresse ausgerichteten Grundsätze, die unter Wahrung von Entscheidungsfähigkeit und Effizienz auf der obersten Unternehmensebene Transparenz

und ein ausgewogenes Verhältnis von Führung und Kontrolle anstreben." Dieser Grundgedanke widerspiegelt sich in den Empfehlungen des Swiss Codes, die in 19 von 29 Artikeln den Verwaltungsrat und die Geschäftsleitung betreffen.

Hofstetter (2002, S. 6) unterscheidet in seinem Bericht zum Swiss Code in der ersten Version von 2002 eine enge und eine weite Definition der Corporate Governance. Die Corporate Governance im weiteren Sinne behandelt allgemeine Organisationsfragen zum Schutz der Aktionäre des Unternehmens. Im engeren Sinne hingegen geht es lediglich um die Ausgestaltung der obersten Führungsstruktur. Dabei steht das häufig verwendete Prinzip der Checks and Balances<sup>1</sup> im Zentrum, also das richtige Verhältnis von Führung und Kontrolle im Unternehmen (Hofstetter (2002), S. 6).

Die vorliegenden Erläuterungen zeigen, dass es keine eindeutige Definition gibt. Relevant ist jedoch, dass diese Thematik eine hohe und weltweite Bedeutung für Unternehmen, Anleger, Behörden und Forscher geniesst. Passend in diesem Zusammenhang ist der Titel vom Buch von Larcker und Tayan (2016): Corporate Governance matters.

### 2.1.2. Unterschiede in der Regulierung

Die spezifischen Regulierungen und die Vorgehensweisen der Behörden betreffend Corporate Governance sind von Land zu Land unterschiedlich ausgeprägt. Ein kurzer Überblick der Regulierungsgrundlage in der Schweiz im Vergleich mit den Vereinigten Staaten von Amerika, Grossbritannien und weiteren Ländern soll helfen, die Rahmenbedingungen zu verstehen, mit denen sich die Unternehmen in ihrer Tätigkeit konfrontiert sehen.

2002 hat der Dachverband der Schweizer Wirtschaft Economiesuisse erstmals den Swiss Code of Best Practice for Corporate Governance veröffentlicht. Der Swiss Code beinhaltet Empfehlungen einer Expertengruppe<sup>2</sup> für die börsenkotierten Unternehmen in der Schweiz. Gleichzeitig veröffentlichte die Schweizer Börse SIX Swiss Exchange (2016), früher SWX Six Exchange, erstmals die Richtlinie betreffend Informationen zur Corporate Governance (RLCG). Der Swiss Code wurde erstellt, da die fortschreitende Globalisierung und Transparenzanforderungen solche Best-Practice-Empfehlungen zur Corporate Governance Anfangs des 21. Jahrhunderts erforderten (Hofstetter (2002), S.1). Ebenfalls verhindern oder verzögern solche Selbstregulierungsrichtlinien im Rahmen von Empfehlungen oftmals strengere behördliche Gesetze (Amstutz (2007), S. 80).

Die erstmalige Fassung des Swiss Codes wurde 2007 durch einen Anhang mit Empfehlungen zur Entschädigung ergänzt und 2014 aufgrund der eidgenössischen Volksinitiative gegen die Abzockerei und Entwicklungen in der Corporate

Governance umfassend revidiert. Eine wesentliche Änderung ist die Einführung des Comply or Explain-Prinzips, welches die RLCG bereits seit 2002 kannte.

Comply or Explain bedeutet, dass die Unternehmen die vorgeschlagenen Leitideen und Empfehlungen entweder annehmen (Comply) oder erklären (Explain) müssen, warum sie es nicht tun (Economiesuisse (2014), S. 6). Diese Änderung geht darauf zurück, dass sich die Governance in der Schweiz seit Einführung des Swiss Codes deutlich verbessert hat und sich der Swiss Code dadurch auch als ein etwas stärkeres Instrument<sup>3</sup> legitimiert hat (Amstutz (2007), S. 82).

International unterscheiden sich die Arten der Regulierungen wesentlich. Beispielsweise hat der amerikanische Kongress als Reaktion auf die Unternehmensskandale von Enron und WorldCom 2002 den Sarbanes-Oxley Act durchgesetzt (Amstutz (2007), S. 249). Dieses Gesetz enthält umfangreiche Weisungen zur obersten Führungsstruktur und zu Tätigkeiten der Manager von amerikanischen Firmen (Larcker und Tayan (2016), S. 31). 2010 wurde mit dem Dodd-Frank Act als staatliche Reaktion auf die globale Finanzkrise das zweite wichtige Governance-Gesetz der USA erlassen (Bainbridge (2012), S. 12). Grossbritannien hingegen hat bereit 1992 mit dem Cadbury Report Best-Practice-Richtlinien herausgegeben (Hofstetter (2002), S. 5). Nach diversen Zusammenschlüssen von Reports zum sogenannten Combined Code gilt heute für börsenkotierte Unternehmen in Grossbritannien der U.K. Corporate Governance Code (Larcker und Tayan (2016), S. 32). Weiter gibt es auch die G20/OECD-Grundsätze der Corporate Governance der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung, die erstmalig 1999 veröffentlicht wurden und als einen angesehenen internationalen Standard für viele Unternehmen und Investoren gelten (OECD (2015), S. 3).

# 2.1.3. Prinzipal-Agenten-Problem

Die Corporate Governance basiert auf einem Problem, welches sich aus der grundsätzlichen Ausgestaltung moderner Kapitalgesellschaften ergibt. Die Trennung von Eigentum und Kontrolle führt zum sogenannten Prinzipal-Agenten-Problem (Volkart und Wagner (2014), S. 1054). Der Aktionär, der Eigentümer der Unternehmung, offeriert sein Kapital und gibt gleichzeitig die Kontrolle an das Management ab. Er ist der Prinzipal und die Geschäftsleitung sein Agent. Obwohl der Aktionär der Prinzipal ist, hat er keinerlei Entscheidungsmacht mehr über die strategische und operative Führung des Unternehmens (Bainbridge (2012), S. 2). Die Trennung von Eigentum und Kontrolle, die bereits von Berle und Means (1932) beschrieben wurde, ist also eine implizite Eigenschaft einer modernen Kapitalgesellschaft (Bainbridge (2012), S. 2).

Jensen und Meckling (1976, S. 308) definieren die Prinzipal-Agenten-Beziehung allgemein als einen Vertrag

<sup>&</sup>lt;sup>1</sup>Das Prinzip der Checks and Balances bedeutet in diesem Zusammenhang die ausgewogene Verteilung der Macht und Kontrolle zwischen den Aktionären, dem Verwaltungsrat, der Geschäftsführung und der Revisionsstelle (Hofstetter (2002), S. 28).

<sup>&</sup>lt;sup>2</sup>In dieser Expertengruppe arbeiten unter anderem Prof. em. Peter Böckli (Universität Basel), Prof. em. Peter Forstmoser und Prof. Karl Hofstetter (beide Universität Zürich) mit (Economiesuisse (2014), S. 5).

<sup>&</sup>lt;sup>3</sup>Mit der Einführung des "Comply or Explain"-Prinzips geht der Swiss Code nun über die völlig freiwillige Art und Weise der Selbstregulierung hinaus und fordert von den Unternehmen eine Erklärung, warum sie gewisse Richtlinien nicht einhalten. Dies wurde bereits früh gefordert, um den Code mit "mehr Muskeln zu versehen" (Amstutz (2007), S. 83).

zwischen dem Prinzipal und dem Agenten. In dieser Beziehung ist dabei entscheidend, dass der Agent Entscheidungsbefugnisse vom Prinzipal erhält. Es findet also die bereits erwähnte Trennung von Eigentum (Aktionär) und Kontrolle beziehungsweise Führung (Management) statt. Die Aktionäre haben ausserhalb der Generalversammlung keinen Einfluss mehr, was das Management mit ihrem Kapital macht (Amstutz (2007), S. 77). Da jedoch auch die Generalversammlung als Zusammenkunft der Eigentümer nur auf gewisse Aspekte Einfluss nehmen kann, delegieren die Aktionäre die Oberleitung und Kontrolle der Gesellschaft gemäss Artikel 716a OR an den Verwaltungsrat. Die Beziehung zwischen dem Prinzipal und dem Agenten führt zur logischen Folgerung, dass der Agent nicht immer im Interesse des Prinzipals handelt und dies in den sogenannten Agency Costs resultiert (Jensen und Meckling (1976), S. 208). Ein Manager eines Unternehmens hat möglicherweise andere Ziele oder Interessen als ein Kleinaktionär, welcher nur wenige Aktien des Unternehmens hält.

Wie steht dies nun im Zusammenhang mit der Governance eines Unternehmens? Das System der Corporate Governance kommt ins Spiel, um diese Agency Costs, die aus dem Prinzipal-Agenten-Problem resultieren, verringern zu können (Larcker und Tayan (2016), S. 4). Corporate Governance soll einerseits die Interessen möglichst aller Aktionäre schützen und andererseits eine angemessene Führung des Unternehmens ermöglichen, die auf Transparenz und ausgewogenen Machtverhältnissen mittels Checks and Balances beruht (Amstutz (2007), S. 77). Ammann et al. (2005, S. 3) betonen dabei die Wichtigkeit der Beziehung zwischen Aktionär und Verwaltungsrat, da der VR eine qualitativ hochstehende Corporate Governance implementieren muss. Die Corporate Governance beeinflusst somit direkt den Shareholder Value, dessen Maximierung die Aktionäre letztendlich anstreben (Amstutz (2007), S. 78).

# 2.1.4. Die Schweizerische Aktiengesellschaft

Die Schweizerische Aktiengesellschaft ist gemäss den Artikeln 620 bis 771 im Obligationenrecht eine Rechtsform einer Kapitalgesellschaft. Artikel 620 Absatz 1 definiert die AG als eine Gesellschaft mit eigener Firma, deren Aktienkapital in Aktien zerlegt ist und für deren Verbindlichkeiten nur das Gesellschaftsvermögen haftet. Die Aktiengesellschaft ist somit nützlich, wenn Firmenanteile beispielsweise an der Börse gehandelt werden sollen, Konzerne gebildet werden und eine klare Struktur der Haftungsverhältnisse angestrebt wird (Volkart und Wagner (2014), S. 655). In der Schweiz können sowohl kleine, mittlere als auch sehr grosse Unternehmen als Aktiengesellschaft organisiert sein.

Die Aktionäre haben grundsätzlich die Wahl, ob die Aktiengesellschaft einen Börsengang, also eine Erstplatzierung oder Initial Public Offering, durchführen soll oder ob die Gesellschaft eine private Eigentümergesellschaft bleibt (Volkart und Wagner (2014), S. 661). Sobald eine Aktiengesellschaft jedoch nach einem IPO an einer Börse kotiert ist und die Aktien von Investoren gekauft werden können, handelt es sich um eine Publikumsgesellschaft. Dabei wird nach Volkart

und Wagner (2014, S. 58) zwischen einer echten und einer unechten Publikumsgesellschaft unterschieden. Eine echte Publikumsgesellschaft besitzt ein breitgestreutes Aktionariat. Eine unechte Publikumsgesellschaft hingegen hat einen Mehrheitsaktionär, welcher einen wesentlichen Anteil am Unternehmen besitzt. Dabei handelt es sich in der Schweiz oft um eine Familie (Volkart und Wagner (2014), S. 58).

Die Ausgestaltung der Anteilsrechte der Aktiengesellschaft hat einen grossen Einfluss auf die Corporate Governance des Unternehmens. Bei einer echten Publikumsgesellschaft mit breitem Aktionariat ist das im vorhergehenden Kapitel besprochene Prinzipal-Agenten-Problem am stärksten ausgeprägt. Zwischen den Kleinaktionären und den Managern kann es zu zahlreichen Ziel- und Interessenskonflikten kommen (Volkart und Wagner (2014), S. 58). Sobald aber eine Person oder eine Familie als Mehrheitsaktionär auftritt, verringert sich dieses Problem, da Vertreter dieser Aktionäre meistens in der operativen Führung auftreten (Hofstetter (2002), S. 8). Somit ist die Trennung von Eigentum und Kontrolle nicht mehr wie bei einer echten Publikumsgesellschaft mit Kleinaktionären gegeben. Diese Struktur wiederum kann aber auch zu anderen Problematiken führen. Beispielsweise können sich Differenzen zwischen Mehrheits- und Minderaktionären oder Problematiken mit grossen institutionellen Investoren ergeben (Hofstetter (2002), S. 8). Hofstetter (2002, S. 8) betont dabei, dass gerade in der Schweiz bei Diskussionen bezüglich der Ausgestaltung der Corporate Governance von Aktiengesellschaften sowohl echte als auch unechte Publikumsgesellschaften berücksichtigt werden müssen. In der Folge wird der Einfachheit halber, wo nicht anders vermerkt, auf diese Unterscheidung verzichtet und unter einer Publikumsgesellschaft ein öffentliches und kotiertes Unternehmen verstanden.

### Generalversammlung und Aktionäre

Die rechtliche Situation für den Aktionär ist in der Schweiz durch das Obligationenrecht geregelt. Die Generalversammlung, also die Zusammenkunft der Aktionäre eines Unternehmens, ist gemäss Art. 698 Abs. 1 OR das oberste Organ der Aktiengesellschaft. Art. 698 Abs. 2 Ziff. 2 nennt beispielsweise die Wahl der Mitglieder des Verwaltungsrates und der Revisionsstelle als eine der sechs unübertragbaren Aufgaben der Generalversammlung.

Fama und Jensen (1983, S. 302) bezeichnen ein Unternehmen als ein Nexus von Verträgen. Die Aktionäre übernehmen dabei die Rolle der Residualhalter, nachdem alle anderen Parteien wie beispielsweise die Mitarbeiter oder Fremdkapitalgeber für ihre gelieferten Ressourcen vertraglich entschädigt wurden. Die einzige Rolle der Aktionäre ist somit das Tragen des Residualrisikos, da ihre Anteilsrechte jederzeit übertragbar sind (Fama und Jensen (1983), S. 303). Der Swiss Code hingegen bezeichnet die Aktionäre als letzte Entscheidungsinstanz, da sie das Eigenkapital bereitstellen (Economiesuisse (2014), S. 7). In der ersten Empfehlung werden analog zum Obligationenrecht die Kompetenzen der Aktionäre festgehalten und sie werden aufgefordert, ihre Mitspracherechte an der GV auch aktiv wahrzunehmen (Economie-

suisse (2014), S. 7).

Die Aktionäre sind in der vorliegenden Arbeit aufgrund folgender Tatsache relevant: Wenn Aktionäre mit der Corporate Governance ihres Unternehmens nicht einverstanden sind, dann haben sie zwar beschränkte Möglichkeiten, direkt Einfluss auf das Management zu nehmen, jedoch können sie den Unternehmenswert, gemessen an den Kapitalmärkten, erheblich beeinflussen. Dies erreichen sie beispielsweise durch das Verkaufen ihrer Aktienposition. Somit kann sich eine Veränderung in der Organisationsstruktur sehr schnell auf den Aktienkurs und somit auf den Unternehmenswert auswirken. In diesem Zusammenhang wird bei Publikumsgesellschaften mit breitem Aktionariat häufig vom Shareholder Value oder Aktionärswert als alternative Bezeichnung für den Unternehmenswert gesprochen.

Verwaltungsrat und Verwaltungsratspräsident

Das Problem der Trennung von Eigentum und Kontrolle in einer Publikumsgesellschaft führt dazu, dass die Aktionäre meistens im Hintergrund bleiben. Sie delegieren die Kontrolle und die Führung des Unternehmens an den von ihnen im Rahmen der Generalversammlung gewählten Verwaltungsrat. Dadurch verhindern sie zwar eine Desorganisation, welche entstehen würde, wenn sehr viele Aktionäre am täglichen Management teilhaben würden. Auf der anderen Seite schaffen sie so das in Kapitel 2.1.3 besprochene Prinzipal-Agenten-Problem (Bainbridge (2012), S. 3).

Ammann et al. (2005, S. 3) nennen die Beziehung zwischen dem Aktionär und dem Verwaltungsrat als eines der wichtigsten Themen der Corporate Governance. Der Schweizerische Verwaltungsrat ist volkswirtschaftlich betrachtet eine Möglichkeit, um das Prinzipal-Agenten-Problem zu lösen oder zu verringern (Ammann et al. (2005, S. 3), S.3). von der Crone (2001, S. 29) beschreibt die klare Trennung zwischen Aktionär und Verwaltungsrat folgendermassen: Die Generalversammlung, als oberstes Organ der Aktiengesellschaft gemäss OR, regelt die Grundlagen betreffend Personalentscheidungen und gibt somit die Richtung vor, während der Verwaltungsrat die Strategie und die Leitung des Unternehmens übernimmt.

Der rechtliche Rahmen für den Verwaltungsrat ist im Obligationenrecht im Artikel 716 geregelt. Gemäss Art. 716a Abs. 1 OR hat der Verwaltungsrat folgende sieben Aufgaben, die er weder an jemanden übertragen noch sich davon entziehen kann:

- 1. die Oberleitung der Gesellschaft und die Erteilung der nötigen Weisungen;
- 2. die Festlegung der Organisation;
- die Ausgestaltung des Rechnungswesens, der Finanzkontrolle sowie der Finanzplanung, sofern für die Führung der Gesellschaft notwendig
- 4. die Ernennung und Abberufung der mit der Geschäftsführung und der Vertretung betrauten Personen;
- die Oberaufsicht über die mit der Geschäftsführung betrauten Personen, namentlich im Hinblick auf die Befolgung der Gesetze, Statuten, Reglemente und Weisungen;

- 6. die Erstellung des Geschäftsberichtes sowie die Vorbereitung der Generalversammlung und die Ausführung ihrer Beschlüsse
- 7. die Benachrichtigung des Richters im Falle der Überschuldung.

Wie sofort erkennbar wird, hat der Verwaltungsrat als einzelnes Organ sowohl die Oberleitung als auch die Kontrolle inne. Diese Konzentration von Leitung und Kontrolle in einer einzigen Institution wird als monistisches System bezeichnet und widerspricht jeglichen Ansätzen der Gewaltentrennung, um ein ausgeglichenes System der Checks and Balances zu schaffen (Amstutz (2007), S. 157). Ein dualistisches System mit getrennten Organen ist beispielsweise in Deutschland zu finden, wo der Aufsichtsrat die Aufsicht und Kontrolle übernimmt, während der Vorstand ausschliesslich für die operative Führung verantwortlich ist. Ein solches System ist gemäss OR in der Schweiz jedoch nicht erlaubt, da der Verwaltungsrat die letztendliche Leitung der Gesellschaft sicherstellt und nicht nur für die Aufsicht verantwortlich sein kann. Falls jedoch nur nicht-exekutive Mitglieder im Verwaltungsrat vertreten sind, dann würde dieser Verwaltungsrat dem Deutschen Aufsichtsrat relativ nahekommen (von der Crone (2001), S. 29).

Dieses grundsätzliche Problem, dass der Verwaltungsrat die Gesellschaft führt und sich selbst gleichzeitig kontrolliert, wird jedoch durch die Möglichkeit der Delegation der operativen Geschäftsleitung im Obligationenrecht relativ elegant gelöst (Amstutz (2007), S. 157). Nach Art. 716b Abs. 1 OR<sup>4</sup> kann der Verwaltungsrat die exekutive Geschäftsführung, also das daily business, an einzelne seiner Mitglieder oder an Dritte übertragen. Dies bedeutet, dass der Verwaltungsrat die operative Geschäftsführung entweder an einen Delegierten des Verwaltungsrates oder an einen Direktor (CEO) bzw. eine Geschäftsleitung übertragen kann. Er behält aber zwingend die strategische Oberleitung und Oberaufsicht über die Gesellschaft und über die vom ihm betraute Geschäftsführung. Wichtig zu betonen ist in diesem Kontext, dass die Hierarchie zwischen diesen beiden Organen klar geregelt ist, da die Geschäftsleitung dem Verwaltungsrat in jedem Fall unterstellt ist (Amstutz (2007), S. 157).

Der Swiss Code befasst sich sehr ausführlich mit der Gestaltung des Verwaltungsrats. Höchst relevant ist dabei die neunte Empfehlung bezüglich der Aufgaben des Verwaltungsrates, die über die aktienrechtliche Definition des Aufgabenbereichs hinausgeht. Diese Empfehlung wurde im Rahmen der Revision des Swiss Codes 2014 angepasst. Der Verwaltungsrat ist demnach für die spezifische Definierung und Umsetzung der Corporate Governance zuständig. Weiter enthält der Swiss Code vielfältige Empfehlungen zur Zusammensetzung, Unabhängigkeit, Arbeitsweise und zu den verschiedenen Ausschüssen des Verwaltungsrates (Economiesuisse (2014), S.9ff).

Gemäss Obligationenrecht besteht der Verwaltungsrat aus mindestens einer Person, meistens jedoch aus mehre-

<sup>&</sup>lt;sup>4</sup>Falls durch die Statuten dazu ermächtigt.

ren Mitgliedern<sup>5</sup>. Die Wahl des Verwaltungsratspräsidenten erfolgt entweder durch den Verwaltungsrat selbst oder die GV wählt ihn, falls dies so in den Statuten bestimmt ist<sup>6</sup>. Der Präsident leitet die mehrmals jährlich stattfindenden Sitzungen des Verwaltungsrates und bestimmt so wesentlich die Agenda dieser Diskussionen. Er muss dementsprechend das Unternehmen und seine Strategie, Märkte, Produkte und Konkurrenten bestens kennen. Ebenfalls vertretet er das Unternehmen nach Aussen und kommuniziert unternehmensspezifische Angelegenheiten gegenüber den Medien und seinen Stakeholdern (Larcker und Tayan (2016), S. 59).

Der Swiss Code bezeichnet den Verwaltungsratspräsidenten passend als "Garant der Information", welcher den Verwaltungsrat im Interesse der Aktionäre und der Gesellschaft führen soll (Economiesuisse (2014), S. 11). Folglich ist der Verwaltungsratspräsident zusammen mit dem ganzen Verwaltungsrat für die verschiedenen Aufgaben verantwortlich. Amstutz (2007, S. 124) fasst diese Aufgaben zu den folgenden fünf Haupttätigkeiten zusammen, die in der Abbildung 1 dargestellt sind:

## Geschäftsleitung und CEO

Wie im vorhergehenden Kapitel erläutert, kann der Verwaltungsrat die Geschäftsführung gemäss Obligationenrecht an Mitglieder des Verwaltungsrats oder auch an Dritte delegieren. Die Entscheidung, ob er die Geschäftsführung delegiert oder nicht, fällt der Verwaltungsrat im Rahmen der zweckmässigen Strukturierung der Führungs- und Kontrollfunktionen in der obersten Führungsebene des Unternehmens (Economiesuisse (2014), S. 10).

Wenn der Verwaltungsrat einen Delegierten aus den eigenen Reihen bestimmt, dann übernimmt dieses Mitglied als exekutives Verwaltungsratsmitglied die operative Führung des Unternehmens. Exekutive Mitglieder des Verwaltungsrats werden gemäss Kommentar zur Corporate Governance-Richtlinie als "diejenigen Personen bezeichnet, welche im Unternehmen operative Führungsaufgaben erfüllen" SWX Swiss Exchange (2007), S. 15). Andererseits kann der Verwaltungsrat auch Dritte mit der Geschäftsführung beauftragen. Damit sind Direktoren im Sinne einer Geschäftsleitung gemeint, die nicht dem Verwaltungsrat angehören.

Grundsätzlich ist der Verwaltungsrat gemäss Art. 716a Abs. 1 Ziff. 4 OR mit der Ernennung der Geschäftsleitung betraut. Er ernennt den Chief Executive Officer als Vorsitzender der GL, welcher meistens noch weitere Direktoren wie beispielsweise der Chief Financial Officer angehören. Die klare Definition und Abgrenzung der Aufgabengebiete und Kompetenzen zwischen Geschäftsleitung und Verwaltungsrat soll gemäss der Empfehlung 11 im Swiss Code im Rahmen eines Organisationsreglements geregelt werden (Economiesuisse (2014), S. 10). Ersichtlich wird diese Kompetenzregelung zwischen den beiden Parteien im jährlichen Corporate Governance Bericht als Teil des Geschäftsberichts unter dem Gliederungspunkt 3.6, falls das Unternehmen die Richtlinien

betreffend Informationen zur Corporate Governance der SIX Swiss Exchange befolgt (SIX Swiss Exchange (2016), S. 3).

Die spezifischen Aufgaben des CEO und der Geschäftsleitung sollen zur Veranschaulichung anhand der Kompetenzregelung im Geschäftsbericht 2015 von Lindt & Sprüngli dargestellt werden. Eine konkrete Aufgabe des CEO als Vorsitzender der Geschäftsleitung von Lindt & Sprüngli umfasst die Informationsbeschaffung und die Informationsweiterleitung an die gesamte Geschäftsleitung und an den Verwaltungsrat. Ebenso ist er für die operativen Geschäfte des Konzerns und die korrekte Ausführung der Beschlüsse des Verwaltungsrats durch die Geschäftsleitung verantwortlich. Die Geschäftsleitung hingegen ist gesamthaft für die Umsetzung der beschlossenen Strategie und zudem alle Mitglieder einzeln für ihre jeweiligen Bereiche verantwortlich (Lindt & Sprüngli (2016), S. 25).

Dieses erste theoretische Kapitel 2.1 behandelt einerseits allgemeine Aspekte der Corporate Governance wie das Prinzipal-Agenten-Problem und andererseits spezifische Themen im Schweizerischen Umfeld wie beispielsweise die sehr häufig anzutreffende Aktiengesellschaft mit ihren einzelnen Elementen. Im nächsten Kapitel wird ein sehr umstrittenes und viel diskutiertes Thema im Rahmen einer Good Corporate Governance beleuchtet und eingeführt: die CEO Duality als Situation, wenn dieselbe Person gleichzeitig als CEO und als Verwaltungsratspräsident amtet.

# 2.2. CEO Duality

# 2.2.1. Einführung und Definition

Sobald ein Unternehmen eine gewisse Grösse erreicht oder sobald es einen Börsengang anstrebt, stellt sich unvermeidlich die Frage nach der optimalen Führungsstruktur auf der obersten Ebene des Unternehmens. Wie gestaltet sich das Management? Wie ist das Verhältnis zwischen Management und Verwaltungsrat geregelt? Die Beantwortung dieser Fragen ist höchst relevant für börsenkotierte Unternehmen im Rahmen ihrer Corporate Governance Bestrebungen und endet in der Grundsatzfrage, wie die Positionen des CEO und des Verwaltungsratspräsidenten optimal besetzt werden sollen (Amstutz (2007), S. 157/158).

Bezüglich der Definitionen und den Begriffen zur CEO Duality kommt es sowohl in der Forschung als auch in der Presse immer wieder zu Verwechslungen und Verwirrungen. Eine sorgfältige Definition ist deshalb besonders wichtig. Gemäss Swiss Code wird die Situation, in welcher eine erste Person als Verwaltungsratspräsident, also als Vorsitzender des VR, und eine zweite Person als CEO, also als Vorsitzender der Geschäftsleitung, amtet, als "Doppelspitze" bezeichnet (Economiesuisse (2014), S. 12). Es gibt mit dem VRP und dem CEO eine doppelte Spitze im Unternehmen, die gemeinsam die Führung sicherstellen. Wenn jedoch dieselbe Person gleichzeitig sowohl als Verwaltungsratspräsident wie auch als CEO amtet, dann handelt es sich um eine "Personalunion" dieser beiden Funktionen (Economiesuisse (2014), S.12).

Im englischen Sprachgebrauch wird in diesem Fall von CEO Duality gesprochen. Duality wird dabei aus der Sicht-

<sup>&</sup>lt;sup>5</sup>Art. 707 Abs. 1 OR.

<sup>&</sup>lt;sup>6</sup>Art. 712 Abs. 1 und 2 OR.



Abbildung 1: Führungsaufgaben des Verwaltungsrats (Quelle: Amstutz (2007), S. 124, eigene Darstellung)

weise des CEO verstanden, welcher gleichzeitig zwei Funktionen innehat. Rechner und Dalton (1991, S. 1) sprechen symbolisch von einer Führungsperson, welche als CEO und als VRP zwei verschiedene Hüte trägt. Nachfolgend wird analog dem Swiss Code von einer Personalunion, falls CEO Duality vorliegt, und von einer Doppelspitze, falls keine CEO Duality vorliegt, gesprochen. Die Entscheidung und die Ausgestaltung der Führungsstruktur als Personalunion oder als Doppelspitze hat speziell für die Aktionäre einen hohen Stellenwert. Amstutz (2007, S. 157/158) weist darauf hin, dass die Aktionäre an ausgeglichenen Machtverhältnissen an der Spitze ihres Unternehmens und einer angemessenen Kontrolle des Managements interessiert sind.

# 2.2.2. CEO Duality im internationalen Kontext Übersicht

Als Michael C. Jensen im Januar 1993 seine Presidential Address vor der American Finance Association hielt, äusserte er sich mit folgenden Worten zur CEO Duality Thematik in den Vereinigten Staaten von Amerika (Jensen (1993), S. 866):

"It is common in U.S. corporations for the CEO to also hold the position of chairman of the board. The function of the chairman is to run board meetings and oversee the process of hiring, firing, evaluating, and compensating the CEO. Clearly, the CEO cannot perform this function apart from his or her personal interest. Without the direction of an independent leader, it is much more difficult for the board to perform its critical function. Therefore, for the board to be effective, it is important to separate the CEO and chairman positions. "

Diese deutliche Aussage und die Meinung, dass diese Positionen getrennt sein sollten, wird weltweit von vielen Corporate Governance Experten unterstützt und vertreten (Larcker und Tayan (2016), S. 112). Etwas überraschend ist dabei die Tatsache, dass die grosse Mehrheit der amerikanischen, börsenkotierten Firmen dennoch eine Personalunion aufweisen. Beispielsweise sind dies über 80% im Sample von 661 Unternehmen von Brickley et al. (1997, S. 197) für das Jahr 1998 und 79% im Sample von 141 Unternehmen von Rechner und Dalton (1991, S. 156) für die Jahre 1978 bis 1983. Auch noch nach der Jahrtausendwende nennen Larcker und Tayan (2016, S. 110) mit 77% in 2003, 61% in 2008 und 55% in 2013 hohe Werte für das Vorliegen von CEO Duality, jedoch ist hier ein abnehmender Trend zu erkennen.

In der Corporate Governance Forschung ist diese Thema-

tik der Ausgestaltung der Führungsstruktur höchst umstritten. Die akademische Literatur ist sich uneinig über die Auswirkungen auf die Performance und die Bewertung von Unternehmen, wenn zwischen Personalunion und Doppelspitze unterschieden wird (Tonello (2011), S. 1). Es gibt keine klare Evidenz, da die Ergebnisse aus diversen Studien und sogar aus Meta-Studien, die die Ergebnisse verschiedener Studien zusammenfassen und daraus Schlüsse ziehen wollen, sehr unterschiedlich ausfallen (Bainbridge (2012), S. 106).

Die bisherige empirische Literatur zu diesem Thema kann grundsätzlich zwei gegensätzlichen Denkschulen zugeordnet werden. Die Gegner der Personalunion befürchten, dass sich der CEO/VRP im Unternehmen nicht mehr kontrollieren lässt. Die Befürworter argumentieren meist mit den Vorteilen einer einheitlichen und effizienten Führungsstruktur (Finkelstein und D'aveni (1994), S. 1079). Die erste Sichtweise kann der Denkschule der Agency Theory zugeordnet werden, während es sich bei letzterer um die Stewardship Theory<sup>7</sup> handelt. Diese beiden Denkschulen werden nun im Folgenden einzeln behandelt. Ebenfalls werden anhand dieser beiden Theorien die Vorteile und die Nachteile bezüglich der Wahl einer Personalunion oder einer Doppelspitze erarbeitet.

### Agency Theory

Die Denkschule der Agency Theory basiert auf dem Prinzipal-Agenten-Problem zwischen den Aktionären und dem Management. Dieses Problem wurde bereits in Kapitel 2.1.3 beschrieben. Dabei wird angenommen, dass die Manager nicht im Interesse der Aktionäre handeln, solange keine entsprechenden Anreize und Kontrollstrukturen geschaffen werden (Jensen und Meckling (1976), S. 308). Der Verwaltungsrat agiert als ein Mechanismus, um dieses Verhalten der Manager einzuschränken, indem er im Auftrag der Aktionäre die Manager bei ihren Tätigkeiten überwacht (Donaldson und Davis (1991), S. 50). Dieses Monitoring des Managements funktioniert gemäss Donaldson und Davis (1991, S. 50) jedoch nur, falls der Verwaltungsrat einen unabhängigen Präsidenten hat, welcher nicht gleichzeitig als CEO exekutiv tätig ist. Die Vertreter der Agency Theory befürworten also eine Doppelspitze, wobei die Funktionen des Chief Executive Officer und des Verwaltungsratspräsidenten getrennt voneinander sind.

Rechner und Dalton (1991, S. 155) finden heraus, dass Unternehmen ohne CEO Duality eine konstant bessere Performance zeigen als Unternehmen mit einer solchen Struktur. Sie untersuchen dabei für 141 zufällig ausgewählte Fortune 500 Unternehmen, inwiefern Unterschiede in der Performance in den sechs Jahren zwischen 1978 und 1983 beobachtbar sind. Für alle drei verwendeten Performancegrössen, ROE, ROI sowie Gewinnspanne und über alle sechs Jahre schneiden die Firmen ohne CEO Duality deutlich besser ab (Rechner und Dalton (1991), S. 157). Dies impliziert einen negativen Zusammenhang zwischen CEO Duality und Un-

ternehmensperformance, wie es die Agency Theory voraussagen würde. Häufige Kritik an dieser Studie ist jedoch die fehlende Kontrolle für Industrieeffekte und für andere Governance Faktoren (Baliga et al. (1996), S. 44). Ausserdem haben Rechner und Dalton (1989, S. 142) nur zwei Jahre vorher mit dem identischen Sample von US-Firmen leicht positive, aber nicht signifikante Differenzen bei Unternehmen mit CEO Duality feststellen können. In dieser Studie haben sie als Unterschied jedoch anstatt Accounting-basierter Performancemassen risikoadjustierte Überrenditen verwendet (Probst und Raisch (2005), S. 142). Pi und Timme (1993, S. 521) untersuchen den Zusammenhang für ein Sample von 112 Banken für die Jahre 1987 bis 1990. Sie finden heraus, dass Banken mit CEO Duality eine tiefere Performance gemessen am ROA haben und dass diese Beziehung auf dem 10% Niveau statistisch signifikant ist. Diese Studien unterstützen somit die Sichtweise der Agency Theory, dass CEO Duality die Kontrolle des Verwaltungsrats untergräbt und die Unternehmensperformance möglicherweise negativ beeinflusst werden könnte (Boyd (1995), S. 303).

Gemäss der Agency Theory hat die Doppelspitze mit einem starken Verwaltungsratspräsidenten zahlreiche Vorteile gegenüber der Personalunion. Die Reduktion der aus dem Prinzipal-Agenten-Problem resultierenden Agency Costs infolge der Überwachung des Top-Managements ist ein erster Vorteil (Schmid und Zimmermann (2008), S. 185). Weiter führt eine Doppelspitze zu einer klaren Abgrenzung von Verantwortlichkeiten resultierend in einem Verwaltungsratspräsidenten, welcher als Vorsitzender des VR über die grösste Autorität verfügt, und einen Chief Executive Officer, welcher sich auf das operative Tagesgeschäft konzentrieren kann (Larcker und Tayan (2016), S. 113). Ebenfalls nennen Larcker und Tayan (2016, S. 113) den grossen Vorteil, dass weniger Konflikte bezüglich der Nachfolgeplanung und der Managemententschädigung bestehen, wenn der CEO nicht gleichzeitig auch Verwaltungsratspräsident ist. Die Doppelspitze kann jedoch auch zu wesentlichen Nachteilen führen, die zu berücksichtigen sind. Eine doppelte Führung kann zwar bei guter Umsetzung eine effiziente Führung ermöglichen, jedoch kann sie auch Verwirrung im Unternehmen und ineffiziente Entscheidungen herbeiführen (Larcker und Tayan (2016), S. 114). Weiter können Kosten entstehen, die bei einer Personalunion in dieser Form nicht zu finden wären. Beispielsweise kann eine Informationslücke zwischen den beiden Funktionen zu erheblichen Kosten führen, da der CEO meistens über ein weitaus grösseres Wissen verfügt als ein externer Verwaltungsratspräsident (Schmid und Zimmermann (2008), S. 185). Ebenfalls können neue Kosten für das Monitoring des Verwaltungsratspräsidenten selbst entstehen, da auch dieser vom Gesamtverwaltungsrat überwacht werden muss (Larcker und Tayan (2016), S. 114). Als weiteres Problem identifiziert Bainbridge (2012, S. 107) die Möglichkeit einer Rivalität zwischen dem CEO und dem Präsidenten aufgrund der meist sehr stark ausgeprägten Persönlichkeiten.

Zusammengefasst gilt, dass die Vertreter der Agency Theory starke Corporate Governance Mechanismen im Unternehmen verlangen und negativ gegenüber Strukturen wie

<sup>&</sup>lt;sup>7</sup>Die Denkschule der Stewardship Theory wird manchmal auch der Organization Theory oder der Administrative Theory zugeordnet.

CEO Duality eingestellt sind (Donaldson (1990), S. 376). Die Sichtweise der Agency Theory betrachtet die CEO Duality Thematik aber aus einer Perspektive, die oftmals als zu einseitig gilt und zu früh als allgemeingültige Sichtweise akzeptiert wird (Donaldson und Davis (1991), S. 50). Im folgenden Kapitel wird nun die zweite, zur Agency Theory alternative Denkschule betrachtet, die Stewardship Theory. Stewardship Theory

Die eben vorgestellte Sichtweise der Agency Theory unterstellt, dass sich die Manager auf Kosten der Aktionäre opportunistisch verhalten und somit kontrolliert werden müssen (Boyd (1995), S. 303). Donaldson und Davis (1991, S. 51) hingegen vertreten die Meinung, dass Manager auch nichtfinanzielle Ziele haben, die beispielsweise durch ihre starke Identifikation mit dem Unternehmen und durch ein natürliches Verantwortungsgefühl gegenüber den Stakeholdern ausgelöst werden. Diese alternative Sichtweise wird als Stewardship Theory bezeichnet (Donaldson (1990), S. 377).

Die Stewardship Theory basiert auf einer grundlegend anderen Auffassung über das Verhalten von Managern, als es die Agency Theory postuliert. Donaldson und Davis (1991, S: 51) definieren diesen Typ Manager folgendermassen:

"The executive manager, under this theory, far from being an opportunistic shirker, essentially wants to do a good job, to be a good steward of the corporate assets."

Der Manager verhält sich also nicht opportunistisch, sondern handelt im Sinne der Aktionäre und maximiert deren Shareholder Value (Davis et al. (1997)). Diese alternative Ansicht hat diverse Implikationen auf die Corporate Governance und speziell auf die Ausgestaltung der obersten Führungsstruktur, die im starken Kontrast zu den Empfehlungen gemäss Agency Theory stehen.

Die Struktur der Corporate Governance im Unternehmen muss unbedingt eine effiziente Koordination und Führung ermöglichen (Donaldson (1990), S. 377). Dies wird dadurch begründet, dass die Manager kein Motivationsproblem mehr haben und sich somit auf eine starke Unternehmensperformance fokussieren können (Donaldson und Davis (1991). S. 51). Bezüglich der Frage nach der Ausgestaltung der Positionen des CEO und des Verwaltungspräsidenten bedeutet dies, dass die Vorteile einer Trennung der Führung auf oberster Ebene weniger ersichtlich sind (Tonello (2011), S. 1). Donaldson und Davis (1991, S. 52) argumentieren, dass eine Personalunion eine äusserst effiziente und starke Führung erlaubt und somit zu höheren Renditen für die Aktionäre führt. Ebenfalls verringert dies Verwirrung im Management, da es nur eine Person gibt, welche führt, entscheidet und kontrolliert (Finkelstein und D'aveni (1994), S. 1083). Somit impliziert die Stewardship Theory einen positiven Zusammenhang zwischen CEO Duality und der Performance und ist eine mögliche Erklärung für die Praxis der Vereinigung der Führung im Unternehmen in einer einzigen Person.

Statistische Evidenz für die Stewardship Theory konnten einige Studien feststellen. Donaldson und Davis (1991, S.

54) analysieren in ihrem vielbeachteten Paper den Effekt von CEO Duality auf die Aktionärsrendite von insgesamt 321 Unternehmen. Sie erhalten Resultate, die die Stewardship Theory unterstützen. Unternehmen mit CEO Duality haben einen höheren ROE als Unternehmen ohne CEO Duality, wobei diese Differenzen von 3.26% bzw. 2.38% auf dem 5% Niveau statistisch signifikant sind (Donaldson und Davis (1991), S. 56). Brickley et al. (1997, S. 218) finden ebenfalls einen positiven Einfluss für Unternehmen mit CEO Duality und merken an, dass bisherige Studien zahlreiche Kosten, die bei der Trennung der Funktionen anfallen, nicht genügend berücksichtigt haben.

Auch aktuellere Studien finden ähnliche Resultate. Beispielsweise erhalten Dey et al. (2011, S. 29) negative Returns für Unternehmen, die die beiden Funktionen aufgrund von Investorendruck trennen. Elsayed (2007, S. 1210) analysiert den Effekt von CEO Duality für 92 ägyptische Unternehmen in den Jahren 2000 bis 2004 und erhält signifikant positive Koeffizienten für 6 von 19 Industrien<sup>9</sup>, nachdem für das Gesamtsample ein schwach positiver Koeffizient resultierte. Einen hilfreichen Überblick über weitere Studien mit positiven Koeffizienten für die CEO Duality-Performance-Beziehung liefert Boyd (1995, S. 303) in seiner ausführlichen Meta-Analyse.

Zusammengefasst gilt, dass die Stewardship Theory eine Organisationstruktur mit einer autoritären Person als CEO und Verwaltungsratspräsident begrüsst, um effektives und effizientes Management zu ermöglichen. Finkelstein und D'aveni (1994, S. 1080) sprechen dabei von einer sogenannten "unity of command at the top of the firm".

## Zusammenführung der Theorien

In den zwei vorhergehenden Kapiteln wurden die beiden unterschiedlichen Sichtweisen, Agency Theory und Stewardship Theory, vorgestellt, ihre wichtigsten Vertreter und Studien dargelegt und die Implikationen daraus auf die Frage der CEO Duality erarbeitet. Dieses Kapitel soll nun diese Theorien zusammenführen und den Literaturüberblick auf internationaler Ebene abrunden. Ebenfalls werden noch einige zusätzliche Gedanken zur CEO Duality Thematik diskutiert.

Neben den Studien, die signifikant positive oder negative Ergebnisse erhalten haben, gibt es zahlreiche Studien, die keine signifikanten Unterschiede ausmachen konnten. Chaganti et al. (1985, S. 413) untersuchen den Einfluss von CEO Duality auf das Scheitern von 21 Unternehmenspaaren aus der Retail-Branche. Sie finden keinen signifikanten Zusammenhang, während beispielsweise die Grösse des amerikanischen Board of Directors eine signifikante Rolle spielt. Rechner und Dalton (1989, S: 142) haben interessanterweise mit ihrem Sample von 141 amerikanischen Unternehmen im Zeitraum von 1978 bis 1983 ebenfalls keine statistisch signifikanten Unterschiede erhalten. Weiter finden auch Baliga et al. (1996, S. 51) keine signifikanten Unterschiede in der

<sup>&</sup>lt;sup>8</sup>Die Differenzen in der Performance reduzieren sich von 3.26% auf 2.38%, wenn für Industrieeffekte kontrolliert wird.

<sup>&</sup>lt;sup>9</sup>In einer Industrie resultierte ein signifikant negativer Koeffizient. Die restlichen 12 Industrien ergaben keine signifikanten Ergebnisse.

langfristigen Performance von Unternehmen mit oder ohne CEO Duality Struktur.

Larcker und Tayan (2016, S: 116) schlussfolgern, dass die Trennung von CEO und Verwaltungsratspräsident die Unternehmensleistung nicht automatisch und zwingend verbessert, da die empirische Evidenz sehr schwach ist. Dieses Problem der Inkonsistenz der Resultate begründet Boyd (1995, S. 304) mit der Tatsache, dass sowohl die Agency Theory als auch die Stewardship Theory extreme Sichtweisen darstellen. In seinem Modell versucht er, diese beiden Denkschulen zu integrieren und er erhält Resultate, die beide Sichtweisen unterstützen. In einem ersten Schritt kombiniert Boyd elf Studien zur CEO Duality-Performance Beziehung mit inkonsistenten Resultaten zu einer übergreifenden Meta-Analyse. Er führt zudem selbst eine Analyse mit 192 Unternehmen aus 12 Industrien durch und erhält für beide Analysen wie erwartet nicht signifikante Zusammenhänge. In einem zweiten Schritt untersucht er den Einfluss von CEO Duality unter bestimmten Bedingungen wie beispielsweise hohe Komplexität oder Ressourcenmangel (Boyd (1995), S. 309). Er findet heraus, dass sich CEO Duality sowohl positiv wie auch negativ unter gewissen Bedingungen für die Unternehmen auswirkt und somit beide Sichtweisen unterstützen kann (Boyd (1995), S. 309).

Finkelstein und D'aveni (1994, S. 1101) betonen in diesem Zusammenhang die Wichtigkeit, die Grenzen der Agency Theory zu beachten und andere Sichtweisen zuzulassen. Sie ergänzen, dass die beiden Theorien nicht im totalen Gegensatz zueinanderstehen müssen, sondern dass beide ihre Berechtigung in gewissen Situationen und unter gewissen Bedingungen haben (Finkelstein und D'aveni (1994), S. 1102). Ausserdem weisen Larcker und Tayan (2016, S. 115) darauf hin, dass amerikanische Firmen eine Trennung eher aus Gründen der Nachfolgeplanung durchführen. Unternehmen nutzen den zusätzlichen Titel des Verwaltungsratspräsidenten sehr häufig als Belohnung und Anerkennung für den CEO, sobald dieser eine gewisse Probezeit als "einfacher" CEO erfolgreich bestanden hat (Brickley et al. (1997), S. 218).

Eine häufig diskutierte Massnahme für Unternehmen, die ihre CEO Duality nicht aufgeben können oder wollen, ist die Einberufung eines sogenannten Lead Director als starkes und unabhängiges Verwaltungsratsmitglied neben dem Präsidenten. Dieser Lead Director gilt gemäss Bainbridge (2012, S. 140) als Kompromiss zwischen den Vertretern und den Gegnern der CEO Duality Thematik, da die Unternehmen in den USA gemäss dem Dodd-Frank Act<sup>10</sup> lediglich die Gründe für eine Personalunion darlegen müssen. Ausserdem kritisiert Bainbridge (2012, S. 106) die Anfangs des Kapitels 2.2.2 zitierte Presidential Address von Michael Jensen, da die Überwachung des CEO eine Aufgabe des Gesamtverwaltungsrates sei und nicht nur des Präsidenten.

Abschliessend gilt es nochmals zu betonen, dass die bisherige internationale Literatur zur CEO Duality Thematik sehr ausführlich, aber höchst inkonsistent und umstritten ist. Im

folgenden Kapitel wird deshalb auf die spezifische Situation in der Schweiz eingegangen.

### 2.2.3. CEO Duality im schweizerischen Kontext

Übersicht und bisherige Forschung

Die Situation in der Schweiz ist sehr verschieden im Hinblick auf die Thematik der CEO Duality - verglichen beispielsweise mit den USA. Im Sample von Schmid und Zimmermann (2008, S. 183) haben nur knapp 20% der Unternehmen in der Schweiz eine Personalunion im Jahre 2002, während die restlichen 80% der Unternehmen die Funktionen des CEO und des Verwaltungsratspräsidenten trennen. Ruigrok et al. (2006, S. 1215) erhalten einen Wert von 31% für das Jahr 1999. Amstutz (2007, S. 164/281) untersucht die Unternehmen des Swiss Market Index für das Jahr 2005. Von den 26 Unternehmen haben sechs davon eine Personalunion, namentlich Ciba, Kudelski, Nestlé, Novartis, Roche und Synthes. Dies entspricht 23% der grössten Unternehmen der Schweiz.

In den letzten Jahren ist die Häufigkeit der CEO Duality Struktur in der Schweiz jedoch stark abnehmend. In einer kürzlich erschienenen Studie haben von den grössten 100 Unternehmen im Jahre 2016 (2009) gerade noch 2% (10%) eine Personalunion als Führungsstruktur, jedoch 14 Unternehmen, wenn alle Unternehmen des SPI betrachtet werden (Ethos Services (2016), S. 48 und S. 49). Dieser abnehmende Trend ist einer seits auf den Swiss Code als wirkungsvolle Richtlinie und andererseits auf den zunehmenden Druck der Öffentlichkeit, der Presse, der Aktionäre und Vereinigungen wie der Ethos Stiftung zurückzuführen. Unternehmen wie Novartis mit Daniel Vasella oder Nestlé mit Peter Brabeck änderten ihre Führungsstruktur mehr oder weniger freiwillig und trennten die Funktionen des Verwaltungsratspräsidenten und des CEO, nachdem die Rufe nach Good Corporate Governance immer lauter wurden.

Forschung zum Thema der CEO Duality wurde auch in der Schweiz durchgeführt, jedoch hauptsächlich vor der Finanzkrise. Von der Crone<sup>11</sup> (2001, S. 29) beschäftigt sich seit Jahren mit diesem Thema und ist der Meinung, dass eine funktionale Sichtweise die Frage nach der Personalunion oder der Trennung der beiden Funktionen besser beleuchtet als eine rein hierarchische Sichtweise. Er unterscheidet zwischen der Leitungsfunktion und der Qualitätssicherungsfunktion. Erstere soll der Geschäftsleitung übertragen werden, welcher die exekutiven Verwaltungsratsmitglieder angehören. Letztere ist die Aufgabe, die den restlichen Verwaltungsräten zusteht, also den nicht-exekutiven Mitgliedern des VR. von der Crone (2001, S. 29) kommt zum Schluss, dass die Frage nicht allgemein, sondern nur unter Beachtung der jeweiligen spezifischen Situation eines Unternehmens beantwortet werden kann. Grundsätzlich ist er jedoch der Meinung, dass eine Personalunion in gewissen Situationen überlegen ist und dass in solchen Situationen unbedingt ein verantwortlicher, nicht-exekutiver Verwaltungsrat eingesetzt

<sup>&</sup>lt;sup>10</sup>Sektion 973 im Dodd-Frank Act.

<sup>&</sup>lt;sup>11</sup>Professor für Privat- und Wirtschaftsrecht an der Universität Zürich.

werden muss, welcher die Qualitätssicherungsfunktion übernehmen kann (von der Crone (2001), S. 29).

Ammann et al. (2005, S. 2, 7) von der Universität St. Gallen untersuchen den Zusammenhang von verschiedenen Eigenschaften des VR auf die Unternehmensperformance, gemessen als Überrendite gegenüber dem SPI. Ein Kriterium darunter ist das Doppelmandat im Unternehmen. Mittels einer multiplen Regression erhalten sie einen leicht positiven und auf dem 10% Niveau signifikanten Einfluss des Doppelmandats auf die Performance (Ammann et al. (2005, S. 3), S. 14). Sie betonen daher in ihrem Schlusswort, dass ein allgemeines Verbot der Personalunion für Unternehmen in der Schweiz aufgrund ihrer Ergebnisse nicht gerechtfertigt wäre.

Eine sehr ausführliche Studie über den Zusammenhang zwischen CEO Duality und der Unternehmensbewertung in der Schweiz liefern Schmid und Zimmermann (2008, S. 182) von den Universitäten St. Gallen und Basel. Sie untersuchen 152 Schweizer Unternehmen zwischen 1993 und 2002. Sie finden keine signifikanten Unterschiede in der Bewertung, gemessen mit der Kennzahl Tobin's Q, für Unternehmen mit oder ohne CEO Duality. Im Gegensatz zu den Ergebnissen von Ammann et al. (2005, S. 3) sind die Koeffizienten negativ, jedoch nicht signifikant (Schmid und Zimmermann (2008), S. 197). Auch Schmid und Zimmermann (2008, S. 202) schlussfolgern, dass ihre Resultate keine eindeutige regulatorische Handlung unterstützen. Ebenfalls erwähnen sie die Tatsache, dass ihr Sample noch nicht oder nur marginal am Ende durch den 2002 eingeführten Swiss Code beeinflusst wurde (Schmid und Zimmermann (2008), S. 190).

Amstutz<sup>12</sup> (2007, S. 282) widmet sich in seinem Buch "Macht und Ohnmacht des Aktionärs" ausführlich der Frage der Organisation der Führung im Unternehmen. Da sich die Empirie sowohl international wie auch in der Schweiz uneinig ist, empfiehlt er stattdessen, anderweitige Überlegungen anzustellen (Amstutz (2007), S. 159). Die erhöhte Notwendigkeit eines umfassenden Risk-Managements, die hohe Belastung und Verantwortung für den VRP und den CEO sowie die veränderte Mentalität der Aktionäre sind seiner Meinung nach die Hauptargumente für eine Doppelspitze.

Einen völlig anderen Ansatz wählen Probst und Raisch (2005, S. 90) von der Universität Genf, die die 100 grössten Unternehmenskrisen der Jahre 2000 bis 2005 aus organisatorischer Sichtweise untersuchen. Sie finden heraus, dass extrem mächtige Manager sehr häufig in diesen Krisen vertreten waren und empfehlen daher, den Verwaltungsrat mit genügend unabhängigen Mitgliedern zu versehen, um das System der Checks and Balances jederzeit sicherzustellen (Probst und Raisch (2005), S. 102).

Nach dieser Übersicht über die bisherige Forschung in der Schweiz soll auf die verschiedenen Regulierungen und die diversen Richtlinien im schweizerischen Umfeld eingegangen werden.

Regulierung und Richtlinien bezüglich CEO Duality

# Obligationenrecht

Das Obligationenrecht<sup>13</sup> gilt als die rechtliche Grundlage für die Aktiengesellschaften in der Schweiz. Relevant sind dabei vor allem die Artikel 620 bis 763 OR im 26. Titel der dritten Abteilung für die Handelsgesellschaften und die Genossenschaft. Bezüglich der Thematik CEO Duality macht das Obligationenrecht jedoch keinerlei Vorschriften, da mit dem Art. 716 OR lediglich die Möglichkeit der Delegierung der Geschäftsleitung geschaffen wurde.

Swiss Code of Best Practice for Corporate Governance Der Swiss Code of Best Practice for Corporate Governance von der Economiesuisse wurde bereits im Kapitel 2.1.2 vorgestellt. In der ersten Version von 2002 befasste sich der Artikel 18 mit der Frage der Personalunion oder Doppelspitze. Dabei wurde im ersten Absatz die folgende neutrale Formulierung gewählt (Hofstetter (2002), S. 69):

> "Der Verwaltungsrat legt fest, ob sein Vorsitz und die Spitze der Geschäftsleitung (Delegierter des Verwaltungsrats, Geschäftsleitungsvorsitzender oder "CEO") einer Person (Personalunion) oder zwei Personen (Doppelspitze) anvertraut werden."

Der zweite Absatz weist darauf hin, dass bei der Wahl der Personalunion gewisse Kontrollmechanismen umgesetzt werden müssen, beispielsweise die Ernennung von einem Lead Director (Hofstetter (2002), S. 69). Begründet wird die gewählte Formulierung gemäss Hofstetter (2002 S. 43) mit der Organisationsfreiheit der Unternehmen und der Tatsache, dass auch Vorteile einer Personalunion erkennbar sind.

Der revidierte Swiss Code enthält nun aber mit der Ziffer 19 eine Präferenz für die Doppelspitze. Der erste Absatz lautet folgendermassen (Economiesuisse (2014), S. 12):

"Der Verwaltungsrat wirkt darauf hin, dass sein Vorsitz und die Spitze der Geschäftsleitung zwei Personen anvertraut werden (Doppelspitze)."

Der zweite Absatz bleibt jedoch unverändert. Somit ist die Wahl der Organisationsform gemäss dem Swiss Code noch immer den Unternehmen überlassen, auch wenn die Empfehlung vonseiten Economiesuisse in Richtung Trennung zeigt.

Richtlinie betreffend Informationen zur Corporate Governance

Praktisch zeitgleich zum Swiss Code wurde 2002 die Richtlinie betreffend Informationen zur Corporate Governance von der Schweizer Börse SWX, heute SIX Swiss Exchange, veröffentlicht (SIX Swiss Exchange (2016)). Diese Richtlinie gilt für alle an der SIX Swiss Exchange kotierten Unternehmen und hat zum Ziel, die Transparenz betreffend Corporate Governance zu erhöhen. Dabei muss unter dem Gliederungspunkt 3.1 "Mitglieder des Verwaltungsrats" jeweils angegeben werden, ob das Mitglied exekutiv tätig ist oder nicht (SIX Swiss Exchange (2016), S. 4). Weiter muss unter Punkt

 $<sup>^{12}</sup>$ Langjähriger CEO und Verwaltungsratspräsident von Unternehmen wie Holcim, Alusuisse-Lonza, Von Roll Holding und SGS.

 $<sup>\</sup>overline{\ \ }^{13}$  Insbesondere die darin enthaltenen Abteilungen zum Gesellschaftsrecht.

3.6 die Kompetenzregelung zwischen dem Verwaltungsrat und der Geschäftsleitung beschrieben werden. Weitergehende Anordnungen oder Empfehlungen bezüglich CEO Duality macht die Richtlinie jedoch keine.

# Bankengesetz

Die einzige konkrete Regelung in der Schweiz betreffend CEO Duality macht das Bankengesetz. Art. 3 Abs. 2a BankG fordert für die Erlangung der Bankenbewilligung seit 1991 eine getrennte Struktur für die Organe der Geschäftsführung und des Verwaltungsrats, welcher die Oberleitung und die Aufsicht innehat. Somit ist eine Personalunion in der Schweiz einzig bei Banken gesetzlich nicht möglich. Amstutz (2007, S. 158) merkt an, dass bei Banken das möglicherweise höhere Risiko einer Personalunion also anerkannt wird.

# 2.3. Zwischenfazit und Hypothesendefinition

# 2.3.1. Zwischenfazit

In diesem Kapitel sollen die bisherigen Erkenntnisse zur CEO Duality im internationalen sowie schweizerischen Umfeld zusammengeführt und ein Zwischenfazit gezogen werden. Dann werden zwei gegensätzliche Hypothesen basierend auf den beiden vorgestellten Denkschulen, Agency Theory und Stewardship Theory, definiert. Anschliessend werden diese Hypothesen im dritten Kapitel getestet und die Ergebnisse der Analysen werden im Kontext der bisherigen Forschung besprochen.

Ähnlich wie bei den zahlreichen internationalen Studien können auch die wenigen Studien mit Unternehmen aus der Schweiz keine einheitliche und abschliessende Antwort geben, ob sich CEO Duality positiv oder negativ auf die Bewertung von Unternehmen auswirkt. Die Frage, ob die Aktionäre solche Strukturen auf der obersten Führungsebene der Unternehmen akzeptieren, bleibt also relativ offen. Schmid und Zimmermann (2008) und Ammann et al. (2005, S. 3) erhalten zwar teilweise signifikante Resultate, jedoch erstere mit negativen und letztere mit positiven Vorzeichen. Solche Inkonsistenzen in den Ergebnissen führen dazu, dass die Leistungsträger in der Wirtschaft und Politik die möglichen Auswirkungen von Unternehmensentscheidungen und neuen Regulierungen und Gesetzen nur sehr schwer einschätzen können.

Mit der vorliegenden Arbeit wird mit einer im Gegensatz zu anderen Studien längeren und vor allem aktuellen Zeitperiode versucht, die Auswirkungen von CEO Duality auf die Bewertung von Unternehmen in der Schweiz zu analysieren. Dabei wird untersucht, inwiefern sich die Bewertungen der Unternehmen durch die Aktionäre unterscheiden. Somit lässt sich ermitteln, ob die Aktionäre solche Praktiken gutheissen und das Unternehmen höher bewertet ist oder ob sie die Unternehmen über einen tieferen Unternehmenswert abstrafen. Dazu werden nun im folgenden Kapitel zwei Hypothesen aufgestellt.

# 2.3.2. Hypothesendefinition *Hypothese 1: Agency Theory*

Die Vertreter der Agency Theory sind kritisch gegenüber der CEO Duality eingestellt, da eine solche Struktur eine effektive Überwachung des Managements durch den Verwaltungsrat verhindert. Somit erwartet diese Denkschule einen negativen Zusammenhang zwischen dem Vorliegen von CEO Duality und der Unternehmensbewertung. Dies bedeutet, dass die Aktionäre eine Personalunion an der Spitze des Unternehmens nicht begrüssen und das Unternehmen somit tiefer bewertet ist am Kapitalmarkt. Wenn ein Unternehmen also von einer Doppelspitze zu einer Personalunion wechselt, dann sollte dies die Bewertung verringern. Die erste Hypothese lautet folgendermassen:

H1: CEO Duality hat einen negativen Einfluss auf die Bewertung von Schweizer Unternehmen.

# Hypothese 2: Stewardship Theory

Die zweite Denkschule, die Stewardship Theory, hat eine gegensätzliche Auffassung bezüglich der Thematik CEO Duality als die Agency Theory. Ihre Vertreter fokussieren sich auf die Vorteile einer starken Führung im Unternehmen mit einer einzigen Person als CEO und Verwaltungsratspräsidenten gleichzeitig. Somit erwartet diese Denkschule einen positiven Zusammenhang zwischen dem Vorliegen von CEO Duality und der Unternehmensbewertung. Die Aktionäre begrüssen in diesem Fall eine Personalunion an der Spitze des Unternehmens und dies führt zu einer höheren Bewertung am Kapitalmarkt. Wenn ein Unternehmen also von einer Doppelspitze zu einer Personalunion wechselt, dann sollte sich der Unternehmenswert aufgrund des positiven Effekts erhöhen. Die zweite Hypothese lautet folgendermassen:

H2: CEO Duality hat einen positiven Einfluss auf die Bewertung von Schweizer Unternehmen.

# 3. Empirie

# 3.1. Variablen und Datensatz

# 3.1.1. Definition der Variablen

In der vorliegenden Arbeit wird der Einfluss von CEO Duality auf die Bewertung von Unternehmen in der Schweiz untersucht. Die unabhängige Variable von Interesse, der Regressor im Regressionsmodell, ist somit die Variable CEO Duality, im Modell CEO2 genannt. CEO2 ist als eine binäre Variable, also eine sogenannte Dummyvariable, mit den Ausprägungen 1 und 0 codiert. Wenn für ein bestimmtes Unternehmen i im Jahr t die gleiche Person als Chief Executive Officer und als Verwaltungsratspräsident im Amt ist, dann nimmt die CEO Duality Variable CEO2it den Wert 1 an, ansonsten 0. Diese Codierung ist in der Forschung bezüglich CEO Duality weit verbreitet (vgl. Schmid und Zimmermann (2008); Ammann et al. (2005, S. 3); Rechner und Dalton (1991); Boyd (1995); Finkelstein und D'aveni (1994)). Konstellationen, wo ein Verwaltungsratspräsident beispielsweise auf ad interim Basis den CEO-Posten übernimmt, werden in dieser Arbeit für das betreffende Jahr als CEO2 = 1 codiert. sobald er für mindestens 9 Monate eines Jahres (≥ 75%)

beide Posten innehat. Eine kürzere Dauer der Personalunion wird entsprechend als CEO2 = 0 codiert.

Die abhängige Variable, der Regressand im Regressionsmodell, ist in dieser Analyse eine Kennzahl für die Unternehmensbewertung. Hierzu wird die betriebswirtschaftliche Grösse Tobin's Quotient, kurz Tobin's Q, verwendet. Tobin's Q ist definiert als Quotient aus Marktwert und Wiederbeschaffungskosten eines Unternehmens (Chung und Pruitt (1994), S. 70). Aufgrund der komplizierten Berechnung schlagen Chung und Pruitt (1994, S. 70) eine Approximation von Tobin's Q vor, die mit grundlegenden Angaben aus den Finanzberichten der Unternehmen berechnet werden kann. Dabei berechnet sich Tobin's Q als Quotient aus der Marktkapitalisierung, den totalen Schulden (Fremdkapital) und dem Wert der ausstehenden Vorzugsaktien zum Buchwert aller Aktiven (Chung und Pruitt (1994), S. 71). Die gleiche Approximation nutzen auch Schmid und Zimmermann (2008, S: 187). Die Interpretation von Tobin's Q basiert auf der Annahme, dass die beiden Werte Marktwert und Wiederbeschaffungswert langfristig den gleichen Wert annehmen sollten und das Unternehmen ansonsten entweder über- oder unterbewertet ist (Bloomberg (2017)). Bloomberg (2017) bezeichnet Tobin's O als eine nützliche Kennzahl für die Unternehmensbewertung und berechnet ihn ebenfalls mittels der genannten Approximation. Obwohl Ammann et al. (2005, S. 11) eine Performancegrösse und nicht eine klassische Bewertungsgrösse in ihrer Studie verwenden, thematisieren sie die Kennzahl Tobin's Q. Sie erklären, dass der Tobin's Q neben der Verwendung als Kennzahl für die Unternehmensbewertung oft auch als Mass für die Performance gewählt wird, da er die Wertsteigerung misst, die beispielsweise durch eine gute Unternehmensführung resultiert.

Neben der unabhängigen Variable CEO2 und der abhängigen Variable TOBQ werden weitere Kennzahlen als Kontrollvariablen in das Modell eingegliedert. Diese Kontrollvariablen kontrollieren den Einfluss weiterer Faktoren wie beispielsweise der Grösse oder der Profitabilität eines Unternehmens auf die Bewertung. Sie werden dabei als weitere Regressoren in das multivariate Regressionsmodell eingefügt und erlauben somit, den reinen Einfluss von CEO Duality auf die Unternehmensbewertung zu messen, während die restlichen Kontrollvariablen konstant gehalten werden (Stock und Watson (2003), S. 179).

Die erste Kontrollvariable Leverage beschreibt das Finanzierungsverhältnis eines Unternehmens, oft auch als Verschuldungsgrad oder Kapitalstruktur bezeichnet. Volkart und Wagner (2014, S. 162) definieren das Finanzierungsverhältnis als Verhältnis aus Fremdkapital zum Eigenkapital. Dieses Verhältnis, der Leverage-Faktor, ist insofern relevant, da die Höhe des Fremdkapitals einen Einfluss auf die Höhe der Eigenkapitalrentabilität (Return on Equity, ROE) hat: Je höher der Leverage-Faktor, desto höher ist der ROE, solange die Gesamtkapitalrendite grösser als der Fremdkapitalkostensatz ist (Volkart und Wagner (2014), S. 609).

Die Kontrollvariable ROA bezeichnet die Profitabilitätskennzahl Return on Assets. Diese Kennzahl gibt an, wie profitabel und effizient ein Unternehmen in Prozent der Gesamtaktiven wirtschaftet und seine Mittel einsetzt (Bloomberg (2017)). Bloomberg (2017) berechnet diesen Indikator als Verhältnis aus Netto-Ergebnis und Gesamtaktiven, also der gesamten Bilanzsumme. Diese Kennzahl wird sehr häufig als Kontrollvariable für die kurzfristige Performance eines Unternehmens gebraucht (Finkelstein und D'aveni (1994), S. 1091).

Die Kontrollvariable LogSize soll für die Grösse des Unternehmens kontrollieren. Als Kennzahl wird die Bilanzsumme verwendet, die Bloomberg (2017) als Summe aller kurzfristigen und langfristigen Aktiven definiert und die jeweils in Millionen im jährlichen Finanzteil des Geschäftsberichts angegeben wird. Der natürliche Logarithmus davon ergibt die Kontrollvariable LogSize in einer zu anderen Kontrollvariablen vergleichbaren Grössenordnung. Die Kontrollvariable LogSales wird ebenfalls häufig als Kontrolle für die Unternehmensgrösse ausgewählt (Baliga et al. (1996), S. 49). Bloomberg (2017) berechnet den jährlichen Umsatz in Millionen als Summe der vier Quartalszahlen, Halbjahreszahlen oder Jahreszahlen aus der Erfolgsrechnung. Auch diese Grösse wird mit dem natürlichen Logarithmus transformiert. Eine weitere gebräuchliche Kennzahl für die Unternehmensgrösse wäre die Mitarbeiteranzahl, die jedoch meist hoch korreliert ist mit den Variablen Umsatz und Gesamtaktiven (Finkelstein und D'aveni (1994), S. 1092).

Zusammengefasst besteht das grundlegende Modell der vorliegenden Arbeit aus den sechs Variablen, die in Abbildung 2 in einer Übersicht dargestellt sind. Zudem sind alle Variablen im Anhang 1 beschrieben und definiert.

# 3.1.2. Sample und Datensatz

Als Ausgangslage für das Sample von Schweizer Unternehmen galt die Zusammensetzung des Swiss Performance Index per Ende Januar 2017. Zu diesem Zeitpunkt bestand der breit gefasste SPI aus 205 Unternehmen. Die anschliessend eigenständige Erhebung der Daten zur CEO Duality über den Zeitraum von 2000 bis 2015 erfolgte anhand der öffentlichen Geschäftsberichte sowie Pressemitteilungen. Das Sample wurde infolge fehlender Daten, erst im Jahre 2016 stattfindender Börsengänge sowie zweifacher Indexberücksichtigung um elf Unternehmen verringert, wodurch letztendlich 194 Unternehmen im Sample verbleiben, die gesamthaft betrachtet wurden<sup>14</sup>.

Die 194 Unternehmen würden über den Zeitraum von 16 Jahren 3104 einzelne Beobachtungen ergeben. Da sich die Zusammensetzung des SPI jedoch ständig ändert, weil Unternehmen aus dem SPI ausscheiden oder neu dazukommen, verringert sich die Anzahl effektiver Unternehmen pro Jahr im Sample. Dies führt dazu, dass die 194 Unternehmen mit

<sup>&</sup>lt;sup>14</sup>Die folgenden elf Unternehmen wurden nicht berücksichtigt (Grund in Klammern): Accu (fehlende Daten), Private Equity (Outsourcing-Management), Molecular Partners (fehlende Daten), Investis (IPO 2016), Kuros Biosciences (IPO 2016 nach Fusion), Lindt & Sprüngli (doppelte Nennung mit Partizipationsschein), Schindler (doppelte Nennung), Swissmetal (in Liquidation), Swatch Group (doppelte Nennung), VAT Group (IPO 2016), WISeKey International Holding (IPO 2016).

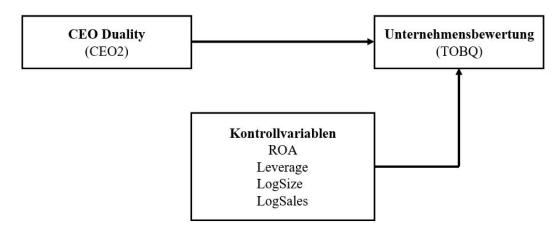


Abbildung 2: Übersicht der Modellvariablen

teils sehr unterschiedlichen Zeiträumen im Sample vertreten sind. Eine detaillierte Auflistung mit den betrachteten Unternehmen und ihrer jeweiligen Beobachtungsdauer ist der R im Anhang 2 zu entnehmen.

Die Abbildung 3 auf der folgenden Seite zeigt die Anzahl pro Jahr effektiv berücksichtigten Unternehmen. Diese Zahl reicht von 132 betrachteten Unternehmen im Jahr 2000 bis zu 193 Unternehmen im Jahr 2015. Über alle Jahre resultiert ein Mittelwert von 163 Unternehmen im Sample. Gesamthaft enthält das Sample über die 16 Jahre somit 2615 einzelne Beobachtungen. Die benötigten Daten für Tobin's Q und die Kontrollvariablen stammen aus Bloomberg. Die unabhängige Variable CEO2 konnte für alle Beobachtungszeitpunkte komplett eruiert und gesammelt werden.

Aufgrund der Struktur der Daten mit 194 Unternehmen als Untersuchungseinheiten und 16 Jahren als Zeiteinheiten handelt es sich um einen sogenannten Paneldatensatz. Dieser ist im "long format" strukturiert und beinhaltet somit pro Zeile eine Unternehmensmessung für ein bestimmtes Jahr (Giesselmann und Windzio (2012), S. 18). Da es ausserdem nicht für jedes Unternehmen die kompletten 16 Jahresbeobachtungen gibt, liegt ein unbalanciertes Panel vor (Giesselmann und Windzio (2012), S. 26).

# 3.2. Methodik und Resultate

# 3.2.1. Deskriptive Statistik

# CEO Duality

Für die unabhängige Variable von Interesse, die binäre Variable für CEO Duality, konnten für die 194 Unternehmen über 16 Jahre insgesamt 2615 Beobachtungen gesammelt werden. In 332 Beobachtungen davon liegt CEO Duality vor. Dies entspricht einem Gesamtanteil von 12.7% für alle Unternehmensbeobachtungen der Jahre 2000 bis 2015. In den restlichen 87.3% der Beobachtungen sind somit die Funktionen CEO und Verwaltungsratspräsident personell getrennt. Die Grössenordnung dieser Werte ist für den Zeitraum nach der Einführung des Swiss Codes in der Schweiz realistisch. Schmid und Zimmermann (2008, S. 60) nennen für die vorangehende Zeitperiode von 1993 bis 2002 Werte für CEO

Duality zwischen knapp 20% bis 25%. In Abbildung 4 ist die prozentuale Entwicklung der CEO Duality in der Schweiz zwischen 2000 und 2015 mit den vorliegenden Daten dargestellt:

Klar ersichtlich wird der abnehmende Trend von CEO Duality in der Schweiz. Während im Jahre 2000 noch 19.7% aller betrachteten Unternehmen eine Personalunion an der Spitze hatten, sind es 2015 gerade noch 8.29%. Der Trend ist in erster Linie auf verschärfte Vorschriften und Regulierungen (Swiss Code), sich für Aktionärsrechte einsetzende Gruppen wie die Ethos Stiftung und den stärkeren Druck zu einer Good Corporate Governance seit der Jahrtausendwende zurückzuführen. Eine interessante kurzfristige Erhöhung zeigt sich in den Jahren 2009 und 2010 während und nach der globalen Finanzkrise.

# Modellvariablen

Für die abhängige Variable Tobin's Q und die Kontrollvariablen resultieren aufgrund fehlender Daten zwischen 2593 und 2614 Beobachtungen für die 194 Unternehmen. Die Tabelle 1 fasst die deskriptive Statistik der Modellvariablen zusammen:

Aufgrund der sehr unterschiedlichen Unternehmen im SPI resultieren bei den Modellvariablen grosse Differenzen. Der Mittelwert für Tobin's Q, die Kennzahl der Unternehmensbewertung, liegt bei 1.6314, während die einzelnen Werte zwischen 0.4578 und 18.5955 stark schwanken. Ebenfalls sehr beachtliche Unterschiede, erkenntlich an der hohen Standardabweichung, zeigen sich beim Return on Assets mit einem Mittelwert von knapp 3% und beim Leverage, dem Verhältnis von Fremdkapital zum Eigenkapital, mit einem Mittelwert von 4.07.

Um negative logarithmierte Werte für Unternehmen mit Umsätzen (Sales) von weniger als einer Million CHF zu vermeiden, wurden bei LogSales die Umsätze in Tausenden anstatt in Millionen gemessen. Size hingegen stellt die gesamten Aktiven dar, die in Millionen gemessen werden. Dadurch entstehen die Differenzen zwischen LogSize und LogSales. Jedoch bestehen sowohl innerhalb LogSales als auch LogSize erhebliche Grössendifferenzen zwischen den Unternehmen.

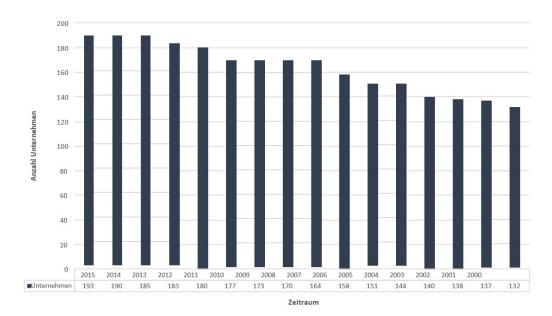


Abbildung 3: Anzahl Unternehmen pro Jahr im Sample

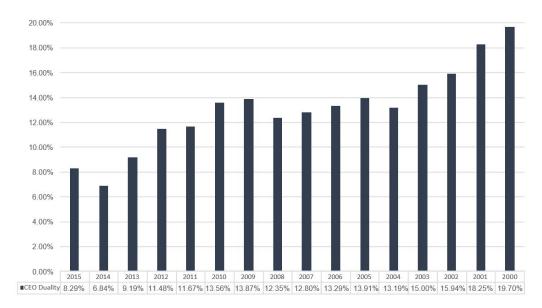


Abbildung 4: CEO Duality in der Schweiz zwischen 2000 und 2015

Tabelle 1: Deskriptive Statistik der Modellvariablen

	# Beob.	MW	SA	MIN	MAX
Tobin's Q	2605	16.314	12.222	0.4578	185.955
ROA	2606	29.718	116.484	-1.220.569	1.121.824
Leverage	2602	40.697	78.643	0.0123	2.081.923
LogSales	2593	57.341	0.9009	0.4771	81.475
LogSize	2614	31.142	0.9680	0.1300	63.796
CEO2	2615	0.1270	0.3330	0	1

Beispielsweise hatte die UBS 2006 eine Bilanzsumme von 2.3965 Billionen CHF (UBS (2007), S. 66). Ihr gegenüber stehen sehr kleine Unternehmen mit Bilanzsummen von wenigen Millionen CHF.

Die Tabelle 2 zeigt die Korrelationen zwischen den einzelnen Modellvariablen. Eine sehr hohe positive Korrelation zeigt sich wie erwartet zwischen LogSales und LogSize. Aus diesem Grund werden diese Kontrollvariablen in den Re-

Tabelle 2: Korrelationstabelle der Modellvariablen

	TOBQ	ROA	Leverage	LogSales	LogSize	CEO2
TOBQ	1.0000					
ROA	0.1054	1.0000				
Leverage	-0.1349	-0.1782	1.0000			
LogSales	-0.0607	0.3233	0.1291	1.0000		
LogSize	-0.1923	0.1602	0.4474	0.7824	1.0000	
CEO2	0.0255	-0.0242	-0.0797	0.0052	-0.0702	1.0000

gressionsanalysen besonders beachtet und auch einzeln verwendet. Erfreulich sind die sehr tiefen Korrelationswerte der unabhängigen Variable CEO2 mit den Kontrollvariablen. Es ergeben sich leicht positive oder negative Werte, jedoch alle sehr nahe um 0. Einen relativ schwach positiven Zusammenhang ergibt sich zwischen der Performancegrösse ROA und der Bewertungsgrösse Tobin's Q. Zwischen Tobin's Q und den restlichen drei Kontrollvariablen, Leverage, LogSales und LogSize, zeigen sich negative Zusammenhänge.

# Zusammenhang CEO Duality und Tobin's Q

Um einen ersten Eindruck über den Zusammenhang zwischen der Variable CEO Duality und der Unternehmensbewertungskennzahl Tobin's Q zu erhalten, werden die Datenpunkte in einem Streudiagramm dargestellt. Da es sich bei CEO Duality um eine Dummyvariable mit den Ausprägungen 0 und 1 handelt, ergibt sich die in Abbildung 5 dargestellte Form.

Die durch die Datenpunkte gezogene, rote Gerade hat einen leicht positiven Verlauf. Diese Steigung zeigt sich auch in der Korrelationstabelle mit dem Wert 0.0255. Diese leicht positive Steigung würde bedeuten, dass ein Unternehmen mit einer Personalunion (CEO2 = 1) eine leicht höhere Unternehmensbewertung aufweist. Es zeigt sich jedoch, dass dieser Korrelationswert mit einem P-Wert von 0.1938 statistisch insignifikant ist. Um den Einfluss von CEO Duality auf die Unternehmensbewertung analysieren zu können, wird im folgenden Kapitel die induktive Statistik behandelt und deren Methodik und Resultate vorgestellt.

# 3.2.2. Induktive Statistik

# Methodik: Analyse von Paneldaten

Werden mehrere Untersuchungseinheiten über mehr als zwei Untersuchungszeitpunkte analysiert, dann liegt ein sogenannter Paneldatensatz vor (Stock und Watson (2003), S. 11). Der Index i bezeichnet dabei jeweils die Untersuchungseinheit, in diesem Fall das einzelne Unternehmen. Der Index t zeigt den Untersuchungszeitpunkt an. Wie bereits in Kapitel 3.1.2 erwähnt, handelt es sich im vorliegenden Sample um ein unbalanciertes Panel, da die Anzahl Beobachtungen kleiner als  $t \times i$  ist. Weiter ist das Sample so aufgestellt, dass es eine "large i, small t"-Struktur hat, also viele Unternehmen zu einer kleineren Anzahl an Untersuchungszeitpunkten beobachtet werden (Giesselmann und Windzio (2012), S. 11).

Die Analyse von Paneldaten hat den grossen Vorteil, dass für Merkmale, die unbeobachtet und unternehmensspezifisch sind, kontrolliert werden kann (Giesselmann und Windzio (2012), S. 10). Beispielsweise können dies Unterschiede in der Geschäftstätigkeit der Unternehmen oder auch kulturelle Faktoren sein. Somit kann für diese individuelle Heterogenität mit relativ einfachen Methoden kontrolliert werden (Torres-Reyna (2007)). Wichtig zu betonen ist die Tatsache, dass es sich dabei um eine Längsschnittfragestellung handelt. Dies bedeutet, dass die Auswirkung einer Veränderung auf Unternehmensebene relevant ist (Giesselmann und Windzio (2012), S. 33). Die Veränderung der Variable CEO Duality stellt einen solchen intraindividuellen Effekt dar. So lässt sich der tatsächliche Effekt einer Veränderung in der Führungsstruktur eines Unternehmens auf die Unternehmensbewertung zurückführen.

Die beiden am häufigsten verwendeten Modelle zur Analyse von Paneldaten sind das Fixed-Effects-Modell und das Random-Effects-Modell (Wooldridge (2010), S. 286). Das FE-Verfahren subtrahiert von jedem Merkmal den Mittelwert und eliminiert so Niveauunterschiede (Giesselmann und Windzio (2012), S. 39). Gemäss Torres-Reyna (2007) wird das Verfahren vor allem angewendet, wenn der Effekt von einer zeitvarianten Variable gemessen werden soll. Die Annahme dahinter ist, dass der Fehlerterm mit den unabhängigen Variablen korreliert und somit den Einfluss auf die abhängige Variable verzerrt oder beeinflusst wird (Wooldridge (2010), S. 286). Das Random-Effects-Modell basiert auf der Prämisse, dass es einen besseren Schätzer für den Effekt gibt, als das FE-Verfahren vorschlägt (Giesselmann und Windzio (2012), S. 80). Es liegt der Annahme zugrunde, dass der Fehlerterm nicht mit den unabhängigen Variablen korreliert und die Variation somit zufällig ist (Wooldridge (2010), S. 286). Während das FE-Verfahren oder eine Variation davon meistens bei Längsschnittfragestellungen angewendet wird, wird das RE-Verfahren häufig bei Querschnittsfragestellungen angewendet (Giesselmann und Windzio (2012), S. 108).

Da es sich in der vorliegenden Fragestellung um eine Längsschnittfragestellung handelt und der Effekt einer zeitvarianten Variable gemessen werden soll, muss mit grosser Wahrscheinlichkeit ein Fixed-Effects-Verfahren angewendet werden. Zur Überprüfung, ob FE oder RE angewendet werden muss, kann der Hausman-Test durchgeführt werden (Giesselmann und Windzio (2012), S. 109). Der Test überprüft, ob die Fehlerterme mit den Regressoren korreliert sind und gibt für die Alternativhypothese somit das Fixed-Effects-Modell als geeignete Wahl an, da in diesem Fall eine

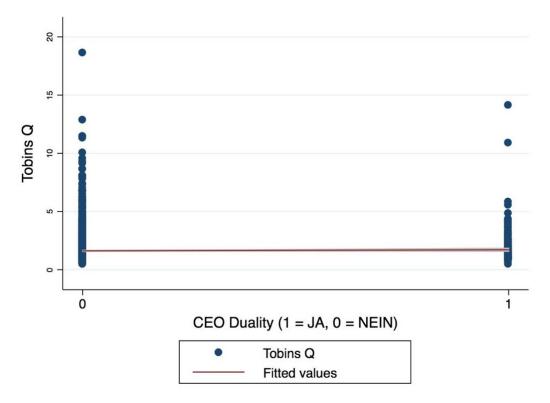


Abbildung 5: Streudiagramm CEO Duality & Tobin's Q

Korrelation vorliegt (Wooldridge (2010), S. 328). Der durchgeführte Test ist im Anhang 3 in Tabelle 6 zu finden. Der hoch signifikante P-Wert führt dazu, dass die Nullhypothese verworfen werden kann und somit ein FE-Modell angewendet werden soll. Ein weiterer Test auf Zeiteffekte<sup>15</sup> (Anhang 4, Tabelle 7) sollte durchgeführt werden, wenn ein FE-Modell verwendet wird. Er zeigt an, dass im vorliegenden Fall für Zeiteffekte kontrolliert werden muss. Somit wird für unbeobachtete Effekte der Unternehmen kontrolliert, die über die Zeitperioden variieren können (Stock und Watson (2003), S. 370).

Die beiden Tests führen dazu, dass in der vorliegenden Arbeit ein sogenanntes "time and entity fixed effects model" aufgestellt werden muss. Dieses Modell kann erstellt werden, indem binäre Variablen für i-1 Unternehmen und t-1 Jahre in das Modell integriert werden (Stock und Watson (2003), S. 370; Torres-Reyna (2007)). Das resultierende Modell (genauer spezifiziert in Anhang 5) lautet folgendermassen:

$$\begin{aligned} & Tobins\,Q_{i,t} = \beta_0 + \beta_1 CEO2_{i,t} + \beta_2 Leverage_{i,t} + \\ & \beta_3 ROA_{i,t} + \beta_4 LogSales_{i,t} + \beta_5 LogSize_{i,t} + \gamma_2 E_2 + \\ & \ldots + \gamma_n E_n + \delta_2 T_2 + \ldots + \delta_n T_t + u_{i,t} \end{aligned}$$

Einfaches Regressionsmodell Nachdem im Kapitel 3.2.1.3 der Zusammenhang zwischen CEO Duality und Tobin's Q grafisch aufgezeigt wurde, wird nun im ersten Schritt der induktiven Analyse eine einfache Regressionsanalyse mit dem vorliegenden Sample durchgeführt. Das Regressionsmodell in diesem Kapitel besteht nur aus der abhängigen Variable TOBQ und der unabhängigen Variable von Interesse CEO2. Damit soll ein erster Eindruck für die Richtung des Vorzeichens des Koeffizienten gewonnen werden. Es werden insgesamt drei Modelle aufgestellt. Der Fokus liegt auf dem Fixed-Effects-Modell mit der Kontrolle für Zeiteffekte (Modell 2). Die Resultate des einfachen Regressionsmodells sind in Tabelle 3 dargestellt. Bei den Werten in den Klammern handelt es sich um den jeweiligen Standardfehler des Schätzers.

Alle Vorzeichen der drei Modelle, (1) Fixed-Effects ohne Zeiteffekte, (2) FE mit Zeiteffekten und (3) Random-Effects, sind sowohl mit normalen wie auch mit robusten Standardfehlern negativ. Wird beim FE-Modell für Zeiteffekte kontrolliert, dann wird sowohl der Koeffizient wie auch der Standardfehler betragsmässig etwas kleiner. Auch das dritte Modell mit RE liefert sehr ähnliche Resultate wie das FE-Modell. Die Koeffizienten mit normalem Standardfehler sind statistisch signifikant auf dem 5% und 10% Signifikanzniveau. Falls jedoch robuste Standardfehler verwendet werden, die die Unternehmenscluster im Sample beachten und für Heteroskedastizität kontrollieren, sind die Koeffizienten statistisch nicht mehr signifikant. In den nachfolgenden Analysen werden ausschliesslich robuste Standardfehler verwendet. Dies ist relevant, da somit berücksichtigt werden kann, dass jedes Unternehmen mehrere Messungen zu unterschied-

<sup>&</sup>lt;sup>15</sup>Die Nullhypothese in diesem Test ist die Joint-Hypothese, dass die Zeit-Dummies für alle Untersuchungsjahre 0 ergeben. Falls die Nullhypothese verworfen werden kann, muss für Zeiteffekte kontrolliert werden (Torres-Reyna (2007)).

**Tabelle 3:** Einfaches Regressionsmodell

Anmerkung: Diese Tabelle zeigt die Schätzungen verschiedener Regressionen von Tobin's Q auf CEO Duality. Der Modelltyp ist entweder Fixed-Effects (FE) oder Random-Effects (RE). Jahr-Dummies werden zur Kontrolle von Endogenität hinsichtlich der Zeit in das Modell integriert. Der Wert in Klammern stellt den normalen (I) oder robusten (II) Standardfehler (SF) dar. \*\*\*, \*\* und \* beschreiben die Signifikanz der Schätzer auf dem 1, 5 oder 10 Prozentniveau.

		Tobin's Q	
	(1)	(2)	(3)
(I) CEO2 (normaler SF)	-0.2023**	-0.1551*	-0.1623**
	(0.0868)	(0.0852)	(0.0821)
(II) CEO2 (robuster SF)	-0.2023	-0.1551	-0.1623
	(0.1559)	(0.1540)	(0.1352)
Modelltyp	FE	FE	RE
Jahr-Dummies	NEIN	JA	NEIN
N	2605	2605	2605
R2	0.001	0.0256	0.0006
Rho	0.5558	0.5724	0.5312

lichen Beobachtungszeitpunkten besitzt (Giesselmann und Windzio (2012), S. 77). Wie in der Tabelle 3 erkennbar ist, erhöhen sich dadurch lediglich die Standardfehler, während die Koeffizienten identisch bleiben. Das sehr tiefe R2 bei allen Modellen widerspiegelt die schwache Aussagekraft des einfachen Regressionsmodells, da sehr viele weitere Faktoren die Bewertung eines Unternehmens beeinflussen.

Wie können aber die negativen Koeffizienten interpretiert werden? Das negative Vorzeichen vor CEO2 bedeutet, dass Unternehmen mit einer CEO Duality Struktur eine um den Koeffizienten tiefere Bewertung haben als die Unternehmen mit einer getrennten Unternehmensspitze. Beim Modell (2) beispielsweise hat ein Unternehmen mit CEO Duality einen um 0.1551 tieferen Unternehmenswert, gemessen am Tobin's Q, als ein Unternehmen ohne CEO Duality. Dieser negative Zusammenhang ergibt eine schwache Zustimmung für die erste Hypothese und somit für die Agency Theory.

# Multiples Regressionsmodell

In diesem zweiten Schritt wird eine multiple Regression von Tobin's Q auf die unabhängige Variable CEO Duality durchgeführt. Dabei werden die in Kapitel 3.1.1 beschriebenen Kontrollvariablen in das Modell integriert, um für spezifische Eigenschaften und Merkmale der Unternehmen kontrollieren zu können. Ebenfalls werden wieder verschiedene Modelle aufgestellt. Die Resultate der multiplen Regressionsanalyse sind in Tabelle 4 dargestellt. Zu beachten ist, dass bei allen fünf Modellen robuste Standardfehler verwendet werden, um unverzerrte Schätzer zu erhalten.

Besonders auffallend ist der konstant negative Zusammenhang zwischen CEO Duality und Tobin's Q. Die Resultate aus der einfachen Regressionsanalyse scheinen sich somit zu bestätigen. Die FE-Modelle (1) und (2) werden mit allen vier Kontrollvariablen gerechnet, einmal jedoch mit und einmal ohne Jahr-Dummies. Analog zur einfachen Regression nimmt die Höhe des Koeffizienten leicht ab, wenn für Zeiteffekte kontrolliert wird. Die FE-Modelle (3) und (4) enthalten nur noch eine der beiden Kontrollvariablen für die Grösse des Unternehmens, LogSize oder LogSales. Das Modell (5) ist ein Random-Effects-Modell mit der Variable LogSize als Kontrol-

le für die Grösse. Es dient als Alternativ- und Vergleichsmodell zu den FE-Modellen, die aufgrund des Hausman-Tests zu bevorzugen sind.

Wird das Modell (1) betrachtet, welches gemäss den durchgeführten Tests zu verwenden ist, zeigt sich ein auf dem 10%-Niveau statistisch signifikanter und negativer Einfluss von CEO Duality auf die Bewertung der Unternehmen. Der Tobin's Q ist um 0.1959 kleiner, falls es sich um ein Unternehmen mit CEO Duality handelt. Dieser Effekt wird im Modell (3) betragsmässig gesehen leicht grösser, wenn nur LogSize verwendet wird. Jedoch ist der Schätzer dann knapp nicht mehr signifikant. Analog verhält es sich mit der Signifikanz im Modell (4), falls LogSales anstatt LogSize verwendet wird. Das RE-Modell (5) liefert ein ähnliches Ergebnis, welches auf dem 10%-Niveau signifikant ist. Noch immer sehr gering ist die Aussagekraft der Modelle mit einem R2 zwischen 1.21% und 4.71%. Dies lässt sich darauf zurückzuführen, dass die Unternehmensbewertung durch die Aktionäre von weitaus mehr Faktoren abhängt, als nur von den hier verwendeten Kennzahlen, die ausserdem vor allem buchhalterisch geprägt sind.

Beachtlich sind jedoch die hohen Werte zwischen 54.82% und 79.46% für die Kennzahl Rho, welche den Anteil der gesamten Varianz darstellt, welcher aufgrund Unterschieden innerhalb der Unternehmen resultiert. Rho wird dabei häufig auch als Intra-Klassen-Korrelation bezeichnet (Torrey-Reyna, 2007). Besonders relevant ist hier jedoch nicht die absolute Höhe der Kennzahl, sondern die relative Veränderung von Rho zwischen den einzelnen Modellen. Die Werte bleiben für die Modelle (1) bis (4) aber relativ konstant und mit über 70% vergleichsweise hoch, was typisch für Fixed-Effects-Modelle ist. Das Rho für das RE-Modell ist mit 54.82% deutlich tiefer.

Die Richtungsweisen der Kontrollvariablen sind wie erwartet positiv für den Leverage und den ROA. Dies bedeutet, dass ein höheres Finanzierungsverhältnis und eine höhere Gesamtrentabilität mit einer höheren Unternehmensbewertung einhergehen. Diese Schätzer sind zudem bis auf den ROA-Schätzer im Modell (3) und (5) höchst signifikant. Log-

**Tabelle 4:** Multiples Regressionsmodell

Anmerkung: Diese Tabelle zeigt die Schätzungen verschiedener Regressionen von Tobin's Q auf CEO Duality. Der Modelltyp ist entweder Fixed-Effects (FE) oder Random-Effects (RE). Jahr-Dummies werden zur Kontrolle von Endogenität hinsichtlich der Zeit in das Modell integriert. Der Wert in Klammern stellt den robusten Standardfehler dar. \*\*\*, \*\* und \* beschreiben die Signifikanz auf dem 1, 5 oder 10 Prozentniveau.

			Tobin's Q		
	(1)	(2)	(3)	(4)	(5)
CEO2	-0.1959*	-0.2793**	-0.2116	-0.1649	-0.2537*
	(0.1146)	(0.1128)	(0.1419)	(0.1115)	(0.1514)
Leverage	0.0142***	0.0140***	0.0131***	0.0129***	0.0131***
	(0.0028)	(0.0036)	(0.0034)	(0.0031)	(0.0043)
ROA	0.0143***	0.0167***	0.0084	0.0141**	0.0088
	(0.0054)	(0.0058)	(0.0061)	(0.0056)	(0.0073)
LogSales	-0.4128	-0.4347	-	-0.8788***	-
	(0.2670)	(0.2925)		(0.1948)	
LogSize	-1.0593***	-0.7837***	-1.5487***	-	-0.6748***
	(0.3262)	(0.3333)	(0.2917)		(0.5162)
Modelltyp	FE	FE	FE	FE	RE
Jahr-Dummies	JA	NEIN	JA	JA	NEIN
N	2572	2572	2586	2572	2586
R2	0.0350	0.0246	0.0471	0.0121	0.0372
Rho	0.7824	0.7144	0.7946	0.7029	0.5482

Sales und LogSize haben einen negativen Einfluss auf die Bewertung. Dies lässt sich damit erklären, dass die Grösse eines Unternehmens ab einem gewissen Punkt den Unternehmenswert aus Sicht der Aktionäre nicht mehr steigern kann und somit senken muss.

Grundsätzlich bestätigt das multiple Regressionsmodell die ersten Ergebnisse aus der einfachen Regression und verstärkt somit die Evidenz für die erste Hypothese der Agency Theory. Sie besagt, dass das Vorhandensein von CEO Duality einen negativen Einfluss auf die Unternehmensbewertung haben könnte, weil die Aktionäre der Meinung sind, dass das Management bei einer CEO Duality nicht genügend kontrolliert wird. Somit bewerten sie das Unternehmen am Kapitalmarkt tiefer und strafen es für ihre Führungsstruktur ab. Es stellt sich die offensichtliche Frage, ob dies für alle Unternehmen gilt oder möglicherweise industrieabhängig ist. Daher wird in den weiteren Analysen eine Subsample-Analyse durchgeführt, bei welcher die Finanzinstitute von den Nicht-Finanzinstituten getrennt werden. Ausserdem wird eine alternative Grösse als abhängige Variable verwendet, um die Ergebnisse weiter testen zu können.

Weiterführende Analysen

# Subsample-Analyse ohne Finanzinstitute

In dieser ersten weiterführenden Analyse soll die multiple Regressionsanalyse mit einem angepassten Sample durchgeführt werden. Da es für Banken seit 1991 gemäss dem Bankengesetz in der Schweiz gesetzlich verboten ist, eine CEO Duality Struktur an der Unternehmensspitze zu installieren, macht es Sinn, das Sample um die Finanzinstitute zu reduzieren, da diese Unternehmen über keine Entscheidungsfreiheit bezüglich ihrer Führungsstruktur verfügen. Ausserdem werden nicht nur Banken exkludiert, sondern allgemein Finan-

zinstitute gemäss dem Global Industry Classification Standard (GICS-Code 40). Dies erfolgt aus dem Grund, dass diese Unternehmen oftmals sehr unterschiedlichen Regulierungen unterworfen sind und eine grundlegend andere Finanzierungsstruktur haben.

Im Gesamtsample fallen von den 194 Unternehmen 44 Unternehmen oder 22.7% in diese Kategorie "Financials". Die verbleibenden 150 "Non-Financial" Unternehmen stellen das Sub-Sample 1 der ersten weiterführenden Analyse dar. Der gesamte Regressionsoutput ist in Tabelle 8 im Anhang 6 zu finden. Die Ergebnisse, verglichen mit den Resultaten vom Gesamtsample, ändern sich nur leicht. Alle Wirkungsrichtungen und Vorzeichen bleiben unverändert. Das Hinzufügen von Jahr-Dummies im FE-Modell führt dazu, dass der Koeffizient betragsmässig leicht kleiner, also weniger negativ wird. Zudem ist der Schätzer dann mit einem P-Wert von 0.16 statistisch nicht mehr signifikant.

Zusammengefasst ändern sich die Resultate für den Zusammenhang zwischen CEO Duality und der Unternehmensbewertung nicht, wenn nur Nicht-Finanzinstitute berücksichtigt werden. Daraus kann gefolgert werden, dass das Gesamtsample keine verzerrten Ergebnisse liefert, weil Unternehmen berücksichtigt werden, die per Gesetz keine CEO Duality haben dürfen.

Analyse mit ROA als abhängige Variable

In dieser zweiten weiterführenden Analyse sollen die bisher erhaltenen Ergebnisse aus dem Kapitel 3.2.2.3 mit einer alternativen Kennzahl als abhängige Modellvariable validiert werden. In der Literatur wird häufig der Return on Assets als Kennzahl verwendet, wenn Zusammenhänge zwischen Themen der Corporate Governance und der Unternehmensleistung untersucht werden. Dabei ist zu beachten, dass es sich beim ROA nicht wie beim Tobin's Q um eine typische Kenn-

zahl für die Unternehmensbewertung handelt, sondern um eine buchhalterische Rentabilitätskennzahl, die die Performance eines Unternehmens misst.

Die Ergebnisse zur multiplen Regression von ROA auf CEO Duality sind der Tabelle 9 im Anhang 7 zu entnehmen. Überraschenderweise zeigt sich ein positiver Zusammenhang zwischen dem ROA und CEO Duality. Dieser Wechsel des Vorzeichens gegenüber den Analysen mit Tobin's Q ist bei allen drei gezeigten Modellen konstant zu finden. Die Koeffizienten bleiben ebenfalls in der gleichen Grössenordnung, wenn LogSales anstatt LogSize verwendet wird und falls nur das Subsample ohne Finanzinstitute berücksichtigt wird.

Dieser positive Zusammenhang zwischen CEO Duality und ROA bedeutet, dass Unternehmen mit CEO Duality einen um 1.48% (Modell 1) höheren Return on Assets haben als Unternehmen ohne CEO Duality. Das Vorhandensein einer starken Führungsperson an der Spitze eines Unternehmens scheint die Performance des Unternehmens positiv zu beeinflussen. Jedoch sind die jeweiligen CEO2-Koeffizienten mit P-Werten von 0.15, 0.11 und 0.18 statistisch nicht signifikant.

Eine mögliche Erklärung für das unterschiedliche Vorzeichen liegt in der Art der Kennzahl. Während Tobin's Q als Kennzahl für den Unternehmenswert oder den Shareholder Value betrachtet wird, ist der ROA eine reine Profitabilitätsgrösse aus dem finanziellen Rechnungswesen. Aktionäre, die eine CEO Duality ablehnen, können die Unternehmen abstrafen, indem sie ihre Aktien verkaufen. Dadurch kann ein tieferer Unternehmenswert resultieren, wenn dies viele Aktionäre gleichzeitig oder im grossen Rahmen tun. Die Auswirkung von CEO Duality auf die operative Leistungsfähigkeit hingegen zeigt sich vor allem in Kennzahlen wie dem ROA. Eine starke Führung als grosser Vorteil von CEO Duality, wie dies die Stewardship Theory oft vertritt, kann sich somit positiv auf die Performance des Unternehmens auswirken. Wie nun diese unterschiedlichen Ergebnisse im Rahmen der Diskussion über CEO Duality zu verstehen und einzuordnen sind, wird im nächsten Kapitel 3.2.3 aufgezeigt.

# 3.2.3. Diskussion und Einordnung der Resultate

Die erhaltenen Resultate für den Zusammenhang zwischen CEO Duality und der Unternehmensbewertung aus den Regressionsanalysen sprechen für die erste Hypothese aus Kapitel 2.3.2.1. Der negative Zusammenhang zeigt eine gewisse Evidenz für die Agency Theory, welche eine Trennung der beiden Funktionen Chief Executive Officer und Verwaltungsratspräsident vorschlägt. Amstutz (2007, S. 165) vertritt sogar die Meinung, dass eine Personalunion nur in einer Krisensituation und zeitlich beschränkt eingeführt werden darf. Dies könnte beispielsweise den bereits erwähnten Anstieg der CEO Duality in der Schweiz während und kurz nach der Finanzkrise in Abbildung 4 im Kapitel 3.2.1.1 erklären.

Die Resultate stehen besonders national in Einklang mit der Corporate Governance Forschung zur CEO Duality. Analog zu Schmid und Zimmermann (2008), die ebenfalls Tobin's Q als Kennzahl verwenden, sind die Koeffizienten negativ. Da sie jedoch keine signifikanten Ergebnisse erhalten, schlussfolgern sie, dass es keine empirische Evidenz für systematische Unterschiede der Bewertung gibt (Schmid und Zimmermann (2008), S. 182). Das Sample in der vorliegenden Arbeit besteht einerseits aus einer grösseren Anzahl an Unternehmen und andererseits aus einem deutlich längeren Beobachtungszeitraum. Dieser aktuelle Zeitraum von 2000 bis 2015 berücksichtigt somit die neueren Entwicklungen der Corporate Governance in der Schweiz, hauptsächlich beeinflusst durch den Swiss Code und dem Ruf nach einer Good Corporate Governance. Die Aktionäre scheinen das Vorliegen einer CEO Duality weniger zu tolerieren, da sie Unternehmen mit einer Personalunion tiefer bewerten. International betrachtet stehen die Resultate vor allem im Einklang mit Rechner und Dalton (1991) sowie Pi und Timme (1993), die mit ihren Studien ebenfalls Evidenz für die Agency Theory finden.

Die Resultate aus der zweiten weiterführenden Analyse mit dem ROA als abhängige Variable und als Performancekennzahl erscheinen auf den ersten Blick widersprüchlich zu den Resultaten aus der Hauptanalyse. Allerdings finden auch Ammann et al. (2005, S. 3) einen positiven Zusammenhang zwischen CEO Duality und der Performance von Unternehmen in der Schweiz. Während in der vorliegenden Analyse der ROA als Performancemass verwendet wird, arbeiten Ammann et al. (2005, S. 3) in ihrer Studie mit Überrenditen gegenüber der Rendite des SPI. Da die Resultate der vorliegenden Arbeit jedoch statistisch nicht signifikant sind, müssen solche Interpretationen mit grosser Vorsicht behandelt werden. Der positive Zusammenhang deutet aber eine gewisse Zustimmung für die Stewardship Theory an, die als Alternative zur gängigen Agency Theory betrachtet wurde. Im Sinne dieser Denkschule führt eine CEO Duality zu einer effizienten und effektiven Führung des Unternehmens, die sich in einer höheren operativen Leistungsfähigkeit des Unternehmens widerspiegeln kann. Eine höhere Performance müsste sich somit in Performancekennzahlen wie dem ROA zeigen.

Allgemein sind jedoch empirische Analysen im Rahmen der Corporate Governance Forschung aufgrund ihrer limitierten Aussagekraft eher vorsichtig zu beurteilen (Larcker und Tayan (2016), S. 14). Ammann et al. (2005, S. 11) betonen dabei, dass die Unternehmensperformance von zahlreichen und oft schwer messbaren Faktoren abhängt. Das gleiche Problem besitzen auch Kennzahlen der Unternehmensbewertung. Ebenfalls zu berücksichtigen ist die Tatsache, dass der Zusammenhang auch umgekehrt möglich wäre. Finkelstein und D'aveni (1994, S. 1088) merken an, dass die Unternehmensperformance auch direkt auf die CEO Duality wirken könnte. Unternehmen mit einer guten Performance in der Vergangenheit sind möglicherweise eher legitimiert, eine CEO Duality Struktur einzuführen als Unternehmen mit einer schlechten Performance. Weiter sind die Aktionäre eines Unternehmens, das eine schlechte Leistung zeigt, wohl eher in einer stärkeren Verhandlungsposition, um beispielsweise eine Personalunion zu verhindern. Aus diesem Grund ist auch die Kontrolle für die Profitabilität anhand der Kontrollvariable ROA sehr wichtig im Hauptmodell. Zweifelsohne besteht für diesen Zusammenhang und Gedankengang grossen Bedarf, weiterführende Forschung und Analysen durchzuführen

# 4. Schlussfolgerungen

# 4.1. Zusammenfassung

Die vorliegende Arbeit liefert eine quantitative Untersuchung der Auswirkung von CEO Duality auf die Bewertung von Unternehmen in der Schweiz. Damit soll einerseits empirisch gezeigt werden, ob die Empfehlungen im Swiss Code of Best Practice for Corporate Governance präzise genug formuliert sind und ob die heutige Präferenz für eine Trennung der beiden Funktionen CEO und Verwaltungsratspräsident gerechtfertigt ist. Andererseits sollen Empfehlungen für die Politik und Wirtschaft erstellt werden können, da die bisherige Forschung auf diesem Gebiet sowohl international wie auch vor allem in der Schweiz umstritten und inkonsistent ist. Nach einer Aufarbeitung der bisherigen Forschung und Literatur zur CEO Duality, liegt der Fokus im zweiten Teil der Arbeit auf der Analyse des Zusammenhangs zwischen CEO Duality und Unternehmensbewertung für 194 Schweizer Publikumsgesellschaften aus dem sehr breit gefassten Swiss Performance Index. Die Analyse erfolgt über 16 Jahre und basiert auf einem handgesammelten Datensatz über das Vorhandensein von CEO Duality in der Schweiz.

Die zwei aufgestellten Hypothesen basieren auf den beiden vorherrschenden Denkschulen in der Literatur, die Agency Theory und die Stewardship Theory. Erstere postuliert einen negativen Zusammenhang zwischen CEO Duality und der Unternehmensbewertung und letztere einen positiven Einfluss. Eine multiple Regressionsanalyse des Paneldatensatzes mit diversen Modellen führt zu konstant negativen Koeffizienten für den Einfluss von CEO Duality auf die Bewertung. Diese Koeffizienten sind teilweise statistisch signifikant und bestätigen somit die erste Hypothese und die Vertreter der Agency Theory, die eine Trennung der beiden Funktionen an der Unternehmensspitze bevorzugen.

Die Ergebnisse bestätigen somit die Änderung der Formulierung im Swiss Code. Seit 2014 beinhaltet die Empfehlung unter der Ziffer 19 eine klare Präferenz für die Trennung und somit für eine Doppelspitze, während vorher eine neutrale Formulierung beide Systeme gleichbehandelte (Economiesuisse (2014), S. 12). Nichtsdestotrotz lässt der Swiss Code eine Personalunion unter gewissen Umständen immer noch zu. Ausser bei den Banken gibt es bisher keine Gesetze oder Regulierungen in der Schweiz, die das Doppelmandat verbieten. Dies hängt damit zusammen, dass die Frage der Ausgestaltung der obersten Führungsstruktur auch immer stark vom Einzelfall und somit von den betreffenden Personen abhängig ist (von der Crone (2001), S. 29). Amstutz (2007, S. 164) ist der Meinung, dass eine strikte gesetzliche Vorgabe der Doppelspitze und somit ein Verbot von CEO Duality nicht zielführend wäre. Er forderte die Einführung der Doppelspitze als Regel, jedoch mit der Möglichkeit für die Unternehmen, die Wahl einer Personalunion angemessen zu erklären. Genau dies wurde 2014 von der Economiesuisse auch umgesetzt.

Ein weiteres Indiz gegen die Einführung einer gesetzlichen Pflicht der Doppelspitze liefern die Ergebnisse der weiterführenden Analyse aus der vorliegenden Arbeit. Es scheint einen schwach positiven Zusammenhang zwischen der CEO Duality und der Performance der Unternehmen zu geben. Unternehmen mit einer Personalunion scheinen einen leicht höheren ROA zu haben. Jedoch sind diese Ergebnisse statistisch nicht signifikant und somit mit grosser Vorsicht zu behandeln. Das beste Beispiel für ein Schweizer Unternehmen mit einer starken Performance während einem Doppelmandat ist Lindt & Sprüngli. Ernst Tanner, CEO und Verwaltungsratspräsident seit 1994, hat erst im Juni 2016 sein Doppelmandat niedergelegt und gleichzeitig versichert, als exekutiver VRP weiterhin sehr aktiv zu sein (Finanzen.ch (2016)). Eine solche Machtkonzentration an der Spitze eines grossen Unternehmens ist in solchem Umfang wahrscheinlich nur möglich, wenn die Performance gleichzeitig sehr hoch ist. Ein gesetzliches Verbot würde die Gestaltungsmöglichkeiten der Unternehmen stark einschränken, was wiederum nicht im Interesse der Politik und der Wirtschaft sein kann.

# 4.2. Fazit und Ausblick

Die Frage nach der optimalen Führungsstruktur von börsenkotierten Unternehmen ist und bleibt eine sehr spannende Thematik. Auch in Zukunft wird das Thema CEO Duality Investoren, Politiker und Manager beschäftigen und kontroverse Diskussionen anregen. Die momentane Rechts- sowie Regulierungsgrundlage kann im Lichte der erhaltenen Resultate aus der vorliegenden Arbeit als vernünftig beurteilt werden. Für ein allgemeines gesetzliches Verbot der Personalunion gibt es zu wenig Konsistenz in der Forschung. Zudem hat die CEO Duality durchaus auch diverse Vorteile in bestimmten Situation. Ausserdem gibt es häufig Konstellationen, in denen ein Doppelmandat stärker gerechtfertigt ist. Beispielsweise ist das Vorhandensein von CEO Duality bei einer unechten Publikumsgesellschaft mit einem institutionellen Investor oder einer Familie als grossen Mehrheitsaktionär eher legitimiert als bei einer Gesellschaft mit einem breitgestreuten Aktionariat. Das Prinzipal-Agenten-Problem reduziert sich oder verschwindet sogar ganz bei solchen Unternehmen. Nicht zu vergessen ist aber die Tatsache, dass das Bankengesetz mit dem Verbot des Doppelmandats für Banken in der Schweiz zwar als Ausnahme gilt, jedoch damit das möglicherweise höhere Risiko bei einer Machtakkumulation explizit berücksichtigt (Amstutz (2007), S. 158).

Zukünftige Analysen und Studien auf diesem Gebiet der Corporate Governance Forschung sollten sich neben der Frage der Legitimation von CEO Duality für bestimmte Arten der Publikumsgesellschaft auch Situationen widmen, die einer CEO Duality nahekommen. Wenn ein langjähriger CEO beispielsweise seinen Posten abgibt und dafür Präsident vom gleichen Unternehmen wird, dann liegt de facto eine ähnliche Situation vor wie bei einer typischen Personalunion (Amstutz (2007), S. 147). Solche Konstellationen sind in der Schweiz sehr häufig zu finden, wenn die Geschäftsberichte der Unternehmen genau analysiert werden. Lindt & Sprüngli ist auch hier ein gutes Beispiel. Als neuer CEO folgte im Juni

2016 Dieter Weisskopf, welcher seit 1995 als CFO bei Lindt & Sprüngli amtete und gemäss Ernst Tanner in circa fünf Jahren idealerweise das Verwaltungsratspräsidium übernehmen wird (Finanzen.ch (2016)).

In der vorliegenden Analyse wurden lediglich vier Kontrollvariablen in das Modell integriert. Zudem erfolgt eine Subsample-Analyse anhand der Separierung in Finanzinstitute und Nicht-Finanzinstitute. Rein technisch macht es für weiterführende und ausführlichere Analysen Sinn, zusätzliche Kontrollvariablen in das Modell aufzunehmen, um beispielsweise für Industriezugehörigkeit und Governance-Variablen kontrollieren zu können. Dies würde den Erklärungsgehalt des Modells nochmals erhöhen.

Die Analyse der Auswirkung von CEO Duality auf die Bewertung von Unternehmen in der Schweiz hat gezeigt, dass die Entscheidung über die Führungsstruktur weitreichende Implikationen für die Unternehmen und ihre Aktionäre hat. Sie muss daher situationsabhängig und vor allem äussert sorgfältig getroffen werden, um negative Auswirkungen auf den Unternehmenswert verhindern oder verringern zu können.

### Literatur

- Ammann, M., Leuenberger, M., von Wyss, H., et al. Eigenschaften von Verwaltungsräten und Unternehmensperformance. Swiss Journal of Economics and Statistics (SJES), 141(I):1–22, 2005.
- Amstutz, M. D. Macht und Ohnmacht des Aktionärs: Möglichkeiten und Grenzen der Corporate Governance bei der Wahrung der Aktionärsinteressen. Neue Zürcher Zeitung, Zürich, 2007.
- Bainbridge, S. M. Corporate governance after the financial crisis. Oxford University Press, New York, 2012.
- Baliga, B. R., Moyer, R. C., und Rao, R. S. CEO duality and firm performance: What's the fuss? *Strategic Management Journal*, 17(1):41–53, 1996.
- Berle, A. und Means, G. The modern corporate and private property. *McMillian, New York, NY*, 1932.
- Bloomberg. Kennzahlendefinition und Daten zu abhängiger Variable Tobin's Q & Kontrollvariablen aus der Bloomberg Datenbank, 2017.
- Böckli, P., Huguenin, C., und Dessemontet, F. Expertenbericht der Arbeitsgruppe "Corporate Governance" zur Teilrevision des Aktienrechts. Zürich: Schulthess Verlag, 2004.
- Boyd, B. K. CEO duality and firm performance: A contingency model. *Strategic Management Journal*, 16(4):301–312, 1995.
- Brickley, J. A., Coles, J. L., und Jarrell, G. Leadership structure: Separating the CEO and chairman of the board. *Journal of Corporate Finance*, 3(3): 189–220, 1997.
- Chaganti, R. S., Mahajan, V., und Sharma, S. Corporate board size, composition and corporate failures in retailing industry [1]. *Journal of Management Studies*, 22(4):400–417, 1985.
- Chung, K. H. und Pruitt, S. W. A simple approximation of Tobin's Q. *Financial Management*, 23(3):70–74, 1994.
- Davis, J. H., Schoorman, F. D., und Donaldson, L. Toward a stewardship theory of management. Academy of Management Review, 22(1):20–47, 1997
- Dey, A., Engel, E., und Liu, X. CEO and board chair roles: To split or not to split? *Journal of Corporate Finance*, 17(5):1595–1618, 2011.
- Donaldson, L. The ethereal hand: Organizational economics and management theory. *Academy of Management Review*, 15(3):369–381, 1990.
- Donaldson, L. und Davis, J. H. Stewardship theory or agency theory: CEO governance and shareholder returns. Australian Journal of Management, 16(1):49–64, 1991.
- Economiesuisse. Swiss code of best practice for Corporate Governance. Zürich: Economiesuisse, 2014.
- Elsayed, K. Does CEO duality really affect corporate performance? Corporate Governance: An International Review, 15(6):1203–1214, 2007.
- Ethos Services. Ethos Studie 2016. Generalversammlungen, Vergütungen und Corporate Governance. Genf: Ethos Stiftung, 2016.
- Fama, E. F. und Jensen, M. C. Separation of ownership and control. *The Journal of Law and Economics*, 26(2):301–325, 1983.
- Finanzen.ch. Lindt & sprüngli: Ernst Tanner gibt Doppelmandat ab Finanzchef neuer CEO, 2016. URL http://www.finanzen.ch/nach richten/aktien/Lindt-&-Spruengli-Ernst-Tanner-gibt-Dopp elmandat-ab-Finanzchef-neuer-CEO-1001239645. Abgerufen am 05.04.2017.
- Finkelstein, S. und D'aveni, R. A. CEO duality as a double-edged sword: How boards of directors balance entrenchment avoidance and unity of command. Academy of Management Journal, 37(5):1079–1108, 1994.
- Giesselmann, M. und Windzio, M. Regressionsmodelle zur Analyse von Paneldaten, volume 1. Springer-Verlag, 2012.
- Hofstetter, K. Corporate Governance in der Schweiz Bericht im Zusammenhang mit den Arbeiten der Expertengruppe "Corporate Governance". Zürich: Economiesuisse, 2002.
- Jensen, M. C. The modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance*, 48(3):831–880, 1993.
- Jensen, M. C. und Meckling, W. H. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3 (4):305–360, 1976.
- Larcker, D. und Tayan, B. Corporate governance matters: A closer look at organizational choices and their consequences. New Jersey: Pearson Education, 2016.
- Lindt & Sprüngli. Geschäftsbericht 2015. Kilchberg, 2016.
- OECD. G20/OECD-Grundsätze der Corporate Governance. Paris: OECD Publishing, 2015.

- Pi, L. und Timme, S. G. Corporate control and bank efficiency. *Journal of Banking & Finance*, 17(2):515–530, 1993.
- Probst, G. und Raisch, S. Organizational crisis: The logic of failure. *The Academy of Management Executive*, 19(1):90–105, 2005.
- Rechner, P. L. und Dalton, D. R. The impact of CEO as board chairperson on corporate performance: Evidence vs. rhetoric. *The Academy of Management Executive*, 3(2):141–143, 1989.
- Rechner, P. L. und Dalton, D. R. CEO duality and organizational performance: A longitudinal analysis. *Strategic Management Journal*, 12(2):155–160, 1991.
- Ruigrok, W., Peck, S. I., und Keller, H. Board characteristics and involvement in strategic decision making: Evidence from swiss companies. *Journal of Management Studies*, 43(5):1201–1226, 2006.
- Schmid, M. M. und Zimmermann, H. Should chairman and CEO be separated? Leadership structure and firm performance in Switzerland. *Schmalenbach Business Review*, 60(2):182–204, 2008.
- SIX Swiss Exchange. Richtlinie betreffend Informationen zur Corporate Governance. Zürich: SIX Swiss Exchange, 2016.
- Stock, J. H. und Watson, M. W. Introduction to Econometrics (3. Aufl.). Boston: Addison Wesley, 2003.
- SWX Swiss Exchange. Kommentar zur Corporate Governance-Richtlinie. Zürich: SWX Swiss Exchange, 2007.
- Tonello, M. Separation of chair and CEO roles. The Harvard Law School Forum on Corporate Governance and Financial Regulation, 2011. URL https://corpgov.law.harvard.edu/2011/09/01/separation-of-chair-and-ceo-roles/. Abgerufen am 17.02.17.
- Torres-Reyna, O. Panel data analysis fixed and random effects using Stata (V. 4.2). Princeton University Data and Statistical Services, 2007. URL https://www.princeton.edu/~otorres/Panel101.pdf. Abgerufen am 08.03.17.
- UBS. Finanzbericht 2006. Zürich: UBS AG, 2007.
- Volkart, R. und Wagner, A. Corporate Finance Grundlagen von Finanzierung und Investition (6. Aufl.). Zürich: Versus, 2014.
- von der Crone, H. Wenn der Präsident auch CEO ist. Wie beurteilt Corporate Governance eine Personalunion? Neue Zürcher Zeitung, page 29, 2001.
- Wooldridge, J. M. Econometric analysis of cross section and panel data. The MIT press, Cambridge, 2 edition, 2010.



# **Junior Management Science**

journal homepage: www.jums.academy



# Die Einstellung kulturinteressierter Personen gegenüber Sponsoren kultureller Events – Eine qualitative Studie

# Maximilian Krebs

Ludwig-Maximilians-Universität München

### Abstract

In dieser Arbeit wird das Sponsoring von kulturellen High-Brow Veranstaltungen behandelt, sowie nach potenziellen Quellen negativer Gefühle der Besucher von kulturellen Veranstaltungen gegenüber Sponsoren dieser Events gesucht. Hier stehen insbesondere Personen mit hohem Interesse an kulturellen Aktivitäten im Fokus. Zur theoretischen Fundamentierung wird Sponsoring zunächst im Allgemeinen definiert und hinsichtlich seiner Ziele und Voraussetzungen erklärt. Verschiedene Theorien, wie der "Mere exposure" Effekt, oder die Kongruenz Theorie, geben einen Einblick in die Wirkungsweise eines Sponsoringengagements. Zugeschnitten auf das Kultursponsoring wird dessen positiver Einfluss auf ein Unternehmen anhand der Corporate Reputation dargelegt. Dafür soll zunächst die Corporate Reputation bezüglich ihrer Dimensionen erläutert werden. Daraufhin werden mögliche Unterschiede zwischen Kultur und Sportsponsoring aufgezeigt. Ausgehend von einer qualitativen Forschung können mögliche Quellen negativer Neigungen gegenüber Sponsoren eines kulturellen Events abgeleitet werden. Mithilfe problemzentrierter Interviews konnten acht Personen befragt werden. Neben der Aufmerksamkeit der Befragten gegenüber einem Sponsor, stehen außerdem der Fit zwischen Sponsor und Event, sowie die Befürchtung einer Einflussnahme und das Auftreten des Sponsors im Vordergrund. Die Ergebnisse wurden zur Veranschaulichung in einem Template zusammengefasst. Limitationen ergeben sich vor allem durch den qualitativen Charakter der Arbeit. Letztendlich lässt sich feststellen, dass trotz der negativen Stimmen über Kultursponsoring, das gesamte Stimmungsbild tendenziell als positiv einzuschätzen ist.

Keywords: Sponsoring, Kultursektor, Firmenimage, Werbewirkung

# 1. Kultursponsoring in Deutschland

Aufgrund der abnehmenden staatlichen Unterstützung müssen in Deutschland einige kulturelle Einrichtungen zur Finanzierung neben den staatlichen Fördergeldern auf alternative Einnahmequellen zurückgreifen (vgl. Rectanus (2002), S. 10). Dazu gehört die Unterstützung aus der privaten Wirtschaft. Hierunter fallen Sponsoren, welchen sich dadurch neue Kanäle des Unternehmensmarketings eröffnen. Die Strategie hinter beschriebenem Vorgang nennt sich Kultursponsoring. Auch wenn der Markt für diese spezielle Art des Sponsorings im Vergleich zum Sportsponsoring verhältnismäßig klein ist und in Deutschland keine signifikant steigenden Wachstumszahlen ersichtlich sind (vgl. Repucom (2015)), stellt er für einige kulturelle Einrichtungen einen wichtigen Teil des finanziellen Budgets dar (vgl. Causales (2013)). Doch auch für den Sponsor bietet das Engagement entscheidende Vorteile. Vergangene Forschungen konnten den positiven Effekt von Sponsoring auf die Corporate Reputation, einer Variable zur Messung des guten Rufs eines Unternehmens, bestätigen (vgl. Schwaiger et al. (2010), S. 85). Dabei fiel allerdings auf, dass kulturinteressierte Personen das Engagement weniger positiv sehen als andere Personen.

Diese Arbeit versucht im Folgenden herauszufinden, worin diese negative Einstellung der Besucher kultureller Veranstaltungen hinsichtlich des Kultursponsorings begründet liegen könnte. Nachdem zu Beginn kurz die Vorteile des Sponsorings für den Gesponserten aufgezeigt werden, wendet sie sich im Anschluss dem Kultursponsoring aus Sicht der Sponsoren zu. Zunächst wird Sponsoring im Allgemeinen definiert sowie mögliche Ziele und Voraussetzungen für deren Erreichen dargelegt. Um die Wirkungsweise von Sponsoring zu erklären, werden verschiedene theoretische Modelle hinzugezogen. Daraufhin wird der Einfluss von Kultursponsoring auf ein Unternehmen anhand der Variable Corporate Reputation aufgezeigt. Nach der bereits angesprochenen Forschung von Schwaiger et al. (2010) hat Kultursponsoring einen positiven Einfluss auf die Corporate Reputation.

Dieser Einfluss wird näher erörtert und die Corporate Reputation in ihren Bestandteilen erklärt. Ausgehend von den Erkenntnissen von Schwaiger et al. (2010) wird in einer eigenen Forschung nach möglichen Quellen negativer Gefühle gegenüber Sponsoren kultureller Events gesucht. Interviews auf qualitativer Basis sollen helfen, Stimmungen und Meinungsbilder von Personen mit hohem Interesse an kulturellen Aktivitäten aufzuzeigen.

# 2. Kultursponsoring - Chance für Gesponserte

Wie bereits erwähnt, müssen einige kulturelle Einrichtungen in Deutschland für ihre Finanzierung auf nichtstaatliche Mittel zurückgreifen. Neben Crowdfunding und eigenen Erlösen spielt das Thema Sponsoring eine wachsende Rolle (vgl. Causales (2013)). Veranschaulicht wird das durch folgende Aussage der Nürnberger Symphoniker über ihre Sponsoren: "Die Arbeit unseres Orchesters und das umfangreiche Angebot, das wir Ihnen als Konzertbesucher bieten können, wäre nicht möglich ohne die institutionelle Förderung des Freistaates Bayern, der Stadt Nürnberg und des Bezirks Mittelfranken. Unser besonderer Dank gilt darüber hinaus unserem Hauptsponsor, der Sparkasse Nürnberg, die uns dabei großzügig finanziell unterstützt." (Nürnberger Symphoniker (2015)). Das Zitat verdeutlicht die Notwendigkeit, neben staatlichen Mitteln auf Unterstützung aus der Wirtschaft zurückzugreifen. Durch diese Unterstützung ist die Institution im Stande namhafte Künstler zu verpflichten, das Event, die Ausstellung, oder ihr Programm zu vergrößern, beziehungsweise die Infrastruktur des Events / der Institution zu verbessern. Dementsprechend ist zu folgern, dass in finanziellen Mitteln der Hauptgrund für kulturelle Einrichtungen liegt, sich einen Sponsoringpartner zu suchen.

Neben finanziellen Mitteln bietet das Sponsoring noch weitere Vorteile für den Gesponserten. So profitieren kulturelle Einrichtungen unter Umständen von organisatorischem und betriebswirtschaftlichem Know-how des Sponsors. Diese Dienstleistungen können entscheidend zum Stattfinden und zum Erfolg insbesondere eines speziellen Events beitragen. Auch mit Sachleistungen unterstützen Unternehmen zum Beispiel Museen. So fand im Kunstmuseum Bochum 2011 eine Ausstellung von Exponaten aus Kunstsammlungen statt, welche ausschließlich unter den Besitztümern von im Ruhrgebiet ansässigen Firmen zu finden sind (vgl. Ruhr.2010 GmbH (2011)).

Einen wichtigen Punkt für den Sponsor stellt, wie in einem späteren Teil der Arbeit noch dargelegt, die Verbreitung des Engagements über die klassischen Medien dar (vgl. Meenaghan (1991), S. 41). Doch ebenso für den Gesponserten birgt das Vorteile. Marketingaktivitäten des Sponsoringpartners könnten eigene Anstrengungen das Event, oder die Institution stärker in den Fokus der Öffentlichkeit zu rücken komplementieren. Dadurch lässt sich unter Umständen ein größeres Publikum generieren und damit verbunden zusätzlich das Ansehen des Events / der Institution erhöhen. Durch die dadurch erhöhten Besucherzahlen steigen ebenfalls die eigenen Einnahmen.

# 3. Kultursponsoring - Chance für Sponsoren

Das vorhergehende Kapitel hat bereits die Wichtigkeit und die Vorteile des Kultursponsorings für den Gesponserten aufgezeigt. Um im Folgenden auf die Chancen für den Sponsor einzugehen, wird zuerst das Kultursponsoring aus Sicht des Sponsors definiert und hinsichtlich seiner Ziele und Wirkungsweise erklärt. Anschließend können die positiven Auswirkungen für Sponsoren anhand der Variable Corporate Reputation analysiert werden.

# 3.1. Definition und konzeptionelle Einordnung des Kultursponsorings

Vor sieben Jahren beschrieb Cornwell (2008, S. 41) in ihrem Artikel "State of the Art and Science in Sponsorship-Linked Marketing" Sponsoring als eine der interessantesten Möglichkeiten, Unterhaltung und Werbung zu kombinieren. Der Begriff bezieht sich auf den Austausch finanzieller Zuwendungen eines Unternehmens gegen das Werbepotenzial einer Organisation oder eines Events. Das Unternehmen erhält dabei das Recht, seinen Namen mit der gesponserten Organisation oder dem Event darzustellen (vgl. O'hagan und Harvey (2000), S. 205). Cornwell (2008, S. 41) zählt Sponsoring, ebenso wie Product Placement (dt. Produktplatzierung) und Advergaming (dt. Werbespiele), zu einer neuen Zeit der Kommunikationsmethoden und bezeichnet diese als Ära des indirektes Marketings. Die Intention hinter dieser Bezeichnung ist deutlich: Anstatt den Konsumenten mit direkter Werbung zu penetrieren, können durch diese Methoden Unterhaltung und Marketing verknüpft und möglicherweise eine für den Konsumenten angenehmere Form von Werbung geschaffen werden. Mehrere Studien bestätigen diesen Eindruck. Nach Quester und Thompson (2001, S. 45) wird Sponsoring vom Konsumenten nicht nur als Werbung, sondern als Chance für den Gesponserten gesehen und daher nicht als so eigensinnig und aufdringlich wie klassische Werbung. Nach McDonald (1991, S. 35-36) sehen Konsumenten im ersten Moment hinter einem Sponsoring den guten Willen des Unternehmens etwas positives zu leisten.

Diese Arbeit konzentriert sich im Folgenden auf das Sponsoring von kulturellen Institutionen oder Events. Genauer auf die Unterstützung von "High-Brow" Kunst, welche beispielsweise neben Museen auch klassische Musik, Kunstaustellungen und Literatur mit einschließt (vgl. Schwaiger et al. (2010), S. 77). Zur Vereinfachung und falls nicht explizit anders erwähnt, verstehen sich die Begriffe Institution und Event im Laufe dieser Arbeit als Synonyme und schließen das jeweilig Nicht-Genannte mit ein. Außerdem beziehen sie sich auf oben genannte Kulturbereiche und erklären Sponsoring aus der Sicht der Sponsoren.

Ein klassisches Beispiel für Kultursponsoring ist das Engagement der Linde AG bei der bayerischen Staatsoper in München. Der Spielzeitpartner der Spielzeit 2015/2016 unterstützt die Oper nicht nur im Allgemeinen, sondern hilft bei der Finanzierung spezieller Projekte wie einem Livestream ausgewählter Veranstaltungen über das Internet (vgl. The

Linde Group (2013)). Die Vorteile für ihre Sponsoren beschreibt die Staatsoper wie folgt: "Das Partnerprogramm bietet die Möglichkeit, eine erlesene und selektierte Zielgruppe – rund 600.000 Besucher p.a. – auf nahezu konkurrenzlose, vornehme und wirkungsvolle Weise anzusprechen. Die Bayerische Staatsoper steht für höchste Qualität und Anspruch – ein Image von dem Unternehmen profitieren können." (Bayerische Staatsoper (2015)) Das Zitat lässt bereits auf mögliche Ziele und Potenziale von Kultursponsoring schließen, auf welche im nächsten Gliederungspunkt eingegangen wird.

# 3.2. Ziele und Voraussetzungen des Kultursponsorings

Nach Meenaghan (1991, S. 36) birgt das Sponsoring zwei essenzielle Potenziale: zum einen hinsichtlich der Zielgruppe des Gesponserten, beispielsweise die Besucher eines klassischen Musikfestivals zu bewerben; zum anderen das Potenzial sich mit dem Image des Gesponserten zu attribuieren. Aufbauend auf diesen beiden Möglichkeiten, arbeitet Meenaghan (1991, S. 40-41) die zwei Hauptziele von Aktivitäten im Bereich Sponsoring heraus. Neben Verbesserung der Bekanntheit und des Images einzelner Marken einer Firma setzen Betriebe Sponsoring ein, um selbige Ziele auf Unternehmensebene zu verfolgen.

Im Laufe dieser Arbeit wird vorwiegend auf die Effekte des Sponsorings auf das Image beziehungsweise auf die Reputation auf Unternehmensebene eingegangen. Neben internen Effekten, beispielsweise der erhöhten Moral von Mitarbeitern, besteht die Chance wichtige Entscheidungsträger aus Politik und Wirtschaft unterschwellig auf Sponsoringevents zu beeinflussen. Daneben vermag es Sponsoring das öffentliche Interesse an der kompletten Firma zu steigern. Neben der Interessenssteigerung besteht die Möglichkeit, je nach gesponserter Institution, auch Einfluss auf das Image der Firma zu nehmen (vgl. Meenaghan (1991), S. 40-41). Dafür ist die Auswahl der richtigen Art von Sponsoring von Bedeutung. Um erfolgreiches Sponsoring zu gewährleisten, ist es von Vorteil, unternehmensweite Regeln für die Wahl von geeigneten Kooperationspartnern festzulegen. Diese Regeln vermeiden Inkonsistenzen und gewährleisten eine Übereinstimmung zwischen den Werten der gesponserten Institutionen und der Unternehmensphilosophie in der öffentlichen Wahrnehmung (vgl. Meenaghan (1991), S. 41). Nach Meenaghan (1991, S. 42-43) sollten Sponsoringaktivitäten hinsichtlich der Zielgruppe drei Kriterien entsprechen: Die Demographie der Zielgruppe beeinflusst maßgeblich die Art des gesponserten Events beziehungsweise der Institution. Ein Hersteller für Videospiele ist beispielsweise dazu Angehalten bei der Auswahl eines geeigneten Events darauf zu achten, dass dieses Ereignis von seiner Zielgruppe besucht wird. Dazu kommt die geographische Abdeckung des Events. Regionale Unternehmen dürften sich stärker auf regionale Events konzentrieren, während multinationale Konzerne sich auf Events mit hoher geographischer Verbreitung ausrichten sollten. Das letzte Kriterium umfasst den Lifestyle des Events. Der Gesponserte muss ähnliche Werte wie der Sponsor vertreten, um, wie bereits erwähnt, keine Inkonsistenzen in der öffentlichen Wahrnehmung zu erzeugen. Cornwell et al. (2006, S. 216) fanden eine verbesserte Wirkung des Sponsorings heraus, wenn zum einen Sponsor und Gesponserter gleiche Werte vertreten und zum anderen aus einer ähnlichen Branche stammen. Zuletzt spielt auch das Auftreten des Sponsors auf dem gesponserten Event eine Rolle. Nach Meenaghan und Shipley (1999, S. 343) variiert der Spielraum wie stark ein Event von einem Sponsor für dessen Zwecke ausgenutzt werden darf. Hohes Potenzial ein Event kommerziell auszubeuten, ohne ein Ärgernis bei den Besuchern hervorzurufen bietet dabei eher das Sportsponsoring. Sponsoren im Bereich elitärer Kunst, oder sozialer Angelegenheiten müssen daher ihren Auftritt auf einem Event anpassen.

Einen wichtigen Punkt, welcher bis jetzt vernachlässigt wurde, stellt die Reichweite des Events in den Medien dar (in diesem Punkt liegt der Fokus auf Events, da diese zeitlich begrenzt sind und somit mehr im Fokus der alltäglichen Berichterstattung liegen, allerdings sollte genauso das Sponsoring von Institutionen medial verbreitet werden). Nach Meenaghan (1991, S. 41) versteht sich die Reichweite des Events in der medialen Berichterstattung für Unternehmen mit internationaler Zielgruppe unter Umständen als ein Hauptargument für Sponsoring. Doch auch ein Unternehmen, welches seine Produkte nur in einer bestimmten Region verkauft, ist im Stande von der Verbreitung seines Engagements auf regionaler Basis zu profitieren. Schwaiger et al. (2010, S. 80) setzen in ihrer Studie über Kultursponsoring ebenfalls eine Verbreitung des Sponsoringengagements durch klassische Medien voraus.

Nach einer Studie von O'hagan und Harvey (2000, S. 215-217) nutzen Unternehmen speziell Kultursponsoring zur Verbesserung ihres Images, sowohl auf regionaler als in gleicher Weise auf internationaler Ebene. Auch die Erhöhung der öffentlichen Wahrnehmung, sowohl über das Unternehmen als zusätzlich über bestimmte Marken, erweist sich als Grund für ein Engagement im kulturellen Bereich. Als ein klassisches Beispiel für Kultursponsoring dient das Engagement der Audi AG, welche als Hauptsponsor für die Salzburger Festspiele auftritt. Neben dem Vorteil des möglichen Gewinns an Reputation in der Öffentlichkeit nutzt das Unternehmen das Event, um wichtigen Geschäftspartnern, im Rahmen der "Audi Night", ein exklusives Erlebnis auf den Festspielen zu bieten. Dazu wird das Produkt Audi A8 als Shuttle für geladene Gäste eingesetzt und somit extra beworben. Des Weiteren lädt Audi die Salzburger Festspiele für ein Gastspiel jährlich zum eigenen Sommerfest nach Ingolstadt ein. Damit verfolgt die Aktiengesellschaft nicht nur in Salzburg und Umgebung, sondern ebenso am eigenen Unternehmensstandort regionale Ziele (vgl. Audi AG (2015)). An diesem Beispiel wird deutlich, dass sowohl die von Meenaghan (1991, S. 40-41) aufgestellten als auch die spezifisch für Kultursponsoring vorgestellten Ziele (vgl. O'hagan und Harvey (2000), S. 215-217) für die Praxis von Relevanz sind.

# 3.3. Wirkungsweise von Sponsoring

Die Literatur liefert verschiedene Ansätze hinsichtlich der Wirkung von Sponsoring. Nach Cornwell et al. (2005, S. 22) liegt der Mehrzahl der Sponsoringtheorien ein assoziatives Erinnerungsmodell von Keller (1993, S. 2) zu Grunde. Demnach setzt sich das Wissen über eine Marke (oder im Falle von Sponsoring über ein Unternehmen) aus dem Bewusstsein, bestehend aus Erinnerungen über die Marke, und dem Image, also allen Assoziationen mit der Marke, zusammen. Daher machen es sich die meisten Theorien über Sponsoring zur Aufgabe, diese Determinanten zu verbessern oder zu beeinflussen (vgl. Cornwell et al. (2005), S. 22-23). Im Folgenden werden die wichtigsten dieser Theorien vorgestellt. Die Theorien schließen sich keinesfalls gegenseitig aus, sondern vielmehr komplementieren sie sich unter einander.

Die erste Theorie der "Mere exposure" Effekt, also das bloße Auseinandersetzen von Personen mit einer Marke, wurde von Olson und Mathias Thjømøe (2003, S. 249-250) in einer Studie erarbeitet. Personen, welche in bloßen Kontakt mit einer Marke kamen, hatten daraufhin verbesserte Präferenzen gegenüber dieser Marke. Dieser Effekt trägt möglicherweise zur Wirkung von Sponsoring bei, da Personen auf einem Event einem dezenten Stimulus, beispielsweise dem Namen des Unternehmens, ausgesetzt werden und daraufhin dieses Unternehmen besser bewerten. Ähnlich sind die Erkenntnisse von Pham und Vanhuele (1997, S. 413-414), welche die Reaktivierung bestehender Assoziationen zu einer Marke postulieren, sobald Kontakt mit Sponsorenhinweisen, Titel oder Logos hergestellt wird.

Zwei weitere Theorien konzentrieren sich stärker auf das Verhältnis zwischen dem Image des Gesponserten und des Sponsoren. Die Kongruenz Theorie geht davon aus, dass sich Personen am besten Dinge merken, welche in Beziehung zueinander stehen oder ähnlich sind. Cornwell et al. (2006, S. 216) wiesen in einer Studie ein besseres Sponsoringergebnis für Unternehmen mit kongruentem Partner nach. Mit Hilfe eines Experiments verglichen sie den Recall auf das Sponsoring eines Musikfestivals. Das in der Unterhaltungsbranche tätige Unternehmen Sony erzielte dabei ein besseres Ergebnis als der Lebensmittelhersteller Heinz. Auch die Articulation Theorie bestätigte sich durch besagtes Experiment. Assoziierten sich die beiden Unternehmen mit Werten, welche dem des Musikfestivals ähnelten, realisierten sie ein besseres Resultat für das Sponsoring (vgl. Cornwell et al. (2006), S. 216).

Nach Cornwell et al. (2005, S. 28) spielen darüber hinaus sozialwissenschaftliche Theorien hinsichtlich des Sponsorings eine Rolle. Die Social Identity Theory beschreibt die Identifikation von Individuen mit verschiedenen sozialen Kategorien. Diese umfassen je nach Rahmen sehr viele, wie bei der Zugehörigkeit zu einer Nation, oder weniger Menschen beispielsweise bei Sportvereinen und auch bei der Zugehörigkeit zu einer Organisation (vgl. Tajfel und Turner (1979), S. 40). Ashforth und Mael (1989, S. 25-26) fanden heraus, dass die Identifikation mit einem Unternehmen zu einem Interesse an dessen Erfolg führt. Ein positiver Effekt, in Form einer erhöhten Kaufintension, zwischen der Identifikation mit dem

Gesponserten, beispielsweise einem Fußballverein, und dem Sponsor konnte ebenfalls nachgewiesen werden (vgl. Madrigal (2000), S. 21).

Die für diese Arbeit wichtigste Theorie beschäftigt sich mit dem Transfer des Images des Gesponserten auf den Sponsor. Das Modell wurde von McCracken (1989, S. 314-316) ursprünglich zur Erklärung von Übertragungseffekten von Testimonials auf Produkte konstruiert. Es beschreibt einen dreistufigen Prozess, in welchem die Meinung über die prominente Person auf das Produkt übergeht und letztendlich vom Produkt durch den Konsum auf den Konsumenten. Ganassali und Didellon (1996, S. 43-44) adaptieren in ihrer Arbeit "Le transfert comme principe central du parrainage" das Modell und passen es für das Sponsoring an. Auf affektiver Ebene übertragen sich die mit dem Gesponserten assoziierten Emotionen, auf den Sponsoren. Ebenso wird auf der kognitiven Ebene das bewusst wahrgenommene Image des Gesponserten auf den Sponsor projiziert. Im Falle von kulturellen Events verstehen sich diese Assoziationen auf Begriffe, wie anspruchsvoll, elitär oder seriös (vgl. Meenaghan und Shipley (1999), S. 342).

# 3.4. Einfluss auf die Corporate Reputation

Um den tatsächlichen Effekt von Kultursponsoring auf ein Unternehmen zu bewerten, muss zunächst nach einer messbaren Variable dafür gesucht werden. Da sich die Forschung in dieser Arbeit an dem Artikel "Art for the Sake of the Corporation Audi, BMW Group, DaimlerChrysler, Montblanc, Siemens, and Volkswagen Help Explore the Effect of Sponsorship on Corporate Reputations" von Schwaiger et al. (2010) orientiert, wurde, wie im genannten Artikel, die Corporate Reputation gewählt.

# 3.4.1. Bestandteile einer Corporate Reputation

Es existieren verschiedene Modelle und Definitionen zur Corporate Reputation. Einige beschränken sich nur auf die kognitiven Assoziationen zu einer Firma (vgl. Gray und Balmer (1998), S. 696-697), andere konzentrieren sich auf die Emotionen, welche die Öffentlichkeit mit dem Unternehmen verbindet (vgl. Fombrun und Rindova (2000), S. 78-79). Ein Modell, welches beide Komponenten verbindet, konstruiert Schwaiger (2004) in einem Artikel über die Corporate Reputation.

Als erste Dimension fasst Schwaiger (2004, S. 53-63) alle affektiven oder emotionalen Assoziationen mit einem Unternehmen unter dem Begriff "Likeabillity" zusammen. Zu Deutsch am besten mit Sympathie zu übersetzen, schließt sie Komponenten wie Identifikation oder Vertrauenswürdigkeit mit ein. Bei der zweiten Dimension handelt es sich um die "Competence" (dt. Kompetenz) eines Unternehmens. Geprägt von den kognitiven Verbindungen zu einer Firma, geht diese stärker auf die wirtschaftliche Leistung des Unternehmens ein. Bestandteile sind beispielsweise die Qualität der Produkte oder guter Service (vgl. Schwaiger (2004), S. 63). Die Einflüsse hinter diesen Dimensionen konnten zu vier Treibern zusammengefasst werden. Diese unterteilen sich

in Qualität der Produkte und Services; wirtschaftliche und Management Performance; Respekt für das Unternehmen in Bereichen wie zum Beispiel Umwelt oder Wettbewerbsverhalten (Corporate Social Responsibility) und physische Attraktivität (Marken, Gebäude, Qualität der Mitarbeiter). Die Sympathie eines Unternehmens wird sowohl von Qualität und Respekt, als auch durch die Attraktivität positiv beeinflusst, wobei die beiden letztgenannten als Haupttreiber gelten. Einzig die Komponente Performance wirkt sich negativ auf die Sympathie aus. Schwaiger (2004, S. 65) schließt daraus, dass ein Unternehmen, das sich in der öffentlichen Wahrnehmung besonders auf finanzielle Ziele konzentriert, schwerer emotionale Nähe zur Öffentlichkeit aufbauen kann. Auf die Dimension Kompetenz dagegen haben die Performance sowie Qualität den größten Einfluss. Allerdings weist obendrein die Attraktivität eine positive Korrelation auf. Auf diese Dimension wirkt sich nur die Komponente Respekt negativ aus. Dies begründet sich unter Umständen mit Folgender Argumentation, die Schwaiger (2004, S. 66) erklärt, "being a good corporate citizen, taking over social responsibility and taking care of environmental issues makes the company loved but not necessarily a target for investors." Um das Modell zu veranschaulichen, wurde in Anlehnung an Schwaiger (2004, S. 66) ein Template erstellt, in welchem die verschiedenen Komponenten und die, aus ihnen resultierenden Dimensionen zu sehen sind (vgl. Abbildung 1).

Die positiven Effekte einer starken Corporate Reputation werden in verschiedenen Forschungen demonstriert. Eine Studie von Turban und Cable (2003, S. 745-746) zeigt die Überlegenheit von Unternehmen mit starker Reputation im Bewerbungsprozess um neue Mitarbeiter. Neben einem größeren Bewerberpool nimmt ferner die Qualität der Bewerber bis zu einem gewissen Grad zu. Nach Lafferty und Goldsmith (1999, S. 109-113), welche die Corporate Reputation mit der Glaubwürdigkeit eines Unternehmens auf eine Stufe stellen, hat diese außerdem positiven Einfluss auf die Wahrnehmung der Marken einer Firma. Außerdem erhöht sich mit hoher Reputation das Vertrauen der Konsumenten in die Produkte und Services des Unternehmens. Sie gewährleistet für Kunden sowie Außenstehende und Mitarbeiter die Konsistenz der Handlungen mit dem Image des Unternehmens (vgl. Fombrun und Van Riel (1997), S. 6). Des Weiteren ergeben sich für Unternehmen mit starker Corporate Reputation Vorteile in der Preissetzung, da sie im Stande sind nicht nur zu niedrigeren Beschaffungspreisen einzukaufen, sondern auch zu einem Premiumpreis zu verkaufen (vgl. Schwalbach (2000), S. 289). Zusätzlich fanden Eberl und Schwaiger (2005, S. 851) heraus, dass Firmen mit einer hohen Reputation im Bereich der kognitiven Komponente der Corporate Reputation in der Zukunft eine bessere finanzielle Performance aufweisen. Die Vielzahl von positiven Effekten unterstützt die Hypothese von Schwaiger (2004, S. 51), eine starke Corporate Reputation als strategischen Wettbewerbsvorteil nach Porter zu sehen.

3.4.2. Einfluss des Kultursponsorings auf die Corporate Reputation

Als Hauptvoraussetzung für eine starke Reputation führt Ballen (1992, S. 40) neben qualitativ hochwertigen Produkten und finanziellen Komponenten auch die Performance der Unternehmensleitung an. Darunter fällt das Bewusstsein über Maßnahmen, welche im Stande sind die Corporate Reputation zu verbessern wie möglicherweise Kultursponsoring. Im Folgenden soll erläutert werden, wie Kultursponsoring auf die Corporate Reputation wirkt, wenn die bereits genannten Übereinstimmungen in den Bereichen Demographie, Geographie und Lifestyle gegeben sind und das Event ausreichend durch Massenmedien publiziert wurde.

Basierend auf dem bereits beschriebenen Modell des Imagetransfers von Ganassali und Didellon (1996, S. 43-44) folgern Schwaiger et al. (2010, S. 80), dass sich Kultursponsoring aufgrund des Transfers der affektiven und kognitiven Assoziationen von Gesponserten auf Sponsoren positiv auf die zwei Dimensionen der Corporate Reputation eines Unternehmens auswirkt. Zurückzuführen ist das auf die positiven Auswirkungen von Kultursponsoring in Bezug auf die verschiedenen Treiber der zwei Dimensionen der Corporate Reputation. Die Studie von Schwaiger et al. (2010, S. 85-86) belegt allerdings nur einen signifikanten Einfluss des Kultursponsorings auf die Dimension Sympathie der Corporate Reputation. Auch mit den Treibern hinter den Dimensionen verhält es sich ähnlich; für Quality und Performance konnte kein signifikanter Einfluss nachgewiesen werden. Der Einfluss des Kultursponsorings auf die Variable Respekt oder Corporate Social Responsibility resultiert nach Schwaiger et al. (2010, S. 80) aus der Ähnlichkeit des Sponsorings zur Philanthropie. Auf den ersten Blick scheint Sponsoring für den Außenstehenden möglicherweise nicht als kommerziell genutztes Marketingtool, sondern als altruistische Geste gegenüber dem Gesponserten (vgl. McDonald (1991), S. 35-36). Auch der Treiber Attraktivität wird vom Kultursponsoring positiv beeinflusst (vgl. Schwaiger et al. (2010), S. 86). Dies ist möglicherweise darauf zurückzuführen, dass Sponsoring neben den Zielen, welche an externe Interessensgruppen gerichtet sind, überdies interne Ziele verfolgt. Der Stolz auf das Unternehmen und der Transfer von Werten verstehen sich darauf, die Moral der Mitarbeiter zu erhöhen (vgl. Meenaghan (1991), S. 40). Unter Umständen resultiert daraus eine höhere Attraktivität des Unternehmens bei potenziellen Mitarbeitern.

Daraus lässt sich nach Schwaiger et al. (2010, S. 87) ein positiver Einfluss des Kultursponsoring auf die Sympathie eines Unternehmens folgern. Wie bereits beschrieben, handelt es sich dabei um die affektiven oder emotionalen Verbindungen der Öffentlichkeit zu einem Unternehmen. Diese können folglich unter Umständen über die Treiber Corporate Social Responsibility und Attraktivität durch Kultursponsoring verbessert werden, wobei das Hauptaugenmerk allerdings auf ersterem Treiber liegt. Durch den damit verbundenen Anstieg der Corporate Reputation besteht die Chance die Vorteile einer starken Corporate Reputation zu realisieren. Eine wei-

**Corporate Reputation** 

# CSR affektiv — Attraktivität Qualität Kompetenz

# Abbildung 1: Bestandteile der Corporate Reputation; Quelle: Eigene Darstellung in Anlehnung an Schwaiger (2004), S. 66

Dimensionen

tere interessante und für diese Arbeit essenzielle Erkenntnis sind die Unterschiede der Ergebnisse von Schwaiger et al. (2010, S. 86-87) hinsichtlich der soziodemographischen Variable "Interest in cultural activites". Entgegen der Erwartungen von Schwaiger et al. (2010, S. 87) bewerteten Personen mit einem hohem Interesse am kulturellen Kultursponsoring signifikant negativer als andere. Potenzielle Gründe dafür sollen in dieser Arbeit herausgearbeitet werden.

Performance

# 3.5. Unterschiede zum Sportsponsoring

kognitiv

Komponenten

Den größten Anteil am Sponsoringbudget deutscher Unternehmen hat neben dem Kultursponsoring mit 18,5 % das Sportsponsoring mit 43 % (vgl. Ostfalia (2013)). Im Folgenden sollen die Unterschiede zwischen Sportund Kultursponsoring nach Quester und Thompson (2001, S. 35) aufgezeigt und ergänzt werden. Differenzen existieren zum einen hinsichtlich der gewünschten Ziele und deren Evaluierung. Zum anderen beziehen sie sich auf das transferierte Image und betreffen zuletzt die Verantwortung in der Organisation.

Ziele von kulturellem Engagement richten sich mehr darauf, die Beziehung zwischen der Öffentlichkeit und dem Unternehmen zu verbessern sowie auf die Verbesserung des Images (vgl. Quester und Thompson (2001), S. 35; Witcher et al. (1991), S. 17). Sportsponsoring hingegen konzentriert sich stärker auf die Reichweite des Sponsorings und die Verbesserung der Bekanntheit des Unternehmens beziehungsweise einzelner Marken (vgl. Sparks (2015), S. 4-6; Witcher et al. (1991), S. 17). Zusätzlich evaluieren Firmen die Ziele des Kultursponsorings häufiger (vgl. Witcher et al. (1991), S. 19). Der Grund dafür liegt nach Quester und Thompson (2001, S. 35) in einem kleinerem Budget für Kultursponsoring im Gegensatz zum Sportsponsoring, welches daher genauer auf Sinnhaftigkeit überprüft wird. Ein weiterer Unterschied ist im transferierten Image des Sponsoringengagements zu suchen; Sportevents vertreten Werte

wie energetisch, Jugendlichkeit oder Maskulinität, kulturelle Events hingegen werden eher als elitär, pompös oder seriös wahrgenommen (vgl. Meenaghan und Shipley (1999), S. 342). Folglich hat das Image der Veranstaltung auch Einfluss auf das direkt angesprochene Publikum, welches das Event besucht. Meenaghan und Shipley (1999, S. 340) decken einen weiteren wichtigen Unterschied zwischen Sport und Kultursponsoring auf. Während Sportsponsoring in der öffentlichen Wahrnehmung fast mit klassischer Werbung gleichzusetzen ist, genießt Kultursponsoring ein stärker philanthropisches Image. Ähnlich, aber ein wenig schwächer, wie bei Sponsoring für soziale Institutionen geht die Öffentlichkeit bei Kultursponsoring von einem guten Willen des Unternehmens aus. Zuletzt ergeben sich Abweichungen in der Organisation der Sponsoringaktivität. Die Verantwortung für Kultursponsoring liegt nach Witcher et al. (1991, S. 17-18) vorwiegend in der Public-Relations-Abteilung, wohingegen beim Sportsponsoring die Marketingabteilung im Vordergrund steht. Dies könnte als Erklärung für die Unterschiede in der Zielsetzung fungieren. Während die Public-Relations-Abteilung traditionell für die Beziehung und das Image einer Firma verantwortlich ist, orientiert sich die Marketingabteilung mehr an kommerziellen Zielen.

# 4. Diskrepanzen in der Wahrnehmung von Kultursponsoring unter Berücksichtigung des generellen Interesses an Kultur

# 4.1. Gegenstand der Forschung

Wie bereits erklärt wurde, hat das Kultursponsoring einen positiven Einfluss auf die Corporate Reputation, genauer auf die affektive Dimension der Sympathie eines Unternehmens. In der Forschung von Schwaiger et al. (2010) fiel auf, dass für Personen mit einem höheren Interesse an kulturellen Aktivitäten die Sympathie für das Unternehmen nicht stieg. Die

Aufgabe dieser Forschung ist es daher, potenzielle Gründe herauszufinden, warum kulturinteressierte Menschen möglicherweise negative Gefühle gegenüber Kultursponsoring haben.

Nach Cornwell et al. (2005, S. 31) haben individuelle Faktoren wie Arousal (dt.: Aufgeregtheit) oder Involvement (dt.: Beteiligung) Einfluss auf die Verarbeitung von Sponsoringstimuli. Nach dem Intensity Principle (dt.: Prinzip der Intensität) werden bei zu großer Aufgeregtheit für das Event andere Stimuli, beispielsweise Logos eines Sponsoren, nicht mehr so einfach verarbeitet (vgl. Pavelchak et al. (1988), S. 365-366). Newell et al. (2001, S. 1146) bestätigen diese Theorie mit einer Studie über den Superbowl. Der Recall auf Werbung aus der ersten Hälfte war dabei besser als der Recall aus der zweiten, spannenderen Hälfte des Spiels. Möglicherweise hat diese Theorie Einfluss auf diese Forschung. Personen, welche stärkeres Interesse an dem kulturellen Event zeigen, könnten von diesem abgelenkt sein und somit Sponsoringstimuli nicht so gut verarbeiten. Kongruent wäre das mit den Forschungsergebnissen von Pham (1992, S. 89-90), nach welchen Sponsoringstimuli in aufregenden Events signifikant schlechter aufgenommen werden.

Negative Gefühle oder Verwirrung treten potenziell überdies auf, wenn keine ausreichende Übereinstimmung zwischen Sponsor und Gesponsertem besteht. An den bereits dargelegten Konzepten von Kongruenz und Articulation angelehnt, existiert die Vermutung Personen, welche sich gut mit Kultur auskennen und die Werte der einzelnen Veranstaltungen kennen, bewerten Sponsoren mit einem geringen Fit zum Event schlechter. Zusätzlich besteht die Möglichkeit einer Verbindung zwischen der Glaubwürdigkeit eines Sponsoringengagements und dem Fit.

Auch das affektive Involvement einer Person kann in Verbindung mit der Kongruenz zwischen Sponsor und Event Einfluss auf die Wahrnehmung von Sponsoring haben (vgl. Schaefer und Keillor (1997), S. 30). Unter dem affektiven Involvement versteht man die persönliche Relevanz der emotionalen Nachricht (Zaichkowsky (1994), S. 60), also im Sponsoring den Versuch der Übertragung des Images auf den Sponsoren. Nach Schaefer und Keillor (1997, S. 30) verbessert sich die Einstellung von Personen mit hohem Involvement gegenüber dem Sponsor mit besserer Verbindung zwischen Sponsor und Gesponsertem. Im Umkehrschluss sollten Personen mit hohem Interesse an kulturellen Aktivitäten Sponsoring als weniger positiv empfinden, wenn sie keine Verbindung zwischen Sponsor und Gesponsertem erkennen. Diese Erkenntnis würde die bereits geschilderten Theorien zur Kongruenz und Articulation stützen.

Ein weiterer möglicher Grund für negative Empfindungen gegenüber Kultursponsoring lässt sich aus folgendem Zitat herauslesen: "Nein, das fürchte ich nicht. Unsere Sponsoren sind viel zu intelligent, um nicht zu wissen, dass inhaltliche Einflussnahme ein Bumerang wäre. Im Gegenteil, ich hoffe, unsere Sponsoren nehmen weiter in dem Sinne Einfluss, dass sie uns ermöglichen, Projekte zu verwirklichen, für die uns sonst das Geld fehlte." So antwortet Rabl-Stadler (2015) auf die Frage, ob sie eine Beeinflussung durch die Sponso-

ren der Salzburger Festspiele fürchte. Auch wenn die Präsidentin der Salzburger Festspiele nicht der Meinung ist, ihre Sponsoren könnten Einfluss auf ihr Programm nehmen, empfinden aber unter Umständen die Besucher des Festivals dies so. Eine Verknüpfung der finanziellen Unterstützungen an bestimmte Bedingungen ist durchaus denkbar. Dadurch würden sich Sponsoren aus der Wirtschaft ermächtigen, auf das Programm, die Zielgruppe oder die politische Ausrichtung der kulturellen Institution Einfluss zu nehmen. Als Resultat wäre neben dem Verschwinden von wirtschaftlich kritischer Kunst auch ein Wettbewerb um Sponsorengelder denkbar. Vor allem bei Personen mit einem hohen Interesse an kulturellen Aktivitäten, könnte die Befürchtung einer Einflussnahme auf den Sponsor negative Gefühle auslösen, da diese Personen größeren Wert auf die künstlerische Integrität des Events legen.

Außerdem besteht Potenzial, dass die Befürchtung entsteht ein kulturelles Event diene nur noch als Marketingbeziehungsweise PR-Plattform eines oder mehrerer Unternehmen. Kulturinteressierte Personen sehen diese Punkte unter Umständen kritischer als andere, da sie sich besser mit Kultur auskennen und somit mehr Wert auf die künstlerische und kulturelle Qualität legen.

# 4.2. Methodik der Forschung

Nach Mayring (2002, S. 41-42) stehen Einzelfallanalysen im Mittelpunkt der qualitativen Forschung. Sie ermöglichen es, genaue und tiefgreifende Ergebnisse über Zusammenhänge und Einflüsse zu erhalten, insbesondere im Kontext der Ganzheit der Analyse. Die Fragestellung der Arbeit verlangt genau diese Art der Individualität in einem konkreten Kontext.

Als Verfahren dafür bietet sich das problemzentrierte Interview an. Es zeichnet sich durch seine offene halbstrukturierte Form aus und darf einem offenen Gespräch gleichen. Dadurch sollen ehrliche, ausführliche und auf einer Vertrauensbeziehung fundierte Antworten gewährleistet werden. Besonders für dieses Forschungsthema sind persönliche Meinungen sowie Erfahrungen von großem Wert. Der Interviewer muss allerdings auf die Zentrierung des Gesprächs und auf die spezifische Problemstellung achten. Durch einen Leitfaden wird das Abschweifen des Interviews verhindert. Ad-hoc-Fragen können bei Verständnisproblemen sowie für Vertiefungen des Themas hilfreich sein (vgl. Mayring (2002), S. 67-69). Vor der Durchführung wurde den Befragten die Thematik und der Ablauf des Interviews erklärt und, nach Erklärung des Einverständnisses, Tonaufnahmen angefertigt. Zum Interview erklärten sich neun Personen bereit, von denen allerdings nur acht das Interview nach den Sondierungsfragen fortsetzen konnten. Der Verbleibende erfüllte nach Ansicht des Interviewers die Voraussetzung "hohes Interesse an kulturellen Aktivitäten" im Bereich der High-Brow Kultur nicht. Das vollständige Interview wurde mit sechs Frauen und zwei Männern durchgeführt, wobei das durchschnittliche Alter 37,25 Jahre betrug. Die Verteilung nach sonstigen demographischen Variablen war ausgeglichen.

Für die Aufbereitung der Tonbandaufnahmen eignet sich besonders ein zusammenfassendes Protokoll für jedes der Interviews. Dieses zeichnet sich durch eine sinnvolle Zusammenfassung des Materials während der Transkription aus (vgl. Mayring (2002), S. 94-95). Dabei wurden lange Sätze verkürzt, Füllwörter vermieden und die Aussagen teilweise sprachlich aufgewertet. Außerdem wurden gleiche Begriffe vereinheitlicht sowie unbedeutende Aussagen gestrichen. Es wurde streng darauf geachtet, keine der Antworten inhaltlich zu verändern oder zu verfälschen.

Nach Mayring (2002, S. 94) eignet sich die qualitative Inhaltanalyse zur Auswertung des zusammenfassenden Protokolls. Hierbei vermag das Material mit Hilfe einer induktiven Kategorienbildung qualitativ interpretiert zu werden, ohne der Grundlage zuvor formulierter Theorienkonzepte (vgl. Mayring (2000), S. 75). Dafür wurden zunächst Selektionskriterien für Kategorien aufgestellt und nach einem Materialdurchgang 25 % des Materials angepasst. Diese systematische Vorgehensweise führt zu einer genauen und detaillierten Auswertung der Materialmenge und erlaubt es, das Material zu analysieren und zu interpretieren (siehe Anhang).

# 4.3. Ergebnisse

Die Sondierungsfragen zu Beginn des Interviews sollten dafür sorgen, dass nur Personen das komplette Interview durchführen, welche als interessiert in kulturellen Aktivitäten gelten. Dabei wurde sowohl Wert auf die eigene Einschätzung der Interviewpartner gelegt als auch auf die Regelmäßigkeit der Besuche von kulturellen Events aus dem High-Brow Bereich. Als Beispiel für die Selbsteinschätzung lässt sich folgendes Zitat heranziehen: "Ja, ich bin kulturinteressiert. Ich habe Kunst studiert im Rahmen meines Studiums der Schulpsychologie und gehe, seit ich im Gymnasium war, gerne in Museen und Vorstellungen im Bereich Theater und ähnliches." Vier der Befragten vermochte dadurch ein sehr starkes Interesse an kulturellen Aktivitäten nachgewiesen werden. Auch die verbleibenden vier zeichneten sich entweder durch ein starkes oder mittelstarkes Interesse an kulturellen Aktivitäten aus. Favorisierte Veranstaltungen aus dem Bereich High-Brow fanden sich in der Oper, Kunstaustellungen, Museen und andere musikalische Events. Genannte Beispiele zu den einzelnen Events sind Tabelle 1 zu entnehmen.

Um den Interviewpartnern möglichst persönliche Erinnerungen zu entlocken, sollten sie zunächst von eigenen Erinnerungen mit Sponsoren vergangener Veranstaltungen berichte. Vier der Personen konnten sich an einen Sponsor eines vergangenen Events erinnern. Dies könnte die erste überraschende Erkenntnis dieser Arbeit sein. Es erscheint ungewöhnlich, dass nur die Hälfte der Befragten einen Sponsor zu einem kulturellen Event, an welchem sie selbst teilgenommen haben, benennen kann. Außerdem bezogen sich drei der Erinnerungen auf Sponsoren von jährlich einmalig stattfindenden Großevents und fielen tendenziell positiv aus. Ferner waren dazu die Gefühle zum Sponsor entweder neutral oder positiv, was folgendes Zitat über den Sponsor der Opernfestspiele in München zeigt: "Meine Einstellung zu

dem Event wurde nicht verändert. Und zu BMW auch nicht." Zudem resultiert aus der Befragung eine positive Erinnerung an die Sponsoren der Wormser Nibelungenfestspiele: "Sehr gut, weil es sehr viele waren und diese geholfen haben, dass die Spiele weiterhin in Worms aufgeführt werden."

Nach den generellen Auswirkungen gefragt, welche ein Sponsor auf ein kulturelles Event haben kann, standen für die Befragten vor allem finanzielle Effekte im Vordergrund. Für drei der Personen gingen diese Effekte, wie Vergrößerung oder qualitative Verbesserung des Events ohne negative Auswirkungen wie einer zu starken Kommerzialisierung einher. Die verbleibenden Personen schätzen ebenfalls die finanziellen Zuwendungen, sehen aber ebenso kritische Aspekte: "Gesponserte Events sind meist größer und unpersönlicher, man fühlt sich etwas verloren in der Masse." Des Weiteren haben sechs der Personen die Befürchtung, ein Sponsor könnte Einfluss auf das Programm oder den Ablauf eines kulturellen Events nehmen. Vor allem die folgende Befürchtung birgt Potenzial für negative Gefühle gegenüber Kultursponsoring: "Bestimmt, dafür dass die sponsern, wollen die dann ja auch Einfluss. Besonders in der Kunstbranche könnten darunter kleine Künstler leiden, da Unternehmen eher bekannte Künstler im Programm haben wollen."

Interessante Tendenzen ergaben sich ebenso in der Art des Auftretens des Sponsors. Um dies zu evaluieren, wurden den Interviewpartnern zwei unterschiedliche Formen des Auftritts eines möglichen Sponsors präsentiert. Zum einen gehörte dazu die Firma Linde AG, welche ein Engagement als Spielzeitpartner der Oper München verfolgt. Dieses Engagement beschrieb sich den Befragten als sehr dezent und zurückhaltend, nur in Form von Flyern in Programmheften. Als Vergleich diente der Auftritt der Audi AG auf den Salzburger Festspielen, wo das Unternehmen als Hauptsponsor fungiert. Das dortige Sponsoring wurde als aufdringlicher dargestellt und mit mehr Vorteilen für die Audi AG verbunden. Konfrontiert mit diesen Beispielen sprachen sich fünf der Befragten für ein dezentes Sponsoring aus. Zusätzlich hatten diese Personen während des weiteren Interviews Bedenken, ein Sponsor könnte auf einem kulturellen Event zu stark in den Vordergrund treten. Deutlich wird das durch folgende Aussagen zum allgemeinen Auftreten von Sponsoren; "Der Sponsor könnte in seinem Auftreten Überhand nehmen, das heißt sich zu stark in Szene setzen. Zum Beispiel durch Ablenkung die Sicht auf die Bühne verschlechtern, oder durch Banner ablenken vom eigentlichen Geschehen." Zum anderen lässt sich das konkret auf die Beispiele beziehen: "Ich persönlich finde das Dezente sympathischer, weil immer noch das Kulturevent im Vordergrund stehen soll. Für die Firma ist es wahrscheinlich besser, aggressiver aufzutreten, da es einprägsamer ist, aber das kommt nicht mehr so positiv an, da das eigentliche Geschehen in den Hintergrund gerät." Auch in diesen Aussagen lassen sich potenzielle Auslöser für negative Empfindungen finden. Die Personen, welche Sponsoring schon bei vorherigen Fragen positiver oder gleichgültiger bewerteten, sehen für die aggressivere Form keine Probleme: "Natürlich, je größer ein Sponsor auftritt, umso mehr wird er wahrgenommen. Bei den Salzburger Festspielen mit den

Tabelle 1: Favorisierte Kundenveranstaltungen (High-Brow); Quelle: Eigene Darstellu	Tabelle 1:	: Favorisierte	Kundenveranstalti	ıngen (High-Brov	w): Ouelle: E	ligene Darstellun
---	------------	----------------	-------------------	------------------	---------------	-------------------

	Favorisi	erte Kulturveranstaltungen (High-Brow)
Veranstaltung	Anzahl	Beispiele
Oper	4	Opernfestspiele in München, Entführung aus dem Serail
Museum	2	staatliches Museum Ägyptischer Kunst
Musik	2	Münchner Philharmonie
Literatur	0	
Kunst	3	Private Vernissagen

Shuttleservices ist ja nicht für die Allgemeinheit, sondern nur für einige bestimmte Leute, die man ansprechen will. Ich finde beides aber gut, weil beide Arten des Sponsorings unterstützen das kulturelle Leben."

Das beschriebene Beispiel kann überdies genutzt werden, um die emotionalen Assoziationen zu einem Sponsor zu überprüfen. Zusätzlich sollten die Befragten von ihren spontanen Empfindungen berichten, wenn sie von einem Sponsoring im kulturellen Bereich hören. Es fällt auf, dass im ersten Moment Sponsoring mit positiven Gefühlen bedacht wird wie zum Beispiel Dankbarkeit über den Einsatz für das kulturelle Event: "Im ersten Moment positiv. Wie gesagt, das Unternehmen stellt was auf die Beine und bringt mir einen Künstler her, den ich mir dann anschauen kann. Also ich gehe da sehr unvoreingenommen ran, und denke mir erstmal, das ist was Gutes." Sechs Personen verbanden im ersten Moment positive Gefühle mit dem Sponsor, wobei allerdings drei von diesen ihre positiven Gefühle an Bedingungen, wie zum Beispiel ein dezentes Auftreten oder einen unbeeinflussten Ablauf des Events knüpften. Wieder ist erstaunlich, dass zwei Personen sich ihrer Meinung nach nicht vom Sponsoring in ihrer Wahrnehmung über den Sponsor beeinflussen lassen.

Bereits 2006 beschäftigten sich Cornwell et al. (2006, S. 216) mit der Theorie zur Übereinstimmung der Werte zwischen Sponsor und Gesponsertem. Demnach sollten die Einstellungen der Befragten zu diesem Thema als mögliche Quelle negativer Empfindungen überprüft werden. Dazu wurde zum einen nach der Wichtigkeit einer Übereinstimmung der Werte zwischen Event und Sponsor gefragt, sowie nach einer Übereinstimmung des Umfelds der beiden. Zuvor durch Beispiele veranschaulicht, empfanden fünf Personen den Fit als einigermaßen wichtig oder wichtig. Ein nicht gewährleisteter Fit könnte also durchaus ein Auslöser negative Gefühle sein, wie folgendes Zitat zeigt: "Finde ich schon wichtig. Das Feeling, zum Beispiel sich elegant zu kleiden, welches mit einem Opernbesuch verbunden ist, würde bleiben, wenn da zum Beispiel ein teures Auto als Sponsoring steht. Das passt zusammen mit dem Auftreten von sich selber. Aber wenn da eine große Packung Pommersche Frischwurst stehen würden, dann wäre ich sehr irritiert." Durch den fehlenden Fit besteht das Potenzial das Sponsoring negativ in Erinnerung zu behalten. Die drei Befragten, welche die Übereinstimmung nicht als wichtig einschätzten, beurteilten das

Sponsoring während des Interviews tendenziell rationaler als andere.

Neben den Theorien zum Fit hat zusätzlich die Aufmerksamkeit der Besucher Einfluss auf das Sponsoring (vgl. Cornwell et al. (2005), S. 31). Daher bestand die Vermutung, Personen würden Sponsoren auf ihren bevorzugten Veranstaltungen weniger wahrnehmen, da sie sich mehr auf das Event konzentrieren. Tendenziell nehmen die Befragten den Sponsor auf unterschiedliche Arten wahr. Zwei der Personen, welche im Interview durchgehend durch ihre rationale Einschätzung für Kultursponsoring auffielen, gaben eine positivere Einschätzung für Sponsoren auf ihren bevorzugten Events: Vielleicht in geringer Weise. Da finde ich es besser, da das Event sonst nicht für alle zugänglich wäre." Zwei weitere Personen hatten keine unterschiedliche Wahrnehmung, während sich drei als sensibler auf diesen Events einschätzten: "Ja, aggressives Sponsoring ärgert mich dann mehr, da es die Veranstaltung ein bisschen kaputt macht." Die Erwartungen hinsichtlich der Aufmerksamkeit wurden von einer Person bestätigt, welche aussagte, auf bevorzugten Events weniger auf die Sponsoren zu achten. Die Ergebnisse aus den Sondierungsfragen deuten zusätzlich auf eine tendenziell schwache Wahrnehmung von Sponsoren auf kulturellen Veranstaltungen hin. Auch hier finden sich demnach mögliche Quellen negativer Assoziationen gegenüber dem Sponsor.

Ferner birgt zusätzlich die Frage nach der Glaubwürdigkeit von Kultursponsoring Potenzial für negative Gefühle. Für sechs der Befragten ist Kultursponsoring durchaus glaubwürdig, allerdings für vier nur unter bestimmten Voraussetzungen. Darunter fielen neben dem bereits angesprochenem Fit darüber hinaus die sonstigen Sponsoringaktivitäten des Sponsors. "Bei BMW empfinde ich es als glaubwürdig und gut. Die veranstalten ja auch diese Jazz Matineen. Auch bei Audi, die setzen sich ja auch generell viel für Kunst ein." Eine Person blieb bei der Frage indifferent, für die Verbliebene blieben die Ziele des Unternehmens im Vordergrund: "Man muss schon sagen, dass der Sponsor nicht das Kulturgut weiterverbreiten will, sondern Werbung betreiben will. Das muss klar sein."

Um gegen Ende des Interviews eine allgemeinere Meinung der Befragten zum Verhältnis zwischen Werbung und Kultur zu bekommen, wurden sie zur Vereinbarkeit der beiden befragt. Vier der Personen gaben an Werbung würde sich nur unter gewissen Umständen mit einem kulturellen Events

vereinbaren lassen. Dabei standen die bereits öfters genannten Voraussetzungen, wie dezentes Auftreten oder ähnliche Werte im Vordergrund, was durch diese Aussage verdeutlicht wird: "Ja wenn die Werbung dezent im Hintergrund bleibt." Deutlich negativere Worte fand eine der zwei Personen für die Werbung im kulturellen Bereich nicht zu suchen hat: "Eigentlich nicht, da Kultur für mich für Unabhängigkeit, Kreativität und Freiheit steht. Dafür, dass jeder seine Ideen leben kann und da hat Werbung eigentlich nichts zu suchen, da diese die Konsumgesellschaft und Zwänge verkörpert." Die verbliebenen zwei Befragten sahen keine Probleme in der Vereinbarkeit von Werbung und Kultur.

Letztendlich sollten die Befragten ihre Gesamteinschätzung zum Thema Kultursponsoring abgeben. Tendenziell fielen diese bei allen Interviewpartnern positiv aus. Allerdings, wie schon mehrfach bei vorherigen Fragen, knüpften fünf Personen ihre Wertung an verschiedene Bedingungen. Diese spiegelten oftmals die potenziellen Quellen negativer Gefühle aus anderen Fragen des Interviews wider. Bezeichnend dafür ist folgendes Zitat: "Ich sehe viele Vorteile für die Veranstaltung und den Sponsor, da für beide mehr Publikum kommt. Dennoch muss man aufpassen, dass man es nicht übertreibt und es nicht zu aggressiv ist." Die Personen, die nur Vorteile für das Event sahen, hatten ebenfalls in anderen Fragen eine positive aber sehr rationale Sicht auf Kultursponsoring. Für sie stand meist der finanzielle Aspekt im Vordergrund und darüber hinaus die damit verbundenen persönlichen Vorteile wie günstigere Eintrittskarten oder überhaupt das Stattfinden des Events: "Ich sehe eigentlich nur Vorteile. Das ist eine finanzielle Hilfe für das Event, damit es ausgetragen werden kann."

# 4.4. Diskussion

Die Ergebnisse dieser Studie lassen einige Tendenzen und Meinungsbilder über Kultursponsoring erkennen, welche für Sponsoren kultureller Events von Bedeutung sein können. Ziel der Forschung war es, mögliche Fundierungen für negative Gefühle gegenüber Sponsoren kultureller Events zu finden. Anhand der Ergebnisse wurden bereits Rückschlüsse auf diese potenziellen Quellen gezogen. Diese werden nachfolgend diskutiert.

Wie schon vorgestellt, konnten nur vier der acht Befragten von persönlichen Erfahrungen mit einem Sponsor berichten. Drei dieser Personen erwähnten Sponsoren größerer Events, wie von den Opernfestspielen in München sowie den Niebelungenfestspielen in Worms, mit tendenziell eher aggressiv auftretenden Sponsoren. Bei alltäglichen kulturellen Veranstaltungen konnte nur eine Person von einem Sponsor berichten. Daraus ergibt sich der Verdacht, Sponsoren auf alltäglichen kulturellen Veranstaltungen fänden bei Personen mit einem hohen Interesse an kulturellen Aktivitäten nur wenig Beachtung. Dies könnte durch die Theorie von Cornwell et al. (2005, S. 31) über Arousal und Involvement argumentativ unterstützt werden. Direkt auf die Wahrnehmung von Sponsoren auf ihren bevorzugten Events angesprochen, erscheint zusätzlich die erhöhte Sensibilität gegenüber Sponsoren von Bedeutung. Geht man davon aus, dass den Personen

Sponsoren unbewusst weniger auffallen und verknüpft dies mit einer erhöhten Sensibilität gegenüber zu aufdringlichem Sponsoring, stellt das möglicherweise einen Ursprung für negative Gefühle dar. Ein normalerweise positiv wahrgenommener Sponsor würde aufgrund der verringerten Aufmerksamkeit nicht in Erinnerung bleiben. Zusätzlich führt die gesteigerte Sensibilität gegenüber negativ aufgefasstem Sponsoring zu größerem Unmut darüber. Unter Umständen trägt dies zusätzlich zu einer verzehrten Wahrnehmung des Kultursponsorings bei Betrachtung mit quantitativen Methoden bei. Die nachfolgenden Punkte erhöhen die Wahrscheinlichkeit für eine negative Auffassung des Sponsorings.

Im Allgemeinen erweckten die Interviews eine positive Stimmung gegenüber Sponsoring. Doch mit der finanziellen Unterstützung, welche von allen Befragten wertgeschätzt wird, gehen durchaus mögliche Quellen für negative Gefühle einher. Zum einen bestehen Bedenken über eine zu starke Kommerzialisierung des Events, welches laut einer Befragten unpersönlich oder sich zu einer Massenveranstaltung verändert. Damit verbunden existieren Befürchtungen, ein Sponsor könne sich durch die finanzielle Unterstützung Einfluss auf das Programm und den Ablauf des Events erkaufen. Speziell für Engagements unbekannter Künstler sah ein Befragter Probleme damit, da seiner Meinung nach Unternehmen nur bekanntere Künstler im Programm wollen.

Der nächste Punkt, den diese Arbeit als denkbare Ursache für negative Gefühle aufführt, hängt gleichermaßen mit der Kommerzialisierung des Events zusammen. Die Interviews erweckten den Eindruck, dass neben den positiven finanziellen Auswirkungen eines Sponsors das eigentliche Event nicht in den Hintergrund geraten darf. Die kulturelle Veranstaltung muss Vorrang vor dem Sponsor haben und das soll auf dem Event auch vermittelt werden. Ein dezentes Auftreten des Sponsors wurde tendenziell begrüßt. Demnach ist ein zu aggresssives Auftreten als möglicher Ursprung negativer Assoziationen mit Kultursponsoring auszumachen. Die scheint vorrangig mit dem ersten genannten Punkt in Konflikt zu stehen. Daraus resultiert die Frage, wie sich Sponsoren kultureller Events besser in den Erinnerungen der Besucher verankern, ohne zu aggressiv aufzutreten.

Überdies kann außerdem ein als unpassend wahrgenommenes Sponsoring Grund für negative Empfindungen gegenüber dem Sponsor sein. Unpassend dahingehend, dass entweder die Werte des Sponsoren und Gesponserten nicht übereinstimmen oder das Umfeld der beiden. Wie eine Forschung von Cornwell et al. (2006, S. 216) vermuten lässt, tendierten die Befragten dazu, einen fehlenden Fit als störend oder unpassend zu empfinden. Für Veranstaltungen mit einem elitären Charakter wie aus dem High-Brow Bereich (vgl. Meenaghan und Shipley (1999), S. 342) erscheint ein Sponsoringpartner mit ähnlichen Werten sinnvoll, wie eine der befragten Personen postuliert: "Das Feeling, zum Beispiel sich elegant zu kleiden, welches mit einem Opernbesuch verbunden ist, würde bleiben, wenn da zum Beispiel ein teures Auto als Sponsoring steht. Das passt zusammen mit dem Auftreten von sich selber. Aber wenn da eine große Packung Pommersche Frischwurst stehen würde, dann wäre ich sehr irritiert."



Abbildung 2: Mögliche Quellen negativer Gefühle gegenüber Sponsoren kultureller Events; Quelle: Eigene Darstellung

Die Frage nach der Vereinbarkeit von Werbung und Kultur muss kritisch diskutiert werden, da eine gewisse Schwierigkeit besteht Werbung und Sponsoring gleichzusetzen. Meenaghan und Shipley (1999, S. 340) stellen das Sponsoring elitärer Kunst, wie es die Bereiche der High-Brow Art abdecken, in der öffentlichen Wahrnehmung mehr in Richtung der Philanthropie, als zur klassischen Werbung. Somit wird dem Kultursponsoring ein guter Wille im Gegensatz zur Werbung unterstellt. Diese These lässt sich ergänzend in dieser Forschung erahnen. Beide Personen die sich gegen die Vereinbarkeit von Werbung und Kultur aussprachen, bewerteten Sponsoring im Allgemeinen tendenziell positiv. Es entstand zusätzlich der Eindruck die Vereinbarkeit sei nur unter gewissen Umständen wie einem dezenten Auftreten des Sponsors gewährleistet.

Ebenso ließen sich Hinweise finden, ein nicht gewährleisteter Fit gefährde die Glaubwürdigkeit von Kultursponsoring. Die Glaubwürdigkeit eines Sponsors hängt tendenziell sowohl mit diesem als auch mit dem Auftreten und sonstigen Engagement des Sponsors zusammen. Per se stellt die Glaubwürdigkeit daher keine Quelle für negative Gefühle dar, allerdings die bereits genannten Einflussfaktoren, welche sich in den vorherigen Punkten widerspiegeln. Ähnlich verhält es sich mit den emotionalen Assoziationen der Befragten mit dem Sponsor sowie der abschließenden Wertung für Kultursponsoring. Die gesammelten Meinungsbilder und Tendenzen lassen sich in einem Template zusammenfassen, welches bei der Implementierung von Strategien im Bereich Kultursponsoring zur Vermeidung von negativen Assoziationen gegenüber dem Sponsor helfen kann (vgl. Abbildung 2). Die Ergebnisse aus den Fragen zur Glaubwürdigkeit, Vergleich von Werbung und Kultur, sowie zu den Gefühlen mit einem Sponsor wurden in dem Template nicht berücksichtigt, da sie sich in den erfassten Komponeten wiederspiegeln und diese somit unterstützen.

# 4.5. Limitationen und Ausblick für weitere Forschung

Leider ergeben sich durch die Gruppe der Befragten einige Limitationen. Zum einen ist anzumerken, dass nur vier der acht Befragten persönliche Erlebnisse mit einem Sponsor vorbringen konnten. Möglicherweise sollten sich weitere qualitative Forschungen mehr auf die persönlichen Erfahrungen mit Sponsoren konzentrieren und dieses Kriterium in die Sondierungsfragen für die potenziellen Interviewpartner mit aufnehmen. Dazu konnten nur zwei männliche Personen befragt werden, eine dritte schied bei den Sondierungsfragen aus. Allerdings wurde kein Unterschied in den Antworten zwischen den männlichen und weiblichen Befragten ausgemacht. Dazu leben sieben der acht Befragten im Ballungsraum der Stadt München. Unterschiedliche Meinungen, welche auf geographischen Unterschieden beruhen, ließen sich somit leider nicht erkennen. Dafür wurde auf eine ausgewogene demographische Verteilung geachtet. Überdies ergeben sich durch den qualitativen Charakter dieser Arbeit einige Limitationen. So kann der Umfang von acht Befragten nicht als repräsentativ gelten, was allerdings nicht die Aufgabe dieser Arbeit war. Außerdem vermag diese Arbeit keine Aussagen über die Stärken verschiedener Zusammenhänge zu treffen. Daher könnten weitere Forschungen die möglichen Ursprünge für negative Gefühle gegenüber den Sponsoren kultureller Events quantitativ überprüfen sowie deren Stärke und Einfluss evaluieren. Überdies wäre es interessant zu wissen ab welcher Intensivität des Auftretens des Sponsors, das Sponsoring als zu Aggressiv aufgenommen wird.

# 5. Fazit

Obgleich es das Ziel dieser Arbeit war, mögliche Quellen negativer Gefühle gegenüber Sponsoren kultureller Events zu finden, kann doch eine tendenziell positive Wertung für das Kultursponsoring postuliert werden. Nicht nur für den Gesponserten bietet diese Strategie durch die finanzielle Unterstützung eine interessante Chance, das Event Besuchens werter zu gestalten. Überdies ist der Sponsor selbst in der Lage, unter Einhaltung gewisser Voraussetzungen, von einem positiven Imagetransfer oder einer reinen Reaktivierung seiner eigenen Werte profitieren. Außerdem wurde der positive Einfluss auf die affektive Dimension der Corporate Reputation bereits bestätigt. Zusätzlich waren die abschließenden Wertungen für Kultursponsoring in dieser Arbeit durchweg positiv. Die Befragten empfanden zum Teil Dankbarkeit für das bloße Stattfinden des Events und zum Teil Freude über niedrigere Ticketpreise, was ebenso mit den regelmäßigen Besuchen kultureller Events zusammenhängt.

Dennoch gibt es negative Aspekte, welche es insbesondere für das Kultursponsoring zu beachten gilt. Wenngleich diese Arbeit einigen Limitationen unterliegt, lässt sie trotzdem einige Tendenzen und Meinungsbilder über potenzielle Fundierungen negativer Assoziationen mit Kultursponsoring erkennen. Zum einen wurde während der Interviews der Eindruck erweckt, dass Sponsoren alltäglicher kultureller Events nur schwach in den bewussten Erinnerungen von Besuchern dieser Events bleiben. Die damit verbundene erhöhte Sensibilität gegenüber den potenziell negativen Aspekten des Sponsorings stellt nach Ansicht dieser Arbeit eine zentrale Herausforderung im Kultursponsoring dar. Dazu kommen mögliche negative Folgen eines fehlenden Fits oder eines zu aggressiven Auftritts auf dem Event. Vor allem Letzteres führt unter Umständen zu Problemen, da ein aggressiverer Auftritt unter Umständen für eine bessere Erinnerung an den Sponsor sorgt. Zuletzt bleibt noch die Befürchtung der Befragten über eine Einflussnahme des Sponsors auf das Event zu nennen.

Eine quantitative Überprüfung der möglichen Gründe könnte zu einer allgemeinen Generalisierbarkeit der Ergebnisse dieser Arbeit beitragen. Abschließend muss noch einmal postuliert werden, dass die Befragten dieser Studie Kultursponsoring durchaus positiv gegenüber stehen und einen Sponsor bereitwillig auf den von ihnen präferierten Veranstaltungen hinnehmen würden. Die finanziellen Vorteile und eine damit verbundene Verbesserung des Events überwogen im allgemeinen Stimmungsbild.

### Literatur

- Ashforth, B. E. und Mael, F. Social identity theory and the organization. *Academy of Management Review*, 14(1):20–39, 1989.
- Audi AG. Die Kunst Voraus zu Sein, 2015. URL http://www.audi.de/de/brand/de/audi-artexperience.html#page=/de/brand/de/audi-artexperience/klassik/salzburger-festspiele.html. [Stand: 01.12.2015].
- Ballen, K. America's most admired corporations. (cover story). Fortune, 125(3), 40, 1992.
- Bayerische Staatsoper. Das Partnerprogramm der Bayerischen Staatsoper, 2015. URL https://www.staatsoper.de/partner/partnerprogramm.html. [Stand: 02.11.2015].
- Causales. Wie groß ist aktuell der Anteil einzelner Finanzierungsquellen an den Gesamteinnahmen Ihrer Kultureinrichtung? (2007 bis 2013), 2013. URL http://de.statista.com/statistik/daten/studie/171390/umfrage/anteil-der-einnahmequellen-von-kultureinrichtung en/. [Stand: 02.11.2015].
- Cornwell, T. B. State of art and science in sponsorship-linked marketing. *Journal of Advertising*, 37(3):41–55, 2008.
- Cornwell, T. B., Weeks, C. S., und Roy, D. P. Sponsorship-linked marketing: Opening the black box. *Journal of Advertising*, 34(2):21–42, 2005.
- Cornwell, T. B., Humphreys, M. S., Maguire, A. M., Weeks, C. S., und Tellegen, C. L. Sponsorship-linked marketing: The role of articulation in memory. *Journal of Consumer Research*, 33(3):312–321, 2006.
- Eberl, M. und Schwaiger, M. Corporate reputation: Disentangling the effects on financial performance. *European Journal of Marketing*, 39(7/8):838–854, 2005.
- Fombrun, C. J. und Rindova, V. P. The road to transparency: Reputation management at Royal Dutch/Shell. *The Expressive Organization*, 7:7–96, 2000.
- Fombrun, C. J. und Van Riel, C. B. The reputational landscape. *Corporate Reputation Review*, 1(2):5, 1997.
- Ganassali, S. und Didellon, L. Le transfert comme principe central du parrainage. *Recherche et Applications en Marketing (French Edition)*, 11(1): 37–48, 1996.
- Gray, E. R. und Balmer, J. M. Managing corporate image and corporate reputation. Long Range Planning, 31(5):695–702, 1998.
- Keller, K. L. Conceptualizing, measuring, and managing customer-based brand equity. *The Journal of Marketing*, 57(1):1–22, 1993.
- Lafferty, B. A. und Goldsmith, R. E. Corporate credibility's role in consumers' attitudes and purchase intentions when a high versus a low credibility endorser is used in the ad. *Journal of Business Research*, 44(2):109–116, 1999.
- Madrigal, R. The influence of social alliances with sports teams on intentions to purchase corporate sponsors' products. *Journal of Advertising*, 29(4): 13–24, 2000
- Mayring, P. Qualitative Inhaltsanalyse Grundlagen und Techniken. 7. Auflage. Weinheim: Beltz Verlag., 2000.
- Mayring, P. Einführung in die qualitative Sozialforschung. 5. Auflage. Weinheim und Basel: Beltz Verlag., 2002.
- McCracken, G. Who is the celebrity endorser? Cultural foundations of the endorsement process. *Journal of Consumer Research*, 16(3):310–321, 1989.
- McDonald, C. Sponsorship and the image of the sponsor. *European Journal of Marketing*, 25(11):31–38, 1991.
- Meenaghan, T. The role of sponsorship in the marketing communications mix. *International Journal of Advertising*, 10(1):35–47, 1991.
- Meenaghan, T. und Shipley, D. Media effect in commercial sponsorship. European Journal of Marketing, 33(3/4):328–348, 1999.
- Newell, S. J., Henderson, K. V., und Wu, B. T. The effects of pleasure and arousal on recall of advertisements during the Super Bowl. *Psychology & Marketing*, 18(11):1135–1153, 2001.
- Nürnberger Symphoniker. Ein Dankeschön, 2015. URL http://www.nu ernbergersymphoniker.de/foerderer/danksagung.html. [Stand: 03.11.2015].
- O'hagan, J. und Harvey, D. Why do companies sponsor arts events? Some evidence and a proposed classification. *Journal of Cultural Economics*, 24 (3):205–224, 2000.
- Olson, E. L. und Mathias Thjømøe, H. The effects of peripheral exposure to information on brand preference. *European Journal of Marketing*, 37 (1/2):243–255, 2003.
- Ostfalia. Anteil der Sponsoringarten am Sponsoringbudget deutscher

- Unternehmen im Jahr 2013, 2013. URL http://de.statista.com/statistik/daten/studie/302384/umfrage/anteil-der-sponsoringarten-am-sponsoringbudget/. [Stand: 29.10.2015].
- Pavelchak, M. A., Antil, J. H., und Munch, J. M. The Super Bowl: An investigation into the relationship among program context, emotional experience, and ad recall. *Journal of Consumer Research*, 15(3):360–367, 1988.
- Pham, M. T. Effects of involvement, arousal, and pleasure on the recognition of sponsorship stimuli. *Advances in Consumer Research*, 19(1):85–93, 1992.
- Pham, M. T. und Vanhuele, M. Analyzing the memory impact of advertising fragments. *Marketing Letters*, 8(4):407–417, 1997.
- Quester, P. G. und Thompson, B. Advertising and promotion leverage on arts sponsorship effectiveness. *Journal of Advertising Research*, 41(1):33–47, 2001
- Rabl-Stadler, H. Editorial development, 2015. URL http://www.salzburgerfestspiele.at/hauptsponsoren/audi. [Stand: 01.12.2015].
- Rectanus, M. W. Culture incorporated: Museums, artists, and corporate sponsorships. University of Minnesota Press, 2002.
- Repucom. Sponsoringvolumen für Sport und Kultur in Deutschland von 2011 bis 2016 (in Milliarden Euro), 2015. URL http://de.statista.com/statistik/daten/studie/167709/umfrage/ausgaben-fuer-sponsoring-in-sport-und-kultur/. [Stand: 29.10.2015].
- Ruhr.2010 GmbH. Unternehmensgalerie RUHR.2010 vor Ort, 2011. URL https://www.essen.de/meldungen/pressemeldung\_542054. de.html. [Stand: 01.12.2015].
- Schaefer, A. und Keillor, B. The effective use of endorsements in advertising: The relationship between mmatch-upänd ivolvement. *Journal of Marketing Management (10711988)*, 7(2):23–33, 1997.
- Schwaiger, M. Components and parameters of corporate reputation-an empirical study. *Schmalenbach Business Review (SBR)*, 56(1):46–71, 2004.
- Schwaiger, M., Sarstedt, M., und Taylor, C. R. Art for the sake of the corporation Audi, BMW Group, DaimlerChrysler, Montblanc, Siemens, and Volkswagen help explore the effect of sponsorship on corporate reputations. *Journal of Advertising Research*, 50(1):77–90, 2010.
- Schwalbach, J. Image, Reputation und Unternehmenswert. *Transnational Communication in Europe. Research and Practice. Berlin: Vistas*, pages 287–297, 2000.
- Sparks, R. E. 11.2 the Olympics and the marketing of sports: Special issues in sponsorship. Paper presented at the Proceedings of the 1995 World Marketing Congress, 2015.
- Tajfel, H. und Turner, J. C. An integrative theory of intergroup conflict. The Social Psychology of Intergroup Relations, 33(47):74, 1979.
- The Linde Group. Engagement für die Gesellschaft. Corporate Responsibility Bericht 2012, 2013. URL http://corporateresponsibility.linde.de/cr-bericht/2012/mitarbeitergesellschaft/gesellschaftlichesengagement.html. [Stand: 2.11.2015].
- Turban, D. B. und Cable, D. M. Firm reputation and applicant pool characteristics. *Journal of Organizational Behavior*, 24(6):733–751, 2003.
- Witcher, B., Craigen, J. G., Culligan, D., und Harvey, A. The links between objectives and function in organizational sponsorship. *International Journal of Advertising*, 10(1):13–33, 1991.
- Zaichkowsky, J. L. Research Notes: The personal involvement inventory: Reduction, revision, and application to advertising. *Journal of Advertising*, 23(4):59–70, 1994.