

Sind konkrete Umweltslogans wirksamer als Greenwashing? Eine Experimentaluntersuchung zu Verbraucher-Schlussfolgerungen aus Werbebotschaften

Marc Möller-Boldt



Sind konkrete Umweltslogans wirksamer als Greenwashing? Eine Experimentaluntersuchung zu Verbraucher–Schlussfolgerungen aus Werbebotschaften

Abstract

Damit nachhaltiger Konsum möglich ist, müssen Verbraucherinnen und Verbraucher die Nachhaltigkeit von Produkten angemessen einschätzen können. Ausgehend von der Annahme, dass Greenwashing-Slogans solche Einschätzungen möglicherweise behindern, untersucht dieses zweiteilige quantitative Online-Experiment Schlussfolgerungen von Verbraucherinnen und Verbrauchern, die statische Werbeanzeigen betrachteten. Dabei wurden konkrete, d.h. quantifizierte Umweltaussagen mit vagen Botschaften (Greenwashing-Kondition) bezüglich eines fiktiven TV-Produktes bzw. TV-Herstellers verglichen. Ein Drittel der jungen und gebildeten Stichprobe (N = 163) zog nach Ansicht einer Produktanzeige, die den geringen Stromverbrauch des „Omro UHD-Fernsehers“ bewarb, den Schluss, dass weitere, in der Werbung nicht beobachtbare Umweltmerkmale eher wahrscheinlich als unwahrscheinlich sind. Unternehmensanzeigen zur Energieeffizienz der Produktion von „Nextvision“ entlockten diese Schlussfolgerung sogar 73 Prozent der Probanden. Vage und konkrete Behauptungen unterschieden sich hinsichtlich dieser sog. Halo-Effekte kaum. Dazu berechnete Indizes korrelierten signifikant positiv mit der eingeschätzten Produktqualität, was für die Wirksamkeit eines Gesamteindrucks (General Impression Halos) spricht. Dies galt insbesondere für vage Botschaften. Sie eignen sich daher besonders für Werbetreibende, weil Anzeigen in der Regel nur für kurze Zeit Aufmerksamkeit erhalten. Verbraucherinnen und Verbrauchern kann auf Basis der Ergebnisse geraten werden, stärker auf die theoretische Überprüfbarkeit von Werbeslogans zu achten.

Anmerkung: Diese Forschungsarbeit entstand auf Basis einer Masterarbeit im Studiengang „Technik- und Innovationskommunikation“ (M.Sc.). Die Betreuung führten Frau Prof. Dr. Susanne Keil sowie Frau Prof. Dr. Christine Buchholz durch.

Marc Möller-Boldt

Hochschule Bonn-Rhein-Sieg
Kontakt: m.moeller-boldt@outlook.de

ISBN 978-3-96043-096-4 (Working Paper 21/5)
DOI 10.18418/978-3-96043-096-4 (Working Paper 21/5)
IZNE Working Paper Series
ISSN 2511-0861

Bonn-Rhein-Sieg University of Applied Sciences
International Centre for Sustainable Development (IZNE)

Grantham-Allee 20

53757 Sankt Augustin / Germany

izne.info@h-brs.de

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	IV
1 Problemstellung und Ziel	1
2 Verbraucher-Schlussfolgerungen aus vager Umweltkommunikation	2
2.1 Theoretische und empirische Grundlagen zu Greenwashing	2
2.2 CSR-Halo-Effekt	4
2.3 Kompensatorische Inferenzstrategien bei Kaufentscheidungen	9
3 Methode	10
3.1 Untersuchungsdesign und Stimuli	10
3.2 Erhebungsinstrument und Messungen	13
3.2.1 Erwartete Qualität der Produkte	15
3.2.2 Halo-Effekte	16
3.2.3 Hintergrundvariablen	17
3.3 Stichprobe	20
3.3.1 Datenbereinigung	20
3.3.2 Vermessung der Stichprobe	21
3.3.3 Randomisierungskcheck	22
4 Auswertung und Interpretation der Ergebnisse	24
4.1 Auftreten von Halos (H1.1 _P und H1.1 _U)	24
4.2 Halos bei vagen und konkreten Botschaften im direkten Vergleich (H1.2 _P und H1.2 _U)	27
4.3 Korrelation zwischen Halo-Index und erwarteter Qualität (H2 _P und H2 _U)	30
5 Diskussion	34
Anhang	40
Literaturverzeichnis	56

Tabellenverzeichnis

	Seite
Tabelle 1: Experimentaldesign	12
Tabelle 2: Übersicht über die im Fragebogen verwendeten Messungen	19
Tabelle 3: Unterschiede im Produkt-Halo-Index und in den Items nach Konkretisierungsgrad	28
Tabelle 4: Unterschiede im Unternehmens-Halo-Index und in den Items nach Konkretisierungsgrad	30
Tabelle 5: Überblick über Bewährung und Ablehnung der Hypothesen	35

Hinweis: Zum Zweck der Übersichtlichkeit wird im folgenden Text das generische Maskulinum verwendet. Selbstverständlich umfassen die entsprechenden Formulierungen alle Geschlechter gleichermaßen.

1 Problemstellung und Ziel

Als „Greenwashing“ gelten Kommunikationsmaßnahmen, die Personen zu übermäßig positiven Vorstellungen über die Umweltfreundlichkeit der Praktiken oder Produkte eines Unternehmens verleiten (Lyon & Montgomery, 2015, S. 223). Die Erforschung der Wirkung von Greenwashing auf Verbraucher ist in gesellschaftlicher und praktischer Hinsicht hochgradig relevant. Nachhaltiger Konsum ist nur dann möglich, wenn Verbraucher über die Nachhaltigkeit von Produkten, Dienstleistungen und Unternehmen in angemessener Weise informiert sind (Smith, Read & Lopez, 2018, S. 1). Differenzieren diese nicht ausreichend zwischen Greenwashing und tatsächlich umweltfreundlichen Produkten, profitieren Produkte und Unternehmen, die nur scheinbar umweltfreundlich sind. In der Folge schmälert dies aus Unternehmenssicht die relative Effizienz authentischen Green Marketings, da die dahinterliegenden Nachhaltigkeitsmaßnahmen oft kostspielige Investitionen für Unternehmen darstellen. Somit besitzt Greenwashing das unrühmliche Potenzial, die Nachhaltigkeitsentwicklung von Märkten behindern zu können.

Aber nicht nur die Gesellschaft hat ein Interesse an der Erforschung von Greenwashing: Aus Sicht eines Unternehmens, das um „echte“ Nachhaltigkeit bemüht ist, ist es wichtig zu verstehen, auf welche Weise das Greenwashing der Konkurrenz auf Verbraucher wirkt, um daraufhin die Authentizität der eigenen Umweltmaßnahmen weiterhin signalisieren zu können. Ein großer Vorteil von Umweltkommunikation ist, dass diese sowohl auf Unternehmens- als auch auf Produktebene eingesetzt werden kann (Delmas & Burbano, 2011). Daher besteht aus praktischer Sicht ein Interesse an der Erforschung, ob Greenwashing und authentische Umweltkommunikation je nach Bezugsebene unterschiedlich wirken.

Während in der wissenschaftlichen Literatur Konsens darüber besteht, dass Greenwashing-Praktiken stark verbreitet sind, ist allerdings nach wie vor weitgehend unklar, wie solche Botschaften auf Verbraucher wirken (Jong, Harkink & Barth, 2018; Lyon & Montgomery, 2015). Eine weit verbreitete Praxis besteht darin, unkonkrete Umweltbotschaften für Marketingzwecke oder zur Imagepflege zu verbreiten (bspw. Baum, 2012; Segev, Fernandes & Hong, 2016). Um vor diesem Hintergrund einen Beitrag zur Beantwortung der Frage zu leisten, welche Schlussfolgerungen Verbraucher aus vager Umweltkommunikation im Unterschied zu konkreten Botschaften ziehen, berichtet diese Forschungsarbeit über ein Online-Experiment, das mit 163 Befragten durchgeführt wurde. Das folgende Kapitel widmet sich zunächst den theoretischen Grundlagen und der Hypothesenentwicklung.

2 Verbraucher-Schlussfolgerungen aus vager Umweltkommunikation

2.1 Theoretische und empirische Grundlagen zu Greenwashing

Lyon und Montgomery (2015) sichteten die bis dato bestehende Literatur zum Thema Greenwashing mit dem Ziel, eine präzise Definition des Phänomens entwickeln zu können. Sie identifizieren auf Basis einer Literaturrecherche elf verschiedene Mechanismen irreführender Kommunikation und stellen fest, dass jeder dieser Mechanismen als Greenwashing zutage treten kann (Lyon & Montgomery, 2015, S. 226–227). Die Bandbreite ist hierbei sehr groß und reicht von der selektiven Offenlegung positiver Informationen über das sog. „Astroturf Lobbying“, bei dem eine womöglich glaubwürdige „front group“ für ihren Einsatz im Gegenzug verdeckt finanziert wird, bis hin zur Verwendung betrügerischer Ökosiegel (Lyon & Montgomery, 2015, S. 227).

Entsprechend gelangen die Autoren zu dem Schluss, dass der Begriff als Sammelbezeichnung genutzt werden sollte und fordern die Forschung auf, die zahlreichen Varianten irreführenden Verhaltens von Organisationen, die theoretisch als Greenwashing bezeichnet werden können, in den Fokus zu rücken und zu systematisieren (Lyon & Montgomery, 2015, S. 244). Eine solche Systematisierung stammt von Freitas Netto, Sobral, Ribeiro und Soares (2020), die die in der Wissenschaft bislang genutzten Definitionen und Konzepte von Greenwashing herausarbeiteten. Sie kommen zu dem Schluss, dass nach wie vor keine einheitlich akzeptierte Definition von Greenwashing existiert und führen dies auf die Interdisziplinarität und Vielseitigkeit des Phänomens zurück (Freitas Netto et al., 2020, S. 10). Seele und Gatti (2017, S. 248) weisen darauf hin, dass Greenwashing im „Auge des Betrachters“ liegt und definieren Greenwashing daher als „co-creation of an external accusation toward an organization with regard to presenting a misleading green message“. Dieser Beitrag verdeutlicht weiter, dass sich Greenwashing als objektiver „Tatbestand“ nur in seltenen Fällen wie bspw. klar identifizierbaren Lügen festlegen lässt.

Die Ergebnisse zahlreicher Untersuchungen scheinen darauf hinzudeuten, dass Verbraucher Greenwashing oder wahrgenommenes Greenwashing mit einer verringerten Kaufabsicht und/oder mit negativen Einstellungen zu Produkt bzw. Marke abstrafen (z.B.: Avcilar & Demirgünes, 2016; Chen & Chang, 2013; Chen, Lin & Chang, 2014; Guo, Tao, Li & Wang, 2017; Lee, Bhatt & Suri, 2018; Nyilasy, Gangadharbatla & Paladino, 2012, 2014; Parguel, Benoît-Moreau & Larceneux, 2011; Rahman, Park & Chi, 2015; Szabo & Webster, 2020). Eine kritische Auseinandersetzung relativiert diese Erkenntnis in einigen Fällen jedoch. Bemerkenswert ist nämlich, dass viele Studien mit zahlreichen Limitationen behaftet sind, indem sie durch ihre Messungen und Fragereihenfolgen für das Thema Greenwashing

primen und/oder suggestive Szenarios enthalten (s. Jong et al., 2018, S. 82–86). Besonders kritisch ist die Messung des „Perceived Greenwashing“ zu bewerten, die von Chen und Chang (2013) entwickelt und anschließend von zahlreichen Autoren adaptiert wurde. Alle fünf Items dieser Skala (bspw. Zustimmung zur Aussage: „this product misleads with words in its environmental features“; Chen & Chang, 2013, S. 494) weisen Versuchspersonen auf das mögliche Vorhandensein von Greenwashing implizit hin und verzerren auf diese Weise die Ergebnisse. Sie sind außerdem suggestiv (s. Diekmann, 1997, S. 391–395) und lösen daher nicht die ohnehin mit dem Thema verbundene soziale Erwünschtheit auf (s. Diekmann, 1997, S. 382–386). Teilweise wird diese Messung (in ursprünglicher oder angepasster Weise) sogar *vor* anderen Outcome-Variablen gemessen (bspw. Avçilar & Demirgünes, 2016; Chen & Chang, 2013; Chen et al., 2014; Szabo & Webster, 2020), wodurch Fragereiheneffekte vermutet werden können (s. Diekmann, 1997, S. 398–399).

Unterdessen sind auch Untersuchungen auffindbar (Atkinson & Kim, 2015; Jong et al., 2018; Parguel, Benoit-Moreau & Russell, 2015; Schmuck, Matthes & Naderer, 2018; Spack, Board, Crighton, Kostka & Ivory, 2012), die auf eine teilweise Wirksamkeit von Greenwashing-Kommunikation im Sinne der praktizierenden Unternehmen hindeuten. In diesen Studien zeichnet sich ab, dass Verbraucher die von Unternehmen verwendeten Argumente nicht immer elaboriert prüfen und auch schwache oder vage Argumente akzeptieren bzw. sich von visuellen Informationen wie Naturbildern beeinflussen lassen. Spack et al. (2012) ermittelten in einer quantitativen Experimentalstudie mit Studierenden der Kommunikationswissenschaft in den USA, dass schwache Argumente für ein Waschmittel („umweltfreundlich“, „reduziert Verschmutzung“, „rettet Bäume“) hinsichtlich der wahrgenommenen Glaubwürdigkeit, Umweltfreundlichkeit und Qualität zwar schlechter abschnitten als starke Argumente („100 % biologisch abbaubare Verpackung“, „zu 100 % aus Post-Consumer-Inhaltsstoffen“, „EPA Clean Air Act Approved“), hinsichtlich der Kaufabsicht jedoch ein ähnliches Ergebnis erzielten. Schmuck et al. (2018) berichten in ihrem Forschungsbericht über zwei Untersuchungen. In Studie 1 ging es um Mineralwasser (Stichprobe quotenbasiert für USA, vage Botschaft: „Drink ALSE bottled water to help the environment. Together we can save our nature.“) und in Studie 2 um Kaffee (Stichprobe quotenbasiert für Deutschland, vage Botschaft: „Tu der Umwelt etwas Gutes.“) und Spülmittel („Der Umwelt zuliebe.“). Die vagen Claims bewirkten nicht, dass die Befragten die Anzeigen als Greenwashing bzw. irreführend wahrnahmen, obwohl die Forscher eine erweiterte Version der suggestiven Perceived Greenwashing-Skala von Chen und Chang (2013) nutzten (s.o.). Wie schon bei Spack et al. (2012) schnitten die Greenwashing-Produkte in einer niederländischen Online-Experimentaluntersuchung von Jong et al. (2018) hinsichtlich der Kaufabsicht, hier gemessen anhand von Statements, die eher die Neugier am Produkt im Sinne einer Markteinführung in den Blick nahmen, ähnlich ab wie die „grüne“ Gruppe.

Allerdings wurde Greenwashing in dieser Studie anders als bei Spack et al. (2012) manipuliert, indem hier die gleichen Umweltclaims mit unterschiedlichen Einschätzungen einer fiktiven Verbraucherzentrale als Stimuli bedacht wurden. Bemerkenswert ist in dieser Studie zusätzlich, dass die Greenwashing-Bedingung bessere Einstellungen hinsichtlich der Umweltfreundlichkeit erzeugte als eine Kontrollgruppe ohne Umweltstatements, obwohl die Irreführung in diesen Claims durch die Verbraucherzentrale (scheinbar) aufgedeckt wurde. Eine aktuelle Experimentaluntersuchung von Gosselt, van Rompay und Haske (2019) erforschte, welche Wirkung unterschiedliche Kombinationen aus externem Umweltsiegel und internem Claim, die den 158 niederländischen Befragten innerhalb einer einzigen Anzeige für Kaffee gezeigt wurden, auf Einstellungen gegenüber dem Unternehmen haben. Dieses realistische Beispiel zeigte, dass Verbraucher auch bei einem lediglich internen Claim, der nicht durch ein externes Label verifiziert wird, intrinsische Motive und damit Aufrichtigkeit vermuten.

Während sich die meisten Untersuchungen also mit der Formation von übergreifenden Einstellungen bzgl. eines Produktes oder eines Unternehmens befassen, sind faktische Schlussfolgerungen, die Verbraucher aus der Wahrnehmung von Claims ziehen können, weniger häufig Gegenstand der Forschung. Nach Wahrnehmung kurzer Werbeslogans unbekannter Unternehmen oder Produkte sind Verbraucher jedoch zunächst darauf angewiesen, abzuschätzen, welche Eigenschaften das Werbeobjekt außer den beworbenen innehat. Aus diesem Grund wird in den kommenden Kapiteln auf dazu passende Theorien eingegangen, zu denen Hypothesen für den empirischen Teil dieser Forschungsarbeit formuliert werden.

2.2 CSR-Halo-Effekt

Die Forschung rund um den sogenannten Halo-Effekt nimmt explizit auf die Urteilsbildung unter Bedingungen unvollständiger Informationen Bezug. Unter Verweis auf Cooper (1981), Kahneman (2012) sowie Nisbett und Wilson (1977) beschreiben Smith et al. (2018, S. 1) den Halo-Effekt als Tendenz, einen Eindruck von einer Person oder einem Objekt auf Basis einiger weniger positiver oder negativer Fakten zu gewinnen und diesen Eindruck in der Folge als Anker zur Beurteilung anderer Fakten über die Person oder das Objekt zu nutzen. Beispielsweise könnte ein Verbraucher, nachdem er vom ökologischen Anbau eines Lebensmittelherstellers erfahren hat, darauf schließen, dass dieser auch ausschließlich erneuerbare Energien nutzt (s. Lyon & Montgomery, 2015, S. 228). Cooper (1981, S. 219) stellt heraus, dass der Begriff „Halo-Effekt“ auf zwei verschiedene Arten genutzt wird, die auf unterschiedliche Ursprünge des beschriebenen Phänomens hinweisen. In der ursprünglichen Form – so auch bei Wells (1907) und Thorndike (1920) – wird davon ausgegangen, dass ein genereller Eindruck die Bewertungen einzelner Unterkategorien beeinflusst. Eine andere Verwendung des Begriffs liegt

bei der Vermutung vor, dass saliente, d.h. markante/auffällige Merkmale die Bewertung (scheinbar) zusammenhängender Kategorien beeinflussen. Boatwright, Kalra und Zhang (2008, S. 217) unterscheiden diese beiden Typen mit Hilfe der Begriffe *General Impression Halo* sowie *Interdimensional Similarity Halo*. Die Autoren schreiben dem *Interdimensional Similarity Halo* im Marketing eine größere Bedeutung zu als dem *General Impression Halo*, da es aus Marketingperspektive relevanter sei, die Prozesse, die zur Formation von Gesamtbewertungen führen, nachzuvollziehen. Ein umgekehrtes Argument lässt sich für die Unternehmenskommunikation bzw. *Corporate Social Responsibility* (CSR) ergänzen, da hier das Image und damit ein Gesamteindruck des Unternehmens im Idealfall auf die Bewertung einzelner Produkte zurückfallen kann (s. hierzu Chernev & Blair, 2015). Wie Cooper (1981, S. 219) allerdings zurecht anmerkt, können die beiden unterschiedlichen Konzepte bzw. Ursprünge des Halo-Effekts empirisch kaum auseinandergehalten werden. Nichtsdestotrotz ist diese Unterscheidung aus theoretischer Sicht notwendig und sollte für die Interpretation der Ergebnisse dieser Forschungsarbeit im Hinterkopf behalten werden.

Diese theoretischen Grundlagen lassen sich in Anwendung auf den Themenbereich CSR- und Umweltkommunikation erweitern. So kann nicht grundsätzlich davon ausgegangen werden, dass Halo-Effekte auf Produktebene in gleicher Weise funktionieren wie auf Unternehmensebene. Smith et al. (2018, S. 5) vermuten abgeschwächte Halo-Effekte auf Unternehmensebene, da (1) die Produktfunktionalität und die Konsumerfahrung von CSR auf Unternehmensebene weniger beeinflusst werden als von der Umweltfreundlichkeit der Produkte selbst, (2) weil bezüglich Unternehmen eine potenziell größere öffentliche Skepsis vorhanden sein könnte als gegenüber einzelnen Produkten und (3) weil Verbraucher auf Unternehmensebene stärker an Widersprüche zwischen verschiedenen CSR- bzw. Nachhaltigkeitsbereichen gewöhnt sein und daher sensibler auf entsprechende Informationen reagieren könnten.

Halo-Effekte wurden bei Verbrauchern sowohl auf Unternehmens- als auch auf Produktebene anhand zahlreicher Beispiele ermittelt (Smith et al., 2018, S. 2–3). Besonders prominent sind demnach Untersuchungen, die thematisieren, welche unbegründeten Schlussfolgerungen Verbraucher aus Gesundheitsinformationen zu Lebensmitteln und entsprechenden Claims ziehen. Es existieren allerdings auch einige Studien, die einen Halo-Effekt feststellten, der sich ausgehend von der Unternehmensreputation positiv auf Kaufentscheidungen auswirkt (für einen Überblick siehe Burke, Dowling & Wei, 2018). Dieser Forschungsstrang ist somit in der Lage zu beschreiben, wann und wie Verbraucher von vorhandenen Informationen auf fehlende Informationen schließen (Boatwright et al., 2008; Kardes, Posavac & Cronley, 2004). Smith et al. (2018, S. 4–5) bieten einen Überblick über die Forschung zu

Halo-Effekten, die von CSR-Kommunikation auf Produktebene ausgehen können. Hierbei wird deutlich, dass Verbraucher je nach Kontext sowohl positive als auch negative Halo-Beurteilungen aus CSR-Informationen ziehen, beispielsweise bei der Bewertung der Qualität des Produktes. Die Autoren selbst führten zwei Studien durch, um den Halo-Effekt im Rahmen von Umweltkommunikation zu erforschen. In beiden Untersuchungen wiesen sie den Befragten eine von zwei profilartigen Beschreibungen von Fast-Food-Restaurants zu. Eine Beschreibung charakterisierte ein Restaurant als sehr umweltfreundlich, während dem anderen Restaurant nur einfache, aber nicht unbedeutende Umweltmaßnahmen zugeschrieben wurden. Die Probanden sahen beide Beschreibungen, sollten aber nur eines der Restaurants in der Folge des Experiments bewerten. Den Stimuli nach zu urteilen war Greenwashing somit nicht Teil dieser Untersuchung, sondern lediglich verschiedene Niveaus der Umweltfreundlichkeit. In Studie 1 bewerteten 297 US-Inländer die Wahrscheinlichkeit, dass das jeweilige Unternehmen neben den im Profil beschriebenen noch weitere CSR-Aktivitäten in den Dimensionen Umwelt, Mitarbeiter, Kunden und Gemeinschaft verfolgt. Zudem sollten die Befragten die Performance des Unternehmens bei diesen CSR-Aktivitäten einschätzen. Studie 2 verwendete die gleichen Beschreibungen und befragte 521 US-Inländer, maß aber die vier CSR-Bereiche anhand von Indizes, die insg. 17 CSR-Unterdimensionen umfassten; so wurden für einen Umwelt-Halo-Index bspw. fünf Bewertungsunterdimensionen (z.B. „Recyclen von Altspeiseöl zu Biodiesel“) erstellt. Obwohl den Befragten in den Restaurantbeschreibungen sowohl in Studie 1 als auch in Studie 2 (mit einer Ausnahme¹) zu keiner der Bewertungsdimensionen Informationen vorlagen, wurde das Unternehmen mit den tiefgreifenderen Umweltmaßnahmen auch hinsichtlich dieser weiteren CSR-Dimensionen stets besser bewertet. Es wurden also entgegen der Vermutung der Autoren (s.o.) stabile Halo-Effekte auf Unternehmensebene, die von einer Beschreibung der Umweltaktivitäten eines in bestimmten Bereichen umweltfreundlichen Restaurants ausgingen, ermittelt. Insofern ist es denkbar, dass Verbraucher in einer für Marketing und PR typischen Kommunikationssituation, in der sie nur über die Informationen der Hersteller selbst verfügen, zu positiven Urteilen gegenüber der Umweltfreundlichkeit der Greenwashing-Produkte gelangen, sofern keine Ambiguität durch Informationen anderer Hersteller besteht.

Offen bleibt damit auch nach den bisherigen theoretischen und empirischen Erwägungen, ob und inwiefern vage Werbebotschaften unterschiedlich stark Halo-Effekte im Vergleich zu konkreten Botschaften auslösen. Um Hypothesen hierzu formulieren zu können, ist eine genauere Betrachtung der psychologischen Mechanismen, die Halo-Effekte generell in gegenstandsübergreifender Hinsicht

¹ Ausnahme: die generelle Frage „How likely is it that [Company] is taking action to ensure a clean environment“ in Studie 1, da die Beschreibungen in diesen Bereich fallen, s. Smith, Read und Lopez (2018, S. 7).

begünstigen, notwendig. Cooper (1981) unterscheidet sechs Ursprünge für Halo-Effekte. Zwei dieser Ursprünge betreffen vage Botschaften besonders und sprechen dafür, dass diese ein größeres Potenzial für Halo-Effekte haben als konkrete Botschaften. Der Begriff *Undersampling* (dt. etwa: „unzureichende Auswahl“) bezieht sich auf unzureichende Informationen, die den Bewertenden zu dem Bewertungsgegenstand vorliegen. Dieser Informationsmangel zwingt sie auf diese Weise zu einer größeren Abhängigkeit von einem Gesamteindruck (Cooper, 1981, S. 220). Im Falle von Werbeanzeigen kann argumentiert werden, dass auch eine einzelne konkrete Botschaft in den Bereich des *Undersampling* fällt. Nichtsdestotrotz liegen durch die Konkretisierung von Slogans i.d.R. entscheidende Informationen für die Einordnung der Botschaft in einen größeren Kontext vor, sodass vage Botschaften tendenziell anfälliger für *Undersampling* sein könnten. *Engulfing* (dt.: „Verschlingen“) beschreibt das Zusammenspiel von den Beobachtungen, die die Bewertenden vom Bewertungsgegenstand machen, und den Bewertungskategorien. Es ist dann einschlägig, wenn die Bewertenden glauben, dass die Bewertungskategorien mit einem Gesamteindruck oder besonders auffälligen („salienten“) Eigenschaften kovariieren (Cooper, 1981, S. 220). Somit sind diese beiden Ursprünge eng verwandt. Beide lassen sich theoretisch in ihrem Einfluss reduzieren, indem mehr Informationen über den Bewertungsgegenstand zur Verfügung gestellt werden, die den Bewertenden mehr Möglichkeiten bieten, Informationen wahrzunehmen, die gegen eine Kovarianz sprechen und so den Einfluss eines Gesamteindrucks zu reduzieren (Cooper, 1981, S. 230). Genau dieser Mechanismus könnte durch konkretere Informationen ausgelöst werden, indem zwar nicht unbedingt die Informationsquantität, aber sehr wohl die Informationsqualität erhöht wird. Konkrete Informationen steigern nämlich die Vertrautheit der Bewertenden mit dem Bewertungsgegenstand (s. hierzu Cooper, 1981, S. 230–231). Auf diese Weise ist es also denkbar, dass vage Botschaften ein größeres Potenzial für Halo-Effekte haben, indem sie durch ihre generellen Formulierungen mehr Interpretationsspielraum zulassen und auf diese Weise Widersprüche zwischen den Beobachtungen des Bewertungsgegenstands und den Bewertungskategorien sowie zwischen den einzelnen Bewertungskategorien verschleiern.

Alternativ kann in einer simplen, intuitiven Argumentation jedoch auch vermutet werden, dass konkrete, belastbare Botschaften glaubwürdiger sind und somit das von Verbrauchern wahrgenommene Bewertungsrisiko bei der Einschätzung weiterer Dimensionen (s. hierzu Boatwright et al., 2008) reduzieren. Diese Überbewertung eines einzelnen Aspektes könnte den sog. *Interdimensional Similarity Halo* begünstigen, der dann anzunehmen ist, wenn Personen einen Gegenstand über mehrere Dimensionen hinweg auf Basis von Informationen bezüglich einer Dimension ähnlich bewerten (Boatwright et al., 2008, S. 217). Eine dritte Möglichkeit ist folglich, dass beide Arten von Botschaften mit jeweils idiosynkratischer Wirkweise einen Halo-Effekt erzeugen. Die vorliegende Forschungsarbeit folgt

dieser Ansicht. Allerdings wird postuliert, dass die konkreten Botschaften in einem direkten Vergleich mit den vagen Botschaften stärkere Halo-Effekte auslösen, weil das Bewertungsrisiko unter den hier erforschten Bedingungen geringer Information für Verbraucher sehr hoch ist und diese daher besonders sensibel auf eine Reduktion desselbigen reagieren müssten.

RQ_P: Inwiefern schlussfolgern Verbraucher nach Wahrnehmung eines konkreten bzw. vagen Umweltarguments auf Produktebene, dass das Produkt noch weitere Umwelteigenschaften aufweist (*positiver Halo-Effekt*)?

H1.1_P: Sowohl nach Wahrnehmung eines konkreten Umweltarguments auf Produktebene als auch nach Wahrnehmung eines vagen Umweltarguments auf Produktebene schlussfolgern Verbraucher, dass das Produkt noch weitere Umwelteigenschaften aufweist.

H1.2_P: Nach Wahrnehmung eines konkreten Umweltarguments auf Produktebene schlussfolgern Verbraucher stärker als nach Wahrnehmung eines vagen Umweltarguments auf Produktebene, dass das Produkt noch weitere Umwelteigenschaften aufweist.

RQ_U: Inwiefern schlussfolgern Verbraucher nach Wahrnehmung eines konkreten bzw. vagen Umweltarguments auf Unternehmensebene, dass das Unternehmen noch weitere Umweltinitiativen ergreift (*positiver Halo-Effekt*)?

H1.1_U: Sowohl nach Wahrnehmung eines konkreten Umweltarguments auf Unternehmensebene als auch nach Wahrnehmung eines vagen Umweltarguments auf Unternehmensebene schlussfolgern Verbraucher, dass das Unternehmen noch weitere Umweltinitiativen ergreift.

H1.2_U: Nach Wahrnehmung eines konkreten Umweltarguments auf Unternehmensebene schlussfolgern Verbraucher stärker als nach Wahrnehmung eines vagen Umweltarguments auf Unternehmensebene, dass das Unternehmen noch weitere Umwelteigenschaften aufweist.

Insgesamt postulieren die Hypothesen H1.2_P und H1.2_U also, dass konkrete Umweltargumente „besser“ bei den Verbrauchern abschneiden als die vagen Claims. Allerdings sei im Rückblick auf den Forschungsstand erwähnt, dass die theoretische Herleitung zu dieser Vermutung nicht als gesichert betrachtet werden kann. Die beiden Hypothesen haben damit notwendigerweise in gewisser Hinsicht einen explorativen Charakter, da die theoretischen Erwägungen und der Forschungsstand in Kombination kein eindeutiges Bild zeichnen: Die bisherige Forschung zu Halo-Effekten bezog sich nicht explizit auf vage Umweltkommunikation, während die bisherige Forschung zu Greenwashing und vager Umweltkommunikation keine Messungen etwaiger Halo-Effekte vornahm. Außerdem sei nochmals

darauf hingewiesen, dass die Theorie zu den Ursprüngen des Halo-Effekts (Cooper, 1981) entgegen der explorativen Hypothesen dafür spricht, dass vage Botschaften eine größere Tendenz zu Halo-Effekten auslösen.

2.3 Kompensatorische Inferenzstrategien bei Kaufentscheidungen

Ein bisher noch nicht angeführter Aspekt besteht in der Fragestellung, welche Auswirkungen die Wahrnehmung der Umweltfreundlichkeit von Produkten auf die Bewertung der Qualität hat. Gerade für Unternehmen aus Technologiesektoren ist es zentral, dass potenzielle Kunden einen positiven Eindruck hinsichtlich der Funktionalität ihrer Produkte erlangen. Aus diesem Grund ist diese Fragestellung für die gewählte Spezialisierung auf Technikkommunikation hochgradig relevant.

Es existieren empirische Arbeiten, die ermitteln konnten, dass Verbraucher die Qualität umweltfreundlicher Produkte unter Umständen als schlechter im Vergleich zur Konkurrenz einschätzen, da sie im Herstellungsprozess Kompromisse zwischen Umweltfreundlichkeit und Funktionalität vermuten (für eine Übersicht siehe Pancer, McShane & Noseworthy, 2017). Die hinter diesen Kompromissvermutungen stehenden Mechanismen in der Entscheidungsbildung der Verbraucher sind mit den oben beschriebenen CSR-Halos hinsichtlich der Ausgangslage verwandt. Auch hierbei geht es um Schlussfolgerungen, die die Verbraucher zu nicht beobachtbaren Eigenschaften ziehen. Im Falle der Produktqualität sind jedoch Inferenzen zulasten der beworbenen Produkte wahrscheinlicher. Im Unterschied zu den oben beschriebenen CSR-Halos, die sich auf die interne Konsistenz von Produkt- und Unternehmenseigenschaften beziehen, tritt bei der Beurteilung der Produktqualität das Verständnis der Verbraucher vom kompetitiven Markt hinzu, sodass diese von der Beobachtung der Überlegenheit hinsichtlich einer Kategorie auf die Unterlegenheit hinsichtlich einer anderen schließen können. Chernev und Carpenter (2001, S. 350) bezeichnen diese Vorgehensweise der Verbraucher als *Compensatory Inference Strategy*. CSR-Halos entstehen dagegen entweder durch die sog. *Evaluative Consistency Strategy*, wenn von einem Gesamteindruck auf einzelne, nicht beobachtbare Eigenschaften geschlossen wird oder aber durch die *Probabilistic Consistency Strategy*, wenn die Versuchspersonen von einer Korrelation zwischen der beobachtbaren Eigenschaft und den nicht beobachtbaren Eigenschaften ausgehen² (für eine überblicksartige Unterscheidung der Strategien s. Chernev & Carpenter, 2001, S. 349–350). Bei Versuchspersonen, die einem CSR-Halo-Effekt unterliegen, könnte die kompensatorische Inferenzstrategie und die damit verbundene Abwertung der wahrgenommenen

² In dieser Forschungsarbeit werden für die beiden Phänomene weiterhin die Begriffe *General Impression Halo* und *Interdimensional Similarity Halo* verwendet, da der Begriff der „Strategie“ ein intentionales Verhalten der Verbraucher unterstellt.

Produktqualität erst recht auftreten, da diese Verbraucher von einer Umweltfreundlichkeit hinsichtlich *mehrerer* Dimensionen ausgehen. Auf Basis der beschriebenen Vermutung eines Kompromisses zwischen Umweltfreundlichkeit und Produktqualität, der die Produktbeurteilungen von Verbrauchern lenkt, kann daher ein negativer Zusammenhang zwischen dem Auftreten von CSR-Halo-Effekten und der erwarteten Produktqualität postuliert werden. Fraglich ist, ob Verbraucher die kompensatorische Inferenzstrategie auch auf Unternehmensebene anwenden. So ist es denkbar, dass Verbraucher die Umweltinitiativen auf Unternehmensebene als von den Produkten entkoppelt wahrnehmen, sodass hier keine Abwertung der erwarteten Produktqualität auf Basis der CSR-Halo-Effekte erfolgt. Einige empirische Arbeiten (Chernev & Blair, 2015; Luchs, Naylor, Irwin & Raghunathan, 2010; Newman, Gorlin & Dhar, 2014) stützen diese Argumentation.

Formuliert man Hypothesen zum Zusammenhang zwischen den beiden Variablen, ist zu beachten, dass unabhängige und abhängige Variable im Sinne einer Kausalität nicht unterschieden werden können. Die oben dargelegte Theorie zur *Compensatory Inference Strategy* sieht zwar vor, dass der Eindruck der Umweltfreundlichkeit zu einer Abwertung der erwarteten Produktqualität führt. Allerdings fanden beispielsweise Chernev und Blair (2015) heraus, dass auch von CSR-Aspekten wie der Wohltätigkeit eines Unternehmens Halo-Effekte ausgehen können, die die Einschätzung der Produktqualität begünstigen. Auch wenn diese Forschungsarbeit eine Replikation dieses Befundes angesichts der idiosynkratischen Bedingungen des Untersuchungsgegenstands Technik (s.o.) nicht erwartet, muss diese mögliche Kausalität berücksichtigt werden. Insofern kann in dieser Forschungsarbeit nur die Richtung (positiv oder negativ) und Stärke des Zusammenhangs zwischen CSR-Halo und erwarteter Produktqualität ermittelt werden, es können jedoch nicht Prädiktor und Kriterium unterschieden werden. Für Produkt- und Unternehmensebene werden somit auf Basis der theoretischen Argumentation die folgenden Hypothesen formuliert:

H2_p: Der Halo-Index auf Produktebene hängt negativ mit der erwarteten Produktqualität zusammen.

H2_u: Der Halo-Index auf Unternehmensebene hängt nicht negativ mit der erwarteten Qualität der Produkte des Unternehmens zusammen.

3 Methode

3.1 Untersuchungsdesign und Stimuli

Die oben angeführten Hypothesen dieser Forschungsarbeit implizieren vermutete Wirkungen, die auf kommunikative Botschaften als Ursache zurückgeführt werden sollen. Zur Erforschung solcher

Ursache-Wirkungs-Beziehungen muss die Kommunikationswissenschaft auf Experimente zurückgreifen (Koch, Peter & Müller, 2019, S. 1). Dies bedeutet, dass zur Annahme der aufgestellten Hypothesen ein systematischer Zusammenhang ermittelt werden muss, wobei die erforderliche zeitliche Abfolge von Ursache und Wirkung beachtet wird und alternative Erklärungen, die den systematischen Zusammenhang begründen könnten, ausgeschlossen werden (Koch et al., 2019, S. 2). Diesen Anforderungen kann nur durch eine entsprechende methodische Planung und Operationalisierung Rechnung getragen werden. Dieses und die folgenden Unterkapitel legen die Methode dieser Forschungsarbeit aus diesen Gründen detailliert dar.

Das Untersuchungsdesign lässt sich wie folgt beschreiben: Im Verlaufe des Experiments durchliefen die Versuchspersonen zwei Stimuli. Einer der beiden Stimuli beinhaltete eine Botschaft auf Produktebene, der andere eine Botschaft auf Unternehmensebene. Die Versuchspersonen wurden mit Hilfe des webbasierten Experiments vor Ansicht der jeweiligen Stimuli randomisiert den Gruppen „vage Umweltbotschaft“ bzw. „konkrete Umweltbotschaft“ im Sinne eines Between-Subjects-Designs (s. Koch et al., 2019, S. 68–69) zugewiesen. Die Zuweisung zu diesen Experimentalgruppen erfolgte für jeden Stimulus separat, also sowohl vor dem Stimulus auf Produktebene als auch vor dem Stimulus auf Unternehmensebene. Um Positionseffekte zu neutralisieren, die von der Abfolge der Bezugsebene der Botschaft (Produkt- bzw. Unternehmensebene) ausgehen können, werden die beiden Messzeitpunkte rotiert. Zusammengefasst beinhaltet die Studie also einen Between-Subjects-Faktor (vage vs. konkrete Botschaften), der in zwei Experimenten bei den gleichen Versuchspersonen variiert wird (s. Tab. 1).

Die zentralen Bausteine eines jeden Experiments sind die Stimuli, in denen die Experimentalvariation integriert wird. Sie müssen realistisch und im Forschungsinteresse relevant sein. Vor allem dürfen sich die Stimuli, die statistisch verglichen werden sollen, nur hinsichtlich der Experimentalvariation unterscheiden. Im Folgenden wird aus diesen Gründen die Konzeption der Stimuli, die der Forschungsarbeit in den Anhängen 1 und 2 beigefügt sind, erläutert.

Tabelle 1*Experimentaldesign*

Between-Subjects: Konkretisierung	Experimentalgruppen: Produktebene	Experimentalgruppen: Unternehmensebene
Konkret	EG1 _p	EG1 _u
Vage	EG2 _p	EG2 _u

Anmerkung. Die Zuweisung der Versuchspersonen zu den Experimentalgruppen erfolgte für beide Ebenen separat, sodass im Datensatz alle Kombinationen von vagen und konkreten Botschaften über die Ebenen hinweg enthalten sind (inkl. vage/vage und konkret/konkret).

Für den Stimulus eignet sich Werbung für Fernseher besonders, da TV-Geräte einerseits für nahezu alle Verbraucher relevant sind und sich andererseits hinsichtlich ihrer Umweltverträglichkeit stark unterscheiden (Umweltbundesamt, 2020). Nichtsdestotrotz muss die Interpretation der Ergebnisse der Studie vor dem Hintergrund der Erkenntnis, dass sich die Wirkung von Umweltkommunikation nicht über mehrere Branchen hinweg generalisieren lässt (Torelli, Balluchi & Lazzini, 2020), erfolgen.

Dem Experimentaldesign folgend sind insgesamt zwei vage und zwei konkrete Umweltargumente notwendig. Für die Stimuli eignet sich der Energieverbrauch als Kommunikationsthema, da dieser bei Fernsehern sowohl auf Unternehmensebene in der Produktion als auch auf Produktebene im Rahmen des Jahresstromverbrauchs relevant ist. Eine Quantifizierung bzw. eine in Zahlen gefasste Präzision des Energieverbrauchs in der Produktion oder beim Gerätebetrieb kann als Konkretisierung einer Umweltbotschaft aufgefasst werden, weil dadurch die notwendige Grundlage geschaffen wird, um Bedeutung und Ausmaß einer entsprechenden Umweltinitiative bzw. Produkteigenschaft einordnen zu können. Ausgehend von dieser Argumentation wurden für die Produktebene die folgenden Botschaften entworfen: *Punktet mit erstklassiger Energieeffizienz. Der neue 48“ Omro UHD-Fernseher von Zensoric.* (Produktebene vage), *Nur 66 kWh Jahresstromverbrauch. Der neue 48“ Omro UHD-Fernseher von Zensoric.* (Produktebene konkret), *Konsequent energiesparende Herstellung von Hightech-TV-Geräten. Dafür steht NEXTVISION.* (Unternehmensebene vage) sowie *22% Energieeinsparung bei der Herstellung von Hightech-TV-Geräten seit 2017. Dafür steht NEXTVISION.* (Unternehmensebene konkret). Die vagen Botschaften wurden bewusst vehement formuliert, um den Spielraum, der einem Unternehmen bei der Konzeption eines vagen Slogans auch in der Realität zur Verfügung steht, auszunutzen. Die Botschaften können im Hinblick auf einige Inhaltsanalysen, die sich unter anderem mit vagen Umweltbotschaften befasst haben, eingeordnet werden: Banerjee, Gulas und Iyer (1995) würden die hier als

vage aufgefassten Botschaften wohl als nur moderat konkretisiert einordnen, Schmidt und Donsbach (2012) würden sie womöglich als unvollständig oder vielleicht sogar als unpräzise bezeichnen. Insofern sind die hier als vage verstandenen Botschaften hinsichtlich der inhaltsanalytischen Literatur anschlussfähig.

Diese Botschaften wurden als Slogan in statische Werbeanzeigen eingefügt (s. Anhänge 1 und 2), für die ein Säulenplakatformat (4/1 Bogen im Hochformat) gewählt wurde. In den Anleitungstexten im Fragebogen wurden diese lediglich als „Werbeanzeigen“ vorgestellt (s. Anhang 3). Neben der Botschaft enthielten die Werbeanzeigen jeweils ein fiktives Markenlogo, wobei die Marke „Zensoric“ für die Anzeigen auf Produktebene und die Marke „Nextvision“ für die Anzeigen auf Unternehmensebene genutzt wurde. Die Benutzung fiktiver Marken schließt aus, dass die Versuchspersonen vorher bestehende Einstellungen in die Bewertung des Produktes und des Unternehmens einfließen lassen und ist somit als Teil der Drittvariablenkontrolle zu verstehen. Die übrigen Gestaltungselemente (Produktpräsentation bei den Anzeigen auf Produktebene, Fahrradpiktogramme bei den Anzeigen auf Unternehmensebene) wurden mit dem Ziel erstellt, die Anzeigen möglichst realistisch erscheinen zu lassen³. Für beide Kommunikationsebenen wurde ein grüner Hintergrund entworfen, auf Produktebene in einer Blattoptik und auf Unternehmensebene in einer Wischoptik. Insofern enthalten die Anzeigen neben den Botschaften auch „executional elements (...) evoking nature“ (Parguel et al., 2015, S. 111); diese sind allerdings nicht Teil der Experimentalvariation.

3.2 Erhebungsinstrument und Messungen

Die Erhebung der Daten erfolgte anhand eines Fragebogens (s. Anhang 3), der mit Hilfe des browserbasierten Onlinefragebogentools SoSci Survey (Leiner, 2020) erstellt wurde. Der Fragebogen wurde dabei im Willkommenstext als Verbraucher*innen-Studie beschrieben, in der „junge Unternehmen und ihre Produkte nach einem ersten Eindruck“ bewertet werden sollen. Ein Hinweis auf die Umweltthematik hätte hier bereits zu einer Selektivität in der Probandenrekrutierung geführt. Die Ansprache der Versuchspersonen als Verbraucher*innen versetzt diese in eine aktivere Rolle als dies bspw. bei einem *Framing* als „Marketingstudie“ der Fall gewesen wäre. Nach dieser Einführung in den Fragebogen wurde den Versuchspersonen in Rotation zunächst entweder eine der beiden Anzeigen auf Produktebene oder eine der beiden Anzeigen auf Unternehmensebene vorgestellt. Sowohl nach der Anzeige auf Produktebene als auch nach der Anzeige auf Unternehmensebene wurden

³ Dem Verfasser dieser Forschungsarbeit ist bewusst, dass bei den Produktanzeigen gemäß §6a EnVKV in der Realität die Energieeffizienzklasse angegeben werden müsste. Da unabhängige Siegel jedoch nicht Untersuchungsgegenstand dieser Forschungsarbeit sind, wurde zwecks Reduktion eines weiteren Faktors darauf verzichtet.

Zustimmungsbatterien zur erwarteten Qualität der Produkte und zum Halo-Effekt gestellt. Diese insgesamt vier Batterien wurden zusätzlich durch die folgende Anweisung eingeleitet: „Bitte legen Sie nun Ihren ersten Eindruck des Fernsehers (bzw. Unternehmens) dar. Entscheiden Sie nach Ihrem Bauchgefühl, das Sie durch die Werbeanzeige gewonnen haben.“ Da diese Fragebatterien für die Hypothesen unabdingbar sind, konnten die Versuchspersonen die Beantwortung nicht umgehen. Bei allen anderen Variablen (s.u.) wurde kein Antwortzwang ausgeübt. Innerhalb aller Fragebatterien mit Ausnahme der Soziodemografie sowie der „Weiß nicht“-Angaben im Wissenstest wurden die Items in zufälliger Reihenfolge ausgespielt. Vor der Erhebung wurde der Fragebogen einem umfangreichen technischen Funktionstest, der den Empfehlungen von SoSci Survey folgte (SoSci Survey GmbH, 2019), unterzogen.

Im Messinstrument wurden für einige latente, d.h. nicht direkt beobachtbare Konstrukte (erwartete Qualität, Interesse an Fernsehern, Umweltanliegen) Skalen und für andere latente Konstrukte (Halo-Effekt, Umweltwissen) Indizes verwendet. Skalen entsprechen sogenannten reflektiven Messmodellen, während einem Index stets ein formatives Messmodell zugrunde liegt (für eine Übersicht s. Christophersen & Grape, 2009). Bei Skalen spiegeln die einzelnen Items, die im Fragebogen als Indikatoren des latenten Konstrukts gemessen werden, allesamt das latente Konstrukt wider. Christophersen und Grape (2009, S. 104) folgend lassen sich die latenten Konstrukte als *Ursache* der beobachtbaren Indikatoren bezeichnen, was impliziert, dass eine Veränderung des Konstrukts durch eine Veränderung in *allen* Indikatoren reflektiert wird. Dementsprechend muss die interne Konsistenz dieser Skalen überprüft werden. Alle innerhalb einer Skala genutzten Items müssen in angemessener Weise miteinander korrelieren, damit angenommen werden kann, dass sie reliabel das latente Konstrukt widerspiegeln. In der vorliegenden Forschungsarbeit wird dies anhand von Reliabilitätsanalysen überprüft, mit denen der Koeffizient Cronbachs Alpha berechnet wird (Field, 2018, S. 821–830). Gegebenenfalls werden Items aus der Auswertung ausgeschlossen, sofern sich dadurch die interne Konsistenz der Skalen bedeutend verbessert. Für die Interpretation des Reliabilitätskoeffizienten wird die Diversität des latenten Konstrukts (hierdurch sinkt der Koeffizient natürlicherweise) und die Anzahl der bestehenden Items (hierdurch steigt selbiger natürlicherweise) herangezogen (Field, 2018, S. 823).

Ein Index unterliegt dagegen einer anderen Logik, indem das latente Konstrukt als „gewichtete Zusammensetzung seiner Indikatoren“ betrachtet wird (Christophersen & Grape, 2009, S. 105). Hier ist die Eignung somit nach theoretischen Gesichtspunkten zu bemessen, indem die Indikatoren in der Summe möglichst facettenreich das latente Konstrukt abbilden (Christophersen & Grape, 2009,

S. 111). Die Berechnung der internen Konsistenz ist ausgeschlossen, da eine Korrelation zwischen den Indikatoren nicht notwendig ist (Christophersen & Grape, 2009, S. 106). Im Folgenden wird noch einmal detailliert auf die einzelnen Messungen eingegangen.

3.2.1 Erwartete Qualität der Produkte

Erwartete Qualität der Produkte (Produktebene). In Anlehnung an Pancer et al. (2017, S. 165) und Newman et al. (2014, S. 826) wurde die erwartete Qualität so erhoben, dass jedes Item im Sinne des reflektiven Messmodells (s.o.) die Leistungsfähigkeit des Produktes widerspiegelt. Entsprechend wurden nur Items gewählt, die eine Eigenschaft enthalten, die zentral für die Leistungsfähigkeit von Fernsehern sind oder auf die Produktqualität als Ganzes abzielen. Jeweils siebenstufig schätzten die Versuchspersonen die qualitativen Mängel [umgepolt], den Kontrastumfang, die generelle Qualität, die Konkurrenzfähigkeit, die Vielseitigkeit und die Bildqualität des Fernsehers ein. Durch die Positionierung dieser Skala vor der Erhebung der Halo-Effekte (s.u.) wird ein Reihenstellungseffekt verhindert, der einen Fehler 1. Art hinsichtlich der Hypothesen H2_p und H2_u begünstigen würde.

Reliabilitätsanalysen offenbarten, dass sich der Reliabilitätskoeffizient Cronbachs α durch Entfernen des Items „Der Omro UHD-Fernseher hat qualitative Mängel“ [umcodiert] von einem bereits guten Niveau (Cronbachs $\alpha = .81$) noch einmal deutlich auf ein Cronbachs α von $.85$ erhöhen ließ. Das Item wies nur eine geringe korrigierte Item-Skala-Korrelation von $.19$ auf. Entsprechend wurde die Skalenberechnung für die Variable „erwartete Qualität der Produkte“ auf Produktebene ohne dieses Item durchgeführt, sodass die Skala aus den fünf verbleibenden Items besteht. Die hohe interne Konsistenz weist darauf hin, dass die Skalenkonstruktion und der Zuschnitt der Items auf Fernsehgeräte im Speziellen erfolgreich waren.

Erwartete Qualität der Produkte (Unternehmensebene). Nach der gleichen Logik wie bei der Bewertung der Produktqualität, die die Versuchspersonen nach Ansicht der Anzeigen auf Produktebene einschätzen sollten, wurde auch die erwartete Qualität der Produkte auf Unternehmensebene erhoben. Die Items wurden dabei so umformuliert, dass sie generell auf „die Fernseher von NEXTVISION“ abzielen, ohne ein konkretes Produkt in den Vordergrund zu stellen. Auf diese Weise konnte die Vergleichbarkeit der Messungen auf Unternehmens- und Produktebene sichergestellt werden. Entsprechend kamen Reliabilitätsanalysen zu ähnlichen Ergebnissen wie schon auf Produktebene. Erneut ließ sich die interne Konsistenz durch Entfernen des Items „qualitative Mängel“ (hier formuliert als „Die Fernseher von NEXTVISION haben qualitative Mängel“) [umcodiert] deutlich von einem Cronbachs $\alpha = .83$ auf Cronbachs $\alpha = .87$ erhöhen. Das entfernte Item verzeichnete eine lediglich geringe korrigierte Item-

Skala-Korrelation von .21. Auch auf Unternehmensebene konnte es also bedenkenlos aus der Skalenermittlung entfernt werden. Die resultierende Skala auf Unternehmensebene besteht somit aus den fünf verbleibenden Items und weist wie schon die Skala auf Produktebene eine hohe interne Konsistenz bzw. Reliabilität auf.

3.2.2 Halo-Effekte

Halo-Effekte können als sogenannte *Likelihood Halos* oder als *Performance Halos* gemessen werden. Bei einer Messung der *Likelihood Halos* bewerten Versuchspersonen die Wahrscheinlichkeit, dass eine Person oder ein Objekt eine bestimmte Eigenschaft hat, während die Messung von *Performance Halos* das Abschneiden einer Person oder eines Objektes hinsichtlich einer bestimmten Eigenschaft impliziert (Smith et al., 2018, S. 3). Ein *Likelihood-Halo-Effekt* tritt beispielsweise dann auf, wenn eine Versuchsperson von der Energieeffizienz eines Fernsehers, über die Informationen vorliegen, darauf schließt, dass die Verpackung des Fernsehers *wahrscheinlich* auch zum Teil aus recyceltem Material besteht. Entsprechend schließt eine Versuchsperson, die einem *Performance-Halo-Effekt* unterliegt, von der Energieeffizienz eines Fernsehers darauf, dass die Verpackung des Fernsehers hinsichtlich des Recyclinganteils stark *abschneidet*. Methodisch bedeutet dies, dass eine Messung des *Likelihood Halos* durch die Frage „Wie wahrscheinlich ist es, dass [Bewertungsobjekt] die folgende(n) Eigenschaft(en) hat?“ erfolgen muss, während der *Performance-Halo-Effekt* anhand der Frage „Wie gut schneidet [Bewertungsobjekt] bei [Bewertungsdimension] ab?“ (Smith et al., 2018, S. 7) gemessen wird. Unter den Bedingungen geringer Information, die in dieser Forschungsarbeit durch die Stimuli gegeben sind (s. Kap. 3.1), eignet sich damit die Messung von *Likelihood Halos* eher als die Messung von *Performance Halos*. Für die Versuchspersonen wäre eine Messung von *Performance Halos* angesichts der Tatsache, dass diese von dem fiktiven Unternehmen nur einen oberflächlichen ersten Eindruck erhalten haben, sehr irritierend gewesen. Im Sinne des formativen Messmodells (s.o.) ist die Messung der Halo-Effekte als Index zu verstehen (so auch bei Smith et al., 2018), da sich ein solcher additiv über die Bewertung einzelner Umweltkategorien ergibt.

Likelihood Halos (Produktebene). Dieser Argumentation folgend wurde das Auftreten eines Halo-Effektes anhand eines Indexes, der in Anlehnung an den *Likelihood-Halo-Index* von Smith et al. (2018) auf den im Stimulus verwendeten Produkttyp „Fernseher“ und auf die beiden Bezugsebenen zugeschnitten wurde, gemessen. Um die Bewertungsdimensionen festzulegen, erfolgte eine Recherche zur Identifikation relevanter Umweltfaktoren von Elektrogeräten im Allgemeinen und Fernsehern im Speziellen. Darauf aufbauend wurden die Probanden in einzelnen Items um eine Einschätzung der Wahrscheinlichkeit, dass der Omro UHD-Fernseher ein niedriges Transportgewicht, nur minimale Mengen

an Schadstoffen wie Quecksilber und halogenisierte Flammschutzmittel, eine Verpackung, die zum Teil aus Recyclingmaterial besteht und eine lange Lebensdauer aufweist, gebeten. In einem weiteren Item schätzten die Versuchspersonen ein, wie wahrscheinlich es ist, dass möglichst viele Materialien und Bauteile des Omro UHD-Fernsehers wiederverwertet werden können. Die Energieeffizienz eines Fernsehers bzw. die Reduktion seines Stromverbrauchs ist aus Umweltsicht der wichtigste Faktor bei Fernsehern (Umweltbundesamt, 2020). Aus diesem Grund wurde diese Umweltdimension im Stimulus (und damit entsprechend *nicht* in der Messung der Halo-Effekte) verwendet, um die Werbeanzeigen auf Produktebene möglichst realistisch zu gestalten.

Likelihood Halos (Unternehmensebene). Nach vergleichbaren Gesichtspunkten erfolgte die Messung der *Likelihood Halos* für die Unternehmensebene. Statt von einem bestimmten Produkt war auf Unternehmensebene von „den Fernsehern von NEXTVISION“ die Rede, sodass sich die Umweltbewertungsdimensionen, die auch auf Produktebene verwendet wurden, auf das Unternehmen als Ganzes bezogen. Die Items fragten folglich nach der Wahrscheinlichkeit, dass „Nextvision“ verschiedene produktübergreifende Initiativen ergreift.

3.2.3 Hintergrundvariablen

Nach Ansicht der Stimuli und Messung der abhängigen Variablen schloss der Fragebogen mit Messungen des generellen Interesses an den beiden Produkttypen sowie Erhebungen des Umweltanliegens, des Umweltwissens und der Demographie ab. Diese Messungen dienen als Grundlage für den Randomisierungsscheck (s. Kap. 3.3.3).

Interesse an Fernsehern. Das Interesse der Versuchspersonen an Fernsehern wurde in Anlehnung an eine von Jong et al. (2018) entwickelte Skala über fünf siebenstufige Zustimmungs-Items erhoben. Vier der fünf Items wurden in Zuschnitt auf den hier untersuchten Produkttyp von Jong et al. (2018) übernommen (Interesse an verschiedenen Marken, Wissen über Fernseher, Offenheit gegenüber neuen Marken, kein Interesse [umgepolt]). Das verbleibende Item von Jong et al. (2018) wurde leicht abgewandelt, damit es auf Fernseher angewendet werden konnte (keine Absicht, in näherer Zukunft einen Fernseher zu kaufen [umgepolt]). Durch Entfernen des Items „Offenheit gegenüber neuen Marken“ konnte die interne Konsistenz der Skala in einem ersten Schritt von Cronbachs $\alpha = .66$ auf Cronbachs $\alpha = .73$ erhöht werden. Die Item-Skala-Korrelation dieses Items betrug $.12$. Eine weitere deutliche Verbesserung wurde durch Ausschluss des Items zur Kaufabsicht, das in der zweiten Reliabilitätsanalyse eine Item-Skala-Korrelation von $.32$ aufwies, erzielt, sodass die Skala mit den drei verbleibenden Items eine akzeptable interne Konsistenz von Cronbachs $\alpha = .78$ aufweist.

Umweltanliegen. Als nächstes wurde das Umweltanliegen mit Hilfe von drei Items gemessen, die wie in der Untersuchung von Schmuck et al. (2018) aus einer Studie von Schuhwerk und Lefkoff-Hagius (1995) zum Zweck einer generellen Messung des *Involvements* der Versuchspersonen mit der Umwelt übernommen wurden. Erneut wurde auf einer siebenstufigen Skala mit beschrifteten Polen die Zustimmung erhoben, in diesem Fall zu den Items „Ich bin bereit, Opfer für den Schutz der Umwelt zu bringen“, „ich mache mir Sorgen um die Umwelt“ und „der Zustand der Umwelt beeinflusst meine Lebensqualität“. Die Skala weist eine akzeptable Konsistenz von Cronbachs $\alpha = .75$ auf. Zwar hätte sie durch Ausschluss des Items zur Lebensqualität auf ein Cronbachs $\alpha = .79$ erhöht werden können. Das Item wies jedoch eine gute Item-Skala-Korrelation von $.50$ auf. Um der Diversität des Konstruktes „Umweltanliegen“ Rechnung zu tragen, wurde daher von einem Entfernen des Items abgesehen.

Umweltwissen. Von dieser subjektiven Selbsteinschätzung sind objektive Messungen des *Umweltinvolvements* wie Wissenstests zu unterscheiden (Schmuck et al., 2018, S. 131). Sechs Testfragen wurden hierzu aus zwei Untersuchungen übernommen. Schmuck et al. (2018) entwickelten, ausgehend von den Methoden von Mohr, Eroğlu und Ellen (1998), fünf Fragen für eine deutsche Stichprobe. Vier dieser fünf Fragen waren Single-Choice-Fragen. Diese wurden in den Index eingefügt (Wissensfrage zu Recycling-Symbolen (reduziert um eine Antwortmöglichkeit), zu einem fiktiven Baumsymbol (dieses wurde eigenhändig auf Basis eines lizenzfreien Piktogramms erstellt), zu Siegeln für organische Textilien sowie zu erneuerbaren Energien). Auf die Multiple-Choice-Frage wurde verzichtet, um den Fragebogen möglichst intuitiv zu gestalten. Diese vier Fragen wurden um zwei weitere Wissensmessungen ergänzt, die aus den empirisch validierten Empfehlungen von Geiger und Holzhauer (2020, 24-26; 31; 58; 67) stammen (Frage zu CO₂-Fußabdruck von Produkten sowie zur Energieeffizienz von Leuchtmitteln). Durch diese Ergänzung auf insgesamt sechs Fragen konnte die von Schmuck et al. (2018, S. 137) vorgeschlagene Indexbreite von null richtigen Antworten bis zum Maximalscore von sechs richtigen Antworten übernommen werden. Dies ist für die Erzielung von Varianz und zur Gewährleistung der Messvalidität wichtig. Der Test des Umweltwissens der Versuchspersonen ist folglich nicht als Skala, sondern als Index zu verstehen. Es wird angenommen, dass die Items in der Summe das Konstrukt Umweltwissen reliabel abbilden. Die Items erfordern daher keine Überprüfung der internen Konsistenz (formatives Modell; analoge Vorgehensweise wie Schmuck et al., 2018).

Soziodemographie. Im Übrigen wurden einige soziodemographische Merkmale der Versuchspersonen erfasst. Die Erhebung des biologischen Geschlechts unter Einschluss einer halboffenen Antwortmöglichkeit folgt einer pragmatischen Empfehlung von Döring (2013, S. 98–99), wodurch ein trennscharfes, eindeutiges, exhaustives und gender-theoretisch akzeptables Messinstrument vorliegt. Es folgten

Messungen des Alters und des (hoch-)schulischen Bildungsgrades sowie eine Messung des Einkommens in Form der Auswahl einer Einkommensgruppe, die eine genauere Einordnung der Versuchspersonen als Verbraucher ermöglicht. Tabelle 2 bietet eine Übersicht über alle im Erhebungsinstrument verwendeten Messungen.

Tabelle 2

Übersicht über die im Fragebogen verwendeten Messungen

Variable	Quellen	Rolle im Experiment	Pos.
Bzgl. Stimulus Produktebene^a			
Erwartete Produktqualität	In Anlehnung an Pancer et al., 2017 sowie Newman et al., 2014	H2 _P	1/3
Likelihood-Halo	Smith et al., 2018 ^{b,c}	H1.1 _P , H1.2 _P , H2 _P	2/4
Bzgl. Stimulus Unternehmensebene^a			
Erwartete Produktqualität	In Anlehnung an Pancer et al., 2017 sowie Newman et al., 2014 ^d	H2 _U	1/3
Likelihood-Halo	Smith et al., 2018 ^b	H1.1 _U , H1.2 _U , H2 _U	2/4
Fragebogenabschluss			
Interesse an Fernsehern	Jong et al., 2018 ^b	Daten zur Stichprobe	5
Umweltanliegen	Schuhwerk & Lefkoff-Hagius, 1995	s.o.	6
Umweltwissen	Schmuck et al., 2018 sowie Geiger & Holzhauer, 2020	s.o.	7
Soziodemografie			
Geschlecht	Döring, 2013	s.o.	8
Alter		s.o.	9
Bildung		s.o.	10
Einkommen		s.o.	11

^a Stimuli miteinander rotierend, einschließlich der beiden folgenden Fragebatterien. ^b Zuschnitt auf Fernseher als Produkttyp. ^c Zuschnitt auf Produktebene. ^d Zuschnitt auf Unternehmensebene.

3.3 Stichprobe

Der mittels SoSci Survey generierte Link wurde zur Rekrutierung einer anfallenden Stichprobe nach forschungspragmatischen Gesichtspunkten verbreitet. Die Datenerhebung erfolgte zwischen dem 24. August und dem 02. September 2020. Der Verfasser dieser Forschungsarbeit teilte den Link in seinem persönlichen Netzwerk, über E-Mail-Verteiler seiner Hochschule sowie in einer rheinland-pfälzischen Berufsschule und achtete nach Möglichkeit darauf, die zwangsläufig entstehenden Verzerrungen in der Stichprobe gegenüber der Bezugsgesamtheit (Verbraucher in Deutschland) durch Berücksichtigung nicht-studentischer Versuchspersonen im Rahmen zu halten. Es wurden jedoch keine Quotierungen vorgenommen oder andere Mechanismen einer „bewussten Auswahl“ (Koch et al., 2019, S. 117) angewandt, die den Aufwand dieser Arbeit stark hätten ansteigen lassen. Den entstehenden Verzerrungen werden jedoch zumindest durch die Vermessung der Stichprobe (s. Kap. 3.3.2) anhand der Hintergrundvariablen Interesse an Fernsehern, Umwelthanliegen, Umweltwissen, Geschlecht, Alter, Bildung und Einkommen sowie durch einen Randomisierungsscheck (s. Kap. 3.3.3) Rechnung getragen. Die Ergebnisse der Untersuchung werden außerdem vor dem Hintergrund der mangelnden Generalisierbarkeit eingeordnet.

Entlastend kann außerdem darauf hingewiesen werden, dass in dieser Forschungsarbeit sehr spezifische Hypothesen überprüft werden, in denen der Vergleich zwischen Experimental- und Kontrollgruppe und nicht die Generalisierbarkeit der Befunde auf eine Grundgesamtheit im Vordergrund steht (Koch et al., 2019, S. 118–119; Mook, 1983). Dadurch, dass relative Effekte nachgewiesen werden sollen bzw. keine absoluten Effektstärken generalisiert werden sollen (Koch et al., 2019, S. 119) und dass diese Forschungsarbeit auf verschiedene Weisen sensibel mit entstehenden Verzerrungen umgeht, kann die forschungspragmatische Entscheidung für eine anfallende Stichprobe gerechtfertigt werden.

3.3.1 Datenbereinigung

Der Rohdatensatz mit 262 begonnenen Interviews wurde zunächst um alle Versuchspersonen bereinigt, die nicht die Abschlusseite des Fragebogens erreichten. Der Datensatz mit den verbleibenden 173 Versuchspersonen wurde um diejenigen Fälle reduziert, die mindestens einen fehlenden Wert aufwiesen. Eine Ausnahme hierbei stellte die Variable Einkommen dar, zu der insgesamt 12 Probanden die Angabe entweder aktiv verweigerten oder die Abfrage übersprangen. Auf diese Weise können alle Berechnungen mit Ausnahme derer, die die Variable „Einkommen“ beinhalten, auf Basis des gleichen Datensatzes erfolgen, was eine Vergleichbarkeit der Berechnungen garantiert. Eine Person, die anstelle ihrer schulischen Bildung in der halboffenen Antwortmöglichkeit ihre Berufsbildung angab,

musste aus Gründen der Trennschärfe gesondert vom Datensatz ausgeschlossen werden (eine Ausnahme von dieser Trennschärfe bildet die Hochschulbildung, die bereits in den Antwortmöglichkeiten enthalten war). Diese Bereinigungsmaßnahmen führten zu einem Datensatz mit insgesamt 163 Versuchspersonen ($n = 151$ für Berechnungen mit der Variable „Einkommen“).

In der Folge wurde kontrolliert, ob nur Werte vergeben wurden, die vorkommen dürfen. Dies war der Fall. Anschließend wurde überprüft, ob der Datensatz Versuchspersonen enthält, die den Fragebogen augenscheinlich zu schnell beantwortet haben, um eine valide Erhebung zu gewährleisten. Kein solcher Fall wurde identifiziert. Zwar ist die Spannweite der Zeitdauer hoch (*Min* 127 Sekunden, *Max* 642 Sekunden); dies ist jedoch bei Fragebögen, die einen Wissenstest beinhalten, nicht unüblich (s. Leiner, 2019, S. 242). Nichtsdestotrotz wurden die Versuchspersonen am unteren Ende dieser Spannweite noch einmal gesondert überprüft; es konnten hierbei keine Antwortmuster ausgemacht werden, die auf ein Durchklicken von Fragebatterien hindeuten würden. Zudem war die Betrachtungsdauer der Stimuli bei diesen Versuchspersonen ausreichend und wich nicht von der Betrachtungszeit der anderen Versuchspersonen ab.

3.3.2 Vermessung der Stichprobe

Die auf diese Weise reduzierte anfallende Stichprobe ($N = 163$) bestand aus 100 männlichen (61 %) und 63 weiblichen (39 %) Versuchspersonen. Das Durchschnittsalter der Versuchspersonen beträgt 26.12 Jahre (*Min* = 16, *Max* = 81, *SD* = 9.32) und ist damit als jung einzustufen. 41 Versuchspersonen besitzen einen (Fach-)Hochschulabschluss als höchsten Bildungsabschluss, 95 das Abitur/Hochschulreife und weitere 15 das Fachabitur/Fachhochschulreife. Lediglich 11 Probanden gaben „Mittlere Reife, Real- schul- oder gleichwertiger Abschluss“ als höchsten Bildungsabschluss an und nur eine Person im Sample weist einen Volks- oder Hauptschulabschluss bzw. Quali auf. Entsprechend wurden diese beiden Stufen mit mittleren und niedrigen Bildungsabschlüssen für die weiteren Berechnungen zusammengefasst. Folglich besitzen 83 % der Probanden mindestens das Abitur, woraus geschlossen werden muss, dass gebildete Versuchspersonen in der Stichprobe deutlich überrepräsentiert sind. Von den Versuchspersonen, die ihr Einkommen angegeben haben ($n = 151$), verfügen 29 (19 %) über kein eigenes Einkommen, 28 (19 %) über weniger als 500 Euro, 39 über 500 bis unter 1000 Euro (26 %), 21 (14 %) über 1000 bis unter 1500 Euro und 12 (8 %) über 1500 bis unter 2000 Euro. 22 (15 %) Probanden ordneten sich einer der Kategorien oberhalb der 2000 Euro zu, davon 5 (3 %) in der höchsten Stufe (4000 Euro oder mehr). Schon an dieser Stelle wurden die Kategorien oberhalb der 2000 Euro zusammengefasst, da sich in ihnen ansonsten nur sehr wenige Fälle befunden hätten. Folglich sind die

Einkommen der Stichprobe in den meisten Fällen unterdurchschnittlich, was vor dem Hintergrund des geringen Durchschnittsalters wenig verwundert.

Der Durchschnittswert für das Interesse an Fernsehern beträgt 3.67 ($SD = 1.69$) und befindet sich somit nah an der Skalenmitte (4), was als eine im Mittel neutrale Einstellung gegenüber diesem Produkttyp gewertet werden kann. Entsprechend enthält die Stichprobe in ähnlichen Teilen sowohl Versuchspersonen mit einem geringen als auch Versuchspersonen mit einem hohen Interesse an Fernsehern. Das durchschnittliche Umweltsanliegen der Versuchspersonen, das auf einer siebenstufigen Skala erhoben wurde, beträgt 5.75 ($SD = 1.09$). Insgesamt gaben viele der Probanden folglich an, dass ihnen die Umwelt wichtig ist. Hierbei ist allerdings anzumerken, dass auch eine quotenbasierte Stichprobe, die deutsche Verbraucher abbildete, in einer aktuellen Studie von Schmuck et al. (2018, S. 143) hohe Werte auf der gleichen Skala erzielte ($M = 4.85$, $SD = 1.33$). Im Mittel schlossen die Versuchspersonen den Umweltwissenstest mit 4.28 von 6 möglichen Punkten ab ($SD = 1.26$). Auch wenn nur 13 (8 %) Versuchspersonen mit weniger als 3 Punkten abschlossen, kann daher festgehalten werden, dass sich der Umweltwissenstest eignet, um im Rahmen dieser Variable zwischen den Versuchspersonen zu differenzieren.

Insgesamt sind somit Verzerrungen entstanden, die aufgrund des Rekrutierungsverfahrens (s.o.) erwartbar waren. Für die Interpretation von Ergebnissen sollte daher im Hinterkopf behalten werden, dass die Stichprobe übermäßig jung und gebildet ist, sich durch ein hohes Umweltsanliegen auszeichnet und in den meisten Fällen über ein geringes oder kein eigenes Einkommen verfügt. Dies ändert jedoch nichts daran, dass die Stichprobe ausnahmslos zur Zielgruppe von TV-Herstellern besteht und zusätzlich als Zielgruppe für Umweltkommunikation verstanden werden kann.

3.3.3 Randomisierungskontrolle

Zur Überprüfung, ob die Versuchspersonen hinsichtlich der erhobenen Hintergrundvariablen gleichmäßig auf die Versuchsgruppen (vage bzw. konkrete Botschaft) verteilt wurden, sind Randomisierungskontrollen notwendig. Die Gruppen sollten ein ähnliches Interesse an Fernsehern haben und hinsichtlich Umweltsanliegen, Umweltwissen, Geschlecht, Alter, Bildung und Einkommen vergleichbar sein, um eine erfolgreiche Randomisierung durch die Fragebogensoftware annehmen zu können. Durch die nachträgliche Datenbereinigung (s. Kap. 3.3.1) sind die Gruppengrößen leicht unterschiedlich: 86 Versuchspersonen sahen eine Produktwerbung mit vager Botschaft, 77 eine konkrete

Botschaft. Bei der Anzeige auf Unternehmensebene erhielten 77 Versuchspersonen eine vage Botschaft und 86 ein konkretes Umweltargument⁴.

Randomisierungskontrolle für das Experiment auf Produktebene. χ^2 -Tests zeigten, dass es keine signifikanten Unterschiede zwischen den beiden Gruppen auf Produktebene hinsichtlich Geschlecht ($\chi^2 = 0.32$, $p = .571$), recodierter (s. Kap. 3.3.2) Bildung ($\chi^2 = 0.45$, $p = .931$) und recodiertem (s. Kap. 3.3.2) Einkommen ($n = 151$; $\chi^2 = 5.18$, $p = .394$) gibt. Für die metrischen Hintergrundvariablen (Alter, Interesse an Fernsehern, Umweltanliegen, Umweltwissen) wurde der Randomisierungskontrolle mit Hilfe von t -Tests bei unabhängigen Stichproben durchgeführt. Ein t -Test offenbarte keine signifikanten Unterschiede zwischen den beiden Gruppen auf Produktebene hinsichtlich des Alters (konkret: $M = 25.88$, $SD = 8.01$, vage: $M = 26.34$, $SD = 10.4$; Differenz von 0.46 Jahren nicht signifikant: $t(161) = -0.31$, $p = .757$). Ähnliches gilt hinsichtlich der Skala, mit der das Interesse an Fernsehern gemessen wurde (konkret: $M = 3.66$, $SD = 1.69$, vage: $M = 3.68$, $SD = 1.69$; Differenz von 0.02 Skaleneinheiten nicht signifikant: $t(161) = -0.09$, $p = .928$). Auch bezüglich der Variable Umweltanliegen war die Randomisierung erfolgreich (konkret: $M = 5.84$, $SD = 1.01$, vage: $M = 5.67$, $SD = 1.16$; Differenz von 0.16 Skaleneinheiten nicht signifikant: $t(161) = 0.94$, $p = .349$). Ein weiterer t -Test verglich das Umweltwissen der Versuchsgruppe, die die konkrete Produktanzeige erhielt ($M = 4.3$, $SD = 1.24$) mit der Gruppe, die die vage Produktanzeige betrachtete ($M = 4.26$, $SD = 1.29$). Die resultierende Indextendifferenz von 0.04 Punkten ist nicht signifikant: $t(161) = 0.22$, $p = 0.829$.

Randomisierungskontrolle für das Experiment auf Unternehmensebene. Auch auf Unternehmensebene zeigten die χ^2 -Tests, dass es keine signifikanten Unterschiede zwischen den beiden Gruppen hinsichtlich Geschlecht ($\chi^2 = 0.32$, $p = .571$), recodierter Bildung ($\chi^2 = 2.83$, $p = .419$) und recodiertem Einkommen ($n = 151$; $\chi^2 = 2.51$, $p = .776$) gibt. Wie schon auf Produktebene wurden für die metrischen Hintergrundvariablen (Alter, Interesse an Fernsehern, Umweltanliegen, Umweltwissen) Randomisierungskontrollen mit Hilfe von t -Tests durchgeführt. Diese Vorgehensweise offenbarte keine signifikanten Unterschiede zwischen den beiden Gruppen auf Unternehmensebene hinsichtlich der metrischen Variablen Alter (konkret: $M = 26.92$, $SD = 11.05$; vage: $M = 25.23$, $SD = 6.86$, Differenz von 1.69 Jahren nicht signifikant: $t(161) = 1.15$, $p = .250$) und dem Interesse an Fernsehern (konkret: $M = 3.58$, $SD = 1.73$, vage: $M = 3.77$, $SD = 1.64$; Differenz von 0.19 Skaleneinheiten nicht signifikant: $t(161) = -0.71$, $p = .477$). Die Randomisierung auf Unternehmensebene war auch bezüglich der Variable Umweltanliegen erfolgreich (konkret: $M = 5.82$, $SD = 1.1$, vage: $M = 5.68$, $SD = 1.1$; Differenz von 0.14 Skaleneinheiten nicht

⁴ Diese „umgekehrte“ Verteilung der Gruppengrößen ergab sich rein zufällig. Wie beschrieben (s. Kap. 3.1) waren alle vier möglichen Kombinationen der Konkretisierungsgrade zwischen Produkt- und Unternehmensebene im Datensatz vertreten.

signifikant: $t(161) = 0.83, p = .408$). Aus einem weiteren t -Test ergaben sich keine signifikanten Unterschiede zwischen dem Umweltwissen der Versuchsgruppe, die die konkrete Unternehmensanzeige erhielt ($M = 4.19, SD = 1.33$) und dem Wissen der Gruppe, die die vage Unternehmensanzeige betrachtete ($M = 4.38, SD = 1.17$; Indexdifferenz von 0.19 Punkten ist signifikant: $t(161) = -0.97, p = 0.336$).

Somit kann resümiert werden, dass die Randomisierung erfolgreich war. Sowohl auf Produkt- als auch auf Unternehmensebene sind jeweils die beiden Gruppen (konkret und vage) hinsichtlich der Variablen Geschlecht, Bildung, Einkommen, Alter, Interesse an Fernsehern, Umwelthanliegen und Umweltwissen vergleichbar.

4 Auswertung und Interpretation der Ergebnisse

Nachdem Kapitel 3 die Methode dieser Forschungsarbeit darlegte und somit die Validität und Reliabilität des Messinstrumentes sowie die Vergleichbarkeit der Experimentalgruppen überprüfte bzw. intersubjektiv nachvollziehbar machte, widmen sich dieses und das folgende Kapitel der Datenauswertung zwecks Überprüfung der in Kapitel 2 aufgestellten Hypothesen.

4.1 Auftreten von Halos (H1.1_p und H1.1_u)

Hypothese H1.1_p äußerte die Vermutung, dass Verbraucher sowohl nach Wahrnehmung eines konkreten Umweltarguments auf Produktebene als auch nach Wahrnehmung eines vagen Umweltarguments auf Produktebene schlussfolgern, dass das Produkt noch weitere Umwelteigenschaften aufweist (Unternehmensebene analog). Für die Hypothese H1.1_p und für H1.1_u liegt ein CSR-Halo-Effekt aus statistischer Sicht dann vor, wenn sich die durchschnittliche Einschätzung der Versuchspersonen auf einer Skala von eins bis sieben signifikant positiv vom Skalenmittelpunkt vier unterscheidet. Der Produkt-Halo-Index wurde so konstruiert, dass die Versuchspersonen unterhalb des Skalenmittelpunkts ihre Ablehnung einer Umwelteigenschaft und oberhalb des Skalenmittelpunktes ihre Zustimmung abstufen können (s. Kap. 3.2.2). Somit stellt der Indexmittelpunkt vollkommene Neutralität dar und spiegelt wider, dass eine Versuchsperson eine Umwelteigenschaft als weder wahrscheinlich noch als unwahrscheinlich einstuft. Als Verfahren eignet sich für die Überprüfung der Hypothese H1.1_p und für H1.1_u ein t -Test bei einer Stichprobe. Mit diesem Test können die Indexmittelpunkte statistisch mit dem Indexmittelpunkt (4) verglichen werden, indem etwaige Abweichungen auf Signifikanz überprüft werden.

Die Versuchspersonen, die die vage Produktanzeige betrachteten ($n = 86$), stufen die im Produkt-Halo-Index enthaltenen weiteren möglichen Umwelteigenschaften des Omro UHD-Fernsehers im

Durchschnitt als leicht unwahrscheinlich ein ($M = 3.66$, $SD = 1.13$). Ein t -Test bei einer Stichprobe zeigte, dass diese negative Abweichung von der Indexmitte (4) um -0.34 Indexpunkten signifikant ist, $t(85) = -2.79$, $p = .007$. Ähnliches gilt für die Versuchspersonen, die die konkrete Produktanzeige als Stimulus erhielten ($n = 77$). Sie stufen die Wahrscheinlichkeit, dass der Fernseher noch weitere Umwelteigenschaften hat, ebenfalls im Durchschnitt als leicht unwahrscheinlich ein ($M = 3.59$, $SD = 1.22$). Diese Experimentalgruppe wich somit durchschnittlich um -0.41 Skalenpunkte vom Indexmittelpunkt (4) ab, wodurch in einem weiteren t -Test bei einer Stichprobe eine Signifikanz aufgedeckt werden konnte, $t(76) = -2.91$, $p = .005$.

Somit kann Hypothese H1.1_P, die postulierte, dass sich sowohl bei vagen als auch bei konkreten Botschaften positive Halo-Effekte ergeben würden, abgelehnt werden. Stattdessen sind in beiden Experimentalgruppen negative Halo-Effekte aufgetreten; die Versuchspersonen stufen die im Messinstrument abgefragten weiteren Umwelteigenschaften insgesamt systematisch als eher unwahrscheinlich ein. Zur deskriptiven Unterstützung dieses induktiven Ergebnisses wurde der Produkt-Halo-Index mit Hilfe von Quantilen so recodiert, dass oberhalb des Indexmittelpunktes drei ca. gleich große Gruppen entstanden; oberhalb 4.00 bis 4.40 Indexpunkte („schwacher Halo-Effekt“), mehr als 4.40 bis 5.00 („moderater Halo-Effekt“) sowie oberhalb 5.00 bis 7.00 („starker Halo-Effekt“). Knapp zwei Drittel der Fälle (66 %; $n = 107$) unterlagen keinem positiven Halo-Effekt. Etwas mehr als jeder Dritte (34 %; $n = 56$) hielt also die weiteren Umwelteigenschaften für eher wahrscheinlich als unwahrscheinlich, obwohl darüber in der Anzeige keine Informationen enthalten waren. In der vagen Experimentalgruppe traf dies auf 35 % ($n = 30$) der Versuchspersonen zu. Die konkrete Experimentalgruppe verzeichnete bei 34 % ($n = 26$) einen solchen positiven Halo-Effekt. Von den Versuchspersonen mit positivem Produkt-Halo-Effekt befanden sich gruppenübergreifend 23 (14 %) in der schwachen Gruppe, 17 (10 %) in der moderaten und 16 (10 %) in der starken Gruppe.

Das Umweltargument des geringen Jahresstromverbrauchs ließ also die Mehrheit der Versuchspersonen nicht auf die Existenz weiterer Umwelteigenschaften schließen. Mehrere, unterschiedliche Erklärungen sind für diesen kontratheoretischen Befund denkbar. Einerseits ist es möglich, dass die Versuchspersonen, die überwiegend jung, gebildet und umweltbesorgt waren, skeptisch gegenüber der Anzeige eingestellt waren. Eine weitere Möglichkeit ist, dass der Jahresstromverbrauch aus Sicht der befragten Verbraucher ein wichtiger Leistungs- und Qualitätsaspekt eines Fernsehers ist, der daher keine Wahrnehmung des Produktes als umweltfreundlich induziert, sondern beispielsweise auf eine günstigere Stromrechnung oder modernere Technik schließen lässt. Im Rückblick auf die Theorie könnte auf diese Weise somit ein *General Impression Halo* ausgeblieben sein. Damit

zusammenhängend sind Verbraucher vermutlich daran gewöhnt, dass TV-Hersteller mit der Energieeffizienz ihrer Geräte werben, sodass eine solche Umwelteigenschaft andere, spezifischere Umweltattribute nicht wahrscheinlicher macht, da es aus Sicht der Verbraucher an Ähnlichkeit mit der beworbenen Eigenschaft mangelt. Dies spricht im vorliegenden Fall für ein Ausbleiben eines *Interdimensional Similarity Halo*. Es ist allerdings nicht zu vernachlässigen, dass sich ungefähr jede dritte Versuchsperson zu einem positiven Halo-Effekt verleiten ließ. Aus Unternehmenssicht könnte man die Anzeige daher als Erfolg werten, da zahlreiche Verbraucher positive Eigenschaften vermuteten, die gar nicht kommuniziert werden mussten.

Hypothese H1.1_U postulierte (analog zur Produktebene), dass die befragten Verbraucher sowohl nach Wahrnehmung eines konkreten Umweltarguments auf Unternehmensebene als auch nach Wahrnehmung eines vagen Umweltarguments auf Unternehmensebene schlussfolgern, dass das Unternehmen noch weitere Umweltinitiativen ergreift. Wie schon auf Produktebene wurde die Hypothese anhand von *t*-Tests bei einer Stichprobe überprüft, indem untersucht wurde, ob die Versuchspersonen im Durchschnitt signifikant von der Indexmitte abwichen. Die Versuchspersonen, die die vage Unternehmensanzeige betrachteten ($n = 77$), stuften die im Produkt-Halo-Index enthaltenen weiteren möglichen Umweltinitiativen der fiktiven Firma „Nextvision“ im Durchschnitt als recht wahrscheinlich ein ($M = 4.83$, $SD = 1.31$). Ein *t*-Test bei einer Stichprobe offenbarte, dass diese positive Abweichung von der Indexmitte (4) um 0.83 Indexpunkte signifikant ist, $t(76) = 5.57$, $p < .001$. Ähnliches gilt für die Versuchspersonen, die die konkrete Unternehmensanzeige als Stimulus erhielten ($n = 86$). Sie stuften die Wahrscheinlichkeit, dass „Nextvision“ noch weitere Umwelteigenschaften ergreift, ebenfalls im Durchschnitt als recht wahrscheinlich ein ($M = 4.65$, $SD = 1.23$). Diese Experimentalgruppe wich somit durchschnittlich um 0.65 Skalenpunkte vom Indexmittelpunkt (4) ab, wodurch in einem weiteren *t*-Test bei einer Stichprobe eine Signifikanz aufgedeckt werden konnte, $t(85) = 4.89$, $p < .001$.

Sowohl bei der Bewertung der vagen als auch bei den konkreten Umweltanzeigen erwarteten die Versuchspersonen in systematischer Weise, dass das fiktive Unternehmen neben der beworbenen noch weitere Umweltinitiativen ergreift. Insofern kann die Hypothese H1.1_U akzeptiert werden. Nur 27 % der Probanden ($n = 44$) unterlagen keinem Unternehmens-Halo-Effekt (Indexmittelwert ≤ 4). Weitere 28 % ($n = 45$) verzeichneten einen relativ schwachen positiven Halo-Effekt (Indexmittelwert größer als 4.00 bis 4.80), 21 % ($n = 35$) kann ein vergleichsweise moderater Halo-Effekt zugeschrieben werden (oberhalb 4.80 bis 5.60); 39 (24 %) Versuchspersonen unterlagen einem starken Halo-Effekt (mehr als 5.60 bis 7.00). Dies ergab die Recodierung des Unternehmens-Halo-Index anhand von Quantilen, durch die oberhalb des Indexmittelpunktes drei ca. gleich große Gruppen entstanden. Die konkrete

Anzeige bewirkte bei 72 % der Befragten in dieser Gruppe einen positiven Halo-Effekt ($n = 62$). In der vagen Experimentalgruppe verzeichneten 74 % ($n = 57$) ein solches Bewertungsverhalten. Ausgehend von den Anzeigen auf Unternehmensebene traten somit deutlich häufiger (und stärker, s. Terzile oberhalb des Indexmittelpunktes) positive Halo-Effekte auf als ausgehend von den Anzeigen auf Produktebene. Das Unternehmen „Nextvision“ konnte mit Hilfe der Anzeige im ersten Eindruck also dimensionsübergreifend Umweltfreundlichkeit suggerieren, was „Zensoric“ mit der Produktanzeige verwehrt blieb.

4.2 Halos bei vagen und konkreten Botschaften im direkten Vergleich (H1.2_p und H1.2_u)

In Hypothese H1.2_p wurde die Vermutung geäußert, dass konkrete Umweltbotschaften auf Produktebene stärker als vage Umweltbotschaften auf Produktebene Halo-Effekte erzeugen. Es soll also die Auswirkung der Zuordnung der Versuchspersonen zu einer von zwei Ausprägungen einer kategorialen Variable (Gruppeneinteilung vage bzw. konkret) auf eine metrische abhängige Variable (Produkt-Halo-Index) ermittelt werden. Hierzu eignet sich ein t -Test bei unabhängigen Stichproben, mit dem die Mittelwerte zweier Gruppen verglichen werden. Auf diese Weise kann der Zusammenhang zwischen der Gruppeneinteilung und einer anderen Variable ermittelt werden (Field, 2018, 445-446; 448-451). Ein solcher zweiseitiger t -Test empfiehlt sich trotz der gerichteten Hypothese, da ein umgekehrter Zusammenhang, also ein „besseres“ Abschneiden der vagen Umweltbotschaft, aufgrund von Theorie und Forschungsstand (s. Kap. 2.2) nicht ausgeschlossen werden kann (s. hierzu Field, 2018, S. 79-82).

Im Durchschnitt unterschieden sich die Versuchspersonen, die die konkrete Produktanzeige gesehen hatten, hinsichtlich der Einschätzung der Wahrscheinlichkeit weiterer Umwelteigenschaften des Produktes kaum ($M = 3.59$, $SD = 1.22$) von den Befragten, die die Produktanzeige mit vager Umweltbotschaft betrachteten ($M = 3.66$, $SD = 1.13$). Die nur marginale Differenz von 0.07 Indexpunkten ist nicht signifikant, $t(161) = -0.36$, $p = .722$. Weitere t -Tests bei unabhängigen Stichproben, die mit den einzelnen Items des Produkt-Halo-Index als Outcome-Variablen durchgeführt wurden, brachten ebenfalls keine signifikanten Unterschiede zwischen den Experimentalgruppen hervor (s. Tab. 3). Unter Heranziehung des Mittelwertes schnitt die vage Botschaft in vier der fünf Umweltbewertungsdimensionen „besser“ ab als der konkrete Claim.

Tabelle 3*Unterschiede im Produkt-Halo-Index und in den Items nach Konkretisierungsgrad*

	Konkret (<i>n</i> = 77)		Vage (<i>n</i> = 86)		<i>t</i> (<i>df</i>)	<i>p</i> ^a	Cohens <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Produkt-Halo- Index	3.59	1.22	3.66	1.13	-0.36(161)	.722	-0.06 ^b
Schadstoffmi- nimierung	3.55	1.39	3.69	1.49	-0.62(161)	.536	-0.09 ^b
Lebensdauer	3.97	1.52	3.79	1.45	0.79(161)	.432	0.12 ^b
Transportge- wicht	3.48	1.60	3.56	1.44	-0.33(161)	.745	-0.06 ^b
Verpackungen	3.71	1.75	3.83	1.52	-0.43(151,683)	.667	-0.07 ^c
Wiederver- wendbarkeit	3.26	1.5	3.44	1.43	-0.79(161)	.428	-0.13 ^b

^aZweiseitiger Signifikanztest. ^bMittelwertdifferenz dividiert durch die Standardabweichung der vagen Gruppe. ^c Mittelwertdifferenz dividiert durch gepoolte Standardabweichung angesichts eines signifikanten Levene-Tests, $F(1, 161) = 4.23, p = .041$.

Da die Nullhypothese, dass kein Unterschied zwischen den beiden Experimentalgruppen hinsichtlich des Produkt-Halo-Indexes besteht, anhand des *t*-Tests bei unabhängigen Stichproben nicht abgelehnt werden konnte, muss die H1.2_P verworfen werden. Die konkrete Umweltbotschaft war entgegen der Annahme, dass konkrete Informationen das Bewertungsrisiko aus Sicht der Verbraucher reduzieren, unter den im Experiment gegebenen Bedingungen geringer Information nicht in der Lage, stärkere positive Halo-Effekte als die vage Umweltbotschaft zu erzielen. Verbraucher scheinen bei Betrachtung einzelner Werbeanzeigen hinsichtlich der tatsächlichen weiteren Umwelteigenschaften eines Produktes nicht zwischen konkreten und vagen Botschaften zu unterscheiden. Sie trauen dem Omro UHD-Fernseher, der mit einer nicht belastbaren Aussage beworben wurde, genauso weitere Umwelteigenschaften zu wie dem Fernseher, der mit einer konkreten, in Zahlen gefassten Aussage vorgestellt wurde.

Aus theoretischer Sicht ist es im Anschluss an diese Ergebnisse weiterhin möglich, dass die vage und die konkrete Botschaft auf unterschiedliche Arten Halo-Effekte begünstigen bzw. abschwächen,

sodass beide letztlich ähnliche Werte auf dem Produkt-Halo-Index erzielten. Die Ergebnisse sprechen für die Vermutung, dass vage Botschaften Verbraucher zu einem Gesamteindruck zwingen (Cooper, 1981, S. 220), allerdings gleichzeitig unter Bedingungen der ohnehin schon spärlichen Information ein höheres Bewertungsrisiko (Boatwright et al., 2008) als eine konkrete Botschaft schaffen. Für eine konkrete Botschaft kann ein umgekehrter Mechanismus vermutet werden. Insofern ist das Ergebnis der Überprüfung von H1.2_p im Rückblick auf die Theorie (s. Kap. 2.2) erklärbar. Resümierend bedeutet dies, dass, auch wenn die Hypothesenkonstruktion letztlich der Argumentation zum Bewertungsrisiko folgte und die Hypothese abgelehnt werden musste, die zugehörige Annahme nicht falsifiziert werden konnte. Allerdings ist die Bedeutung des Bewertungsrisikos im direkten Vergleich von vagen und konkreten Botschaften unter Bedingungen geringer Information scheinbar eingeschränkt.

Die Hypothese H1.2_u postulierte, ähnlich wie Hypothese H1.2_p für die Anzeigen auf Produktebene, dass konkrete Umweltanzeigen stärkere Halo-Effekte induzieren als vage Umweltanzeigen. Der Symmetrie der beiden Unterexperimente weiter folgend wurde auch für die Überprüfung der H1.2 auf Unternehmensebene ein *t*-Test bei unabhängigen Stichproben durchgeführt. Entgegen der formulierten Hypothese schätzten die Versuchspersonen, die der vagen Umweltbotschaft zugeordnet wurden, die Wahrscheinlichkeit, dass „Nextvision“ noch weitere Umweltinitiativen ergreift als höher ein ($M = 4.83$, $SD = 1.31$) als die Versuchspersonen, die die quantifizierte Umweltbotschaft erhielten ($M = 4.65$, $SD = 1.23$). Dieser Unterschied von 0.18 Indexpunkten erwies sich als nicht signifikant, $t(161) = -0.92$, $p = .360$. Auch bezüglich der Einzelitems, die gemeinsam den Unternehmens-Halo-Index bildeten, gab es keine signifikanten Unterschiede zwischen den beiden Experimentalgruppen (s. Tab. 4). Allerdings ist bemerkenswert, dass die vage Anzeige bei jedem Item hinsichtlich des Mittelwertes leicht stärker abschnitt als die konkrete Anzeige. Somit wurde auf Unternehmensebene die Tendenz zu verstärkten positiven Halo-Effekten bei vagen Claims, die schon auf Produktebene bei vier der fünf Items zutage trat, repliziert. Wie schon auf Produktebene konnte sich die konkrete Botschaft hinsichtlich der wahrgenommenen dimensionsübergreifenden Umweltfreundlichkeit nicht gegen die vagen Botschaften durchsetzen. Daher kann die Hypothese H1.2_u abgelehnt werden. Auf Unternehmensebene schnitt die vage Anzeige hinsichtlich des Halo-Index sogar „besser“ ab als die konkrete Anzeige, auch wenn dieser Unterschied nur gering (0.18 Indexpunkte) und nicht signifikant war.

Tabelle 4*Unterschiede im Unternehmens-Halo-Index und in den Items nach Konkretisierungsgrad*

	Konkret (n = 86)		Vage (n = 77)		t(161)	p ^a	Cohens d ^b
	M	SD	M	SD			
Unternehmens- Halo-Index	4.65	1.23	4.83	1.31	-0.92	.360	-0.14
Schadstoffmini- mierung	4.66	1.66	4.99	1.66	-1.25	.214	0.20
Lebensdauer	4.19	1.61	4.40	1.64	-0.85	.397	-0.13
Transportge- wicht	4.49	1.49	4.55	1.64	-0.23	.816	-0.04
Verpackungen	5.22	1.51	5.30	1.56	-0.32	.747	-0.05
Wiederverwend- barkeit	4.67	1.68	4.91	1.41	-0.96	.339	-0.17

^aZweiseitiger Signifikanztest. ^bMittelwertdifferenz dividiert durch die Standardabweichung der vagen Gruppe.

Damit verdichten sich die Schlussfolgerungen, die bereits bei der Überprüfung der H1.2_p bezüglich des Verhältnisses der vagen und konkreten Umweltbotschaften getroffen wurden. Vage Umweltbotschaften scheinen unter Bedingungen geringer Information eine ähnliche Tendenz zu positiven Halo-Effekten zu induzieren wie konkrete Umweltbotschaften. Ein Rückgriff auf die Theorie (s. Kap. 2.2) erlaubt die Vermutung, dass die generellere, weniger spezifische Formulierung von vagen Botschaften zum Gesamteindruck einlädt, das Unternehmen ergreife noch weitere Umweltinitiativen, während die konkreten Botschaften ein geringeres Bewertungsrisiko hinsichtlich der weiteren Umweltdimensionen vermitteln. Auch auf Unternehmensebene würdigten die Verbraucher die mit Zahlen versehenen und damit belastbareren Botschaften nicht hinsichtlich der Einschätzungen zu weiteren Umweltdimensionen. Im Sinne des Problemaufrisses dieser Forschungsarbeit (s. Kap. 1) kann dies als besorgniserregend aufgefasst werden.

4.3 Korrelation zwischen Halo-Index und erwarteter Qualität (H2_p und H2_u)

Hypothese H2_p äußerte die Vermutung, dass zwischen dem Halo-Index auf Produktebene und der Einschätzung der Produktqualität ein negativer Zusammenhang besteht. Somit bezieht sich die

Hypothese auf den Zusammenhang zweier metrischer Variablen. Grundsätzlich kommen für die Hypothesenüberprüfung angesichts dieses Datenniveaus somit einerseits die Korrelationsanalyse nach Pearson und andererseits eine Regression in Betracht. Eine Regression setzt voraus, dass (aufgrund theoretischer Erwägungen) bekannt ist, welche Variable Prädiktor ist, die das Kriterium kausal beeinflusst (Field, 2018, S. 371–376). In Kapitel 2.3 wurde erläutert, warum dies für den Zusammenhang zwischen CSR-Halo-Effekten und der Einschätzung der Produktqualität unter den gegebenen Voraussetzungen nicht der Fall sein kann. Eine Produkt-Moment-Korrelationsanalyse lässt dagegen offen, welche Variable unabhängig und welche Variable abhängig ist und ist somit aus theoretischer Sicht für die Überprüfung der vorliegenden Hypothese vorzugswürdig.

Diese deckte einen hochsignifikanten, moderat positiven Zusammenhang zwischen dem Halo-Index auf Produktebene und der Einschätzung der Produktqualität auf, $r = .376, p < .001$. Mit steigender Einschätzung, dass der Omro UHD-Fernseher wahrscheinlich noch weitere Umwelteigenschaften aufweist, stieg also auch die eingeschätzte Produktqualität. Umgekehrt ausgedrückt – und dies ist angesichts der hier vorliegenden *negativen* Halo-Effekte (s. Kap. 4.1) besonders erwähnenswert – fiel bei den Versuchspersonen bei geringer Erwartung, dass der Fernseher weitere Umwelteigenschaften hat, tendenziell auch die Einschätzung der Produktqualität ab. Allerdings zeigen weitere Produkt-Moment-Korrelationsanalysen, dass der Halo-Index vor allem in der Gruppe, die die vage Umweltbotschaft erhielt ($n = 86$), mit der erwarteten Produktqualität korrelierte, $r = .479, p < .001$, während die konkrete Experimentalgruppe ($n = 77$) eine schwächere Korrelation verzeichnete, $r = .272, p < .017$.

Hypothese H2_p kann somit abgelehnt werden; statt des erwarteten negativen Zusammenhangs zwischen CSR-Halo-Effekten und der erwarteten Produktqualität liegt ein positiver Zusammenhang vor. Daraus kann geschlussfolgert werden, dass die Befürchtung, Verbraucher nutzen zur Einschätzung der Produktqualität kompensatorische Inferenzstrategien, indem sie von einer hohen Umweltfreundlichkeit auf eine niedrigere Qualität schließen, im konkreten Beispiel und Datensatz unbegründet ist. Zieht man zusätzlich zur durchgeführten Korrelationsanalyse die Mittelwerte und Standardabweichungen hinsichtlich Produkt-Halo-Index ($M = 3.63, SD = 1.17$) und erwarteter Produktqualität ($M = 3.93, SD = 1.10$) heran, die die Gesamttendenz der Versuchspersonen näher beleuchten, kommt man zu dem Ergebnis, dass die Anzeigen hinsichtlich der eingeschätzten Umweltfreundlichkeit über die Dimensionen hinweg vergleichsweise schlecht abschnitten, was sich in einer nur durchschnittlichen Bewertung der Produktqualität im Sinne eines Gesamteindrucks fortsetzte. Insofern spricht die Korrelation zwischen den beiden Variablen für einen *General Impression Halo*, indem sowohl die Einschätzung der Produktqualität als auch die Einschätzung der Wahrscheinlichkeit weiterer

Umwelteigenschaften auf einen gemeinsamen ersten Eindruck zurückzuführen ist. Dies erscheint plausibler als die Vermutung eines *Interdimensional Similarity Halo*, bei dem Korrelationen in der Bewertung verschiedener Dimensionen auf deren Ähnlichkeit zurückgeführt werden können. Dass die vage Botschaft eine stärkere Korrelation induzierte, spricht folglich für die aufgestellte Annahme (s. Kap. 2.2), dass allgemeinere Formulierungen eher als konkrete einen *General Impression Halo* auslösen. Grundsätzlich kann nämlich angeführt werden, dass sich die funktionelle Qualität und die dimensionsübergreifende Umweltfreundlichkeit bei technischen Produkten nicht gegenseitig bedingen, sodass die Versuchspersonen die beiden Konstrukte kaum aus diesem Grund tendenziell ähnlich bewerteten. Somit weist die Überprüfung der H2_P darauf hin, dass die beiden Bewertungsdimensionen bei technischen Produkten nicht etwa konträr zueinander stehen, sondern durchaus Synergieeffekte erzielt werden können, indem bessere Bewertungen von Umwelteigenschaften mit einer höheren Einschätzung der Qualität einhergehen (auch wenn „Zensoric“ hinsichtlich der beiden Konstrukte insgesamt nicht gut abschnitt). Fraglich bleibt allerdings, ob dieser Zusammenhang auch bei einem älteren, weniger gebildeten Datensatz mit geringerem Umweltschulungsbestand hätte.

Analog zur Überprüfung der H2_P wurde die H2_U anhand einer Produkt-Moment-Korrelationsanalyse untersucht, mit der die Richtung und Stärke eines Zusammenhangs zweier metrisch skalierten Variablen bestimmt werden kann. Im Unterschied zu H2_P postulierte H2_U auf Unternehmensebene, dass kein negativer Zusammenhang zwischen der erwarteten Qualität der Produkte des Unternehmens und der Einschätzung der Wahrscheinlichkeit, dass das Unternehmen noch weitere Umweltinitiativen ergreift, besteht. Die Korrelationsanalyse offenbarte einen positiven und signifikanten Zusammenhang zwischen den beiden Variablen in leichtem bis moderatem Ausmaß, $r = .265$, $p = .001$. Wie schon auf Produktebene korrelieren die beiden Bewertungsdimensionen also dahingehend, dass mit steigenden Werten auf dem Halo-Index auch die Werte auf der Qualitätsskala stiegen. Gleiches gilt umgekehrt, indem geringe Werte auf dem Halo-Index auch tendenziell mit geringen Werten auf der Qualitätsskala korrespondieren. Zwei weitere Produkt-Moment-Korrelationsanalysen offenbarten, dass die beiden Variablen nur bei der Experimentalgruppe, die die vage Umweltbotschaft von „Nextvision“ zu sehen bekam, signifikant miteinander korrelierten, $r = .358$, $p = .001$. Eine Korrelation mit der Stärke $r = .199$ verfehlte bei den Versuchspersonen, die das konkrete Argument lasen, knapp die Grenze zur Signifikanz, $p = .067$.

Hintergrund der Vermutung in H2_U war, dass Umweltinitiativen auf Unternehmensebene von Verbrauchern als von den Produkten entkoppelt wahrgenommen werden, sodass sie sich nicht negativ auf die erwartete Produktqualität auswirken. Dies war im Experiment auch nicht der Fall, sodass H2_U

akzeptiert werden kann. Aus den Daten kann aber noch Weiteres herausgelesen werden, da – wie schon auf Produktebene – ein signifikant positiver Zusammenhang zwischen den beiden Variablen aufgedeckt wurde. Dieser deutet darauf hin, dass die Anzeigen auf Unternehmens- und Produktebene insofern ähnlich wirkten, als dass sie einen ersten Eindruck im Sinne eines *General Impression Halos* (s. Kap. 2.2) vermittelten. Insofern verdichtet sich die Schlussfolgerung, dass die Befürchtung, Verbraucher nutzten zur Einschätzung der Produktqualität kompensatorische Inferenzstrategien, indem sie von einer hohen Umweltfreundlichkeit auf eine niedrigere Produktqualität schließen (Chernev & Carpenter, 2001), im konkreten Beispiel und Datensatz unbegründet ist. Die Anzeigen auf Unternehmensebene waren in der Lage, anhand von Umweltargumenten Synergieeffekte zu erzielen, da bessere Bewertungen von Umwelteigenschaften mit einer höheren Einschätzung der Qualität einhergingen, auch wenn generell eine moderate bis niedrige Qualität der Produkte von „Nextvision“ erwartet wurde (erwartete Qualität der Produkte von „Nextvision“: $M = 3.61$, $SD = 1.08$; Halo-Index: $M = 4.73$, $SD = 1.26$). Durch die Unterschiede der Versuchsgruppen hinsichtlich Korrelationskoeffizient und Signifikanz erhärtet sich die auf Produktebene angeführte Interpretation, dass vage Argumente eher als konkrete einen Gesamteindruck bzw. *General Impression Halo* bewirken können, der selbst auf von der Botschaft thematisch weit entfernte Konstrukte wie die Produktqualität abstrahlen kann.

Bei der Suche nach weiteren Erklärungen für das Ausbleiben kompensatorischer Inferenzstrategien ist ein genauerer Blick auf die Stimuli vielversprechend: Bei den Anzeigen auf Produktebene bezogen sich die Botschaften unmittelbar und in positiver Weise auf die Leistungsfähigkeit des Fernsehers (Energieeffizienz), während die Energieeinsparungen in der Produktion, die auf Unternehmensebene thematisiert wurden, nur mittelbar mit den Produkten selbst zusammenhängt. Dies könnte dazu geführt haben, dass weder beim Produkt noch beim Unternehmen kompensatorische Inferenzstrategien (s. Kap. 2.3) festgestellt wurden; in beiden Beispielen nahm die Umweltfreundlichkeit objektiv keinen negativen Einfluss auf die Leistungsfähigkeit des Produktes, was kompensatorische Inferenzstrategien unwahrscheinlicher macht und im Umkehrschluss einen positiven Halo-Effekt begünstigt (Chernev & Blair, 2015, S. 1414). Während diese Entkopplung der Umweltmaßnahmen von den eigentlichen Produkten gerade kennzeichnend für Unternehmenskommunikation ist (Smith et al., 2018, S. 5), existieren also auch auf Produktebene Umwelteigenschaften, die keine Skepsis gegenüber der Qualität eines Produktes erzeugen. Der Jahresstromverbrauch eines Fernsehers könnte zu diesen Produkteigenschaften gehören. Unklar bleibt allerdings, inwiefern das Ergebnis über die Verzerrungen des Datensatzes hinweg generalisierbar ist; denkbar ist nämlich, dass gerade junge, gebildete Personen mit hohem Umweltanliegen von der wahrgenommenen dimensionsübergreifenden Umweltfreundlichkeit auf eine erhöhte Produktqualität schließen. Daher ist es sinnvoll, die Ergebnisse im

Rahmen der Ansprache einer solchen Zielgruppe zu interpretieren, wenn praktische Schlussfolgerungen abgeleitet werden sollen.

5 Diskussion

Tabelle 5 bietet im Anschluss an diese umfassende Datenauswertung einen Überblick über die abgelehnten und gestützten Hypothesen. Die Hypothesen H1.1_P, H1.1_U, H1.2_P und H1.2_U beschäftigten sich mit der Frage, ob die unterschiedlichen Anzeigen in der Lage waren, den Eindruck zu erzeugen, dass das Unternehmen bzw. das Produkt noch weiteren Umweltansprüchen gerecht wird. Hier gab es deutliche Unterschiede zwischen Unternehmens- und Produktebene, nicht aber zwischen vagen und konkreten Botschaften. Die Anzeigen auf Unternehmensebene veranlasste 73 % der Versuchspersonen zu der Annahme, dass die Firma „Nextvision“ neben der beworbenen noch weitere Umweltinitiativen ergreift, während „Zensoric“ mit seiner Produktanzeige nur bei einem Drittel der Befragten vergleichbare Halo-Schlussfolgerungen erzielte. Eine simple Vermutung zur Erklärung dieses Befundes ist, dass die Befragten die Umweltinitiative „Energieeffizienz in der Produktion“, die auf Unternehmensebene zur Werbung genutzt wurde, als weitreichender wahrnahmen als die Energieeffizienz eines einzelnen Gerätes, sodass entsprechend auch eher weitere Umweltdimensionen als wahrscheinlich erachtet wurden. Der Jahresstromverbrauch ist ohnehin ein gängiges Verkaufsargument bei Fernsehern, an das sich Verbraucher womöglich bereits gewöhnt haben, sodass es eventuell nicht auf weitere, speziellere Umwelteigenschaften abstrahlt. In jedem Fall konnte hier die Erkenntnis von Smith et al. (2018), dass stabile CSR-Halo-Effekte auf Unternehmensebene möglich sind, gestützt werden; in den Daten spricht nichts dafür, dass gegenüber Unternehmen in dieser Hinsicht eine größere Skepsis oder Sensibilität für Widersprüche als gegenüber Produkten besteht (s. hierzu Kap. 2.2).

Dass sowohl auf Produkt- als auch auf Unternehmensebene keine signifikanten Unterschiede zwischen den vagen und konkreten Umweltbotschaften hinsichtlich ihrer Wirkung auf die Halo-Indizes festgestellt werden konnten, ist bemerkenswert. Unter Bedingungen geringer Information veranlassten die konkreten Umweltbotschaften die befragten Verbraucher nicht zu einer besseren Bewertung weiterer Umweltdimensionen als die vagen Aussagen. Im Durchschnitt schnitten die beiden vagen Botschaften sogar besser ab als ihre konkreten Pendanten, auch wenn die Differenzen eine Signifikanz verfehlten. Besonders prägnant trat die Wirkung der Greenwashing-Kondition bei dem Experiment auf Unternehmensebene zutage, weil die Befragten hier mehrheitlich positiven Halo-Effekten unterlagen. Folglich reiht sich diese Forschungsarbeit in diejenigen Studien ein, die bei nicht-suggestiven Szenarios eine teilweise Wirksamkeit von vager Umweltkommunikation feststellten (Schmuck et al., 2018; Spack et al., 2012), auch wenn diese Untersuchungen überwiegend „klassischere“ abhängige

Variablen wie die wahrgenommene Umweltfreundlichkeit, wahrgenommene Irreführung oder die Kaufabsicht zum Gegenstand hatten (s. Kap. 2.1). Insofern besteht in der Untersuchung der Halo-Effekte gegenüber diesen Studien eine Weiterentwicklung des Forschungsstandes. Gleichzeitig stützen die Ergebnisse auch die Vermutung (s. hierzu auch Jong et al., 2018, S. 82–86), dass viele der bisherigen suggestiven Studien, die keine Wirksamkeit von Greenwashing feststellen konnten, Schwächen bezüglich der externen Validität aufweisen (s. Kap. 2.1). Diese kommen häufig zu dem Schluss, Greenwashing erziele keine positive Wirkung, weil die Versuchspersonen hier entweder kein einheitliches Bild der Umweltfreundlichkeit eines Produktes oder Unternehmens erhalten, wie dies in der Realität bei der Rezeption von Werbung der Fall ist (z.B. Nyilasy et al., 2014) oder weil die Messungen selbst in übermäßiger Weise für das Thema Greenwashing sensibilisieren (z.B. Chen & Chang, 2013). Die kommende Forschung sollte Greenwashing daher vor allem mit neutral formulierten Messungen anhand von Stimuli untersuchen, in denen keine gezielte Ambiguität der Informationen über ein Produkt oder Unternehmen vorliegen.

Tabelle 5

Überblick über Bewährung und Ablehnung der Hypothesen

Hyp.	Kurzform der Vermutung	Ergebnis
H1.1 _P	Positiver Halo-Effekt bei konkreten <i>und</i> bei vagen Botschaften	Abgelehnt
H1.1 _U	Positiver Halo-Effekt bei konkreten <i>und</i> bei vagen Botschaften	Gestützt
H1.2 _P	Stärkerer positiver Halo-Effekt bei konkreten Botschaften	Abgelehnt
H1.2 _U	Stärkerer positiver Halo-Effekt bei konkreten Botschaften	Abgelehnt
H2 _P	Negativer Zusammenhang zwischen Halo-Index und Einschätzung der Qualität	Abgelehnt
H2 _U	Kein negativer Zusammenhang zwischen Halo-Index und Einschätzung der Qualität	Gestützt

Eine weitere Weiterentwicklung des Forschungsstandes bestand darin, dass die vagen und konkreten Umweltbotschaften dahingehend untersucht wurden, ob sie auf Seiten der Verbraucher kompensatorische Inferenzstrategien auslösen, d.h. ob die Befragten im Zuge positiver Halo-Effekte die Produktqualität bei ihrer Einschätzung abwerten. Das Gegenteil war der Fall, indem der Halo-Index und die Qualitätsskala miteinander korrelierten. Insofern erinnern die Ergebnisse zu den Hypothesen H2_P und H2_U an eine Studie von Chernev und Blair (2015), bei denen die Versuchspersonen von CSR-Wohltätigkeiten auf eine bessere Produktqualität schlossen. Die Daten verdeutlichen insgesamt das

Potenzial der Werbeanzeigen, einen generellen Eindruck zu erzeugen, der auf unterschiedlichste Bewertungsdimensionen abstrahlt. Zu dieser Interpretation gelangt man, wenn man bedenkt, dass es unwahrscheinlich ist, dass die Versuchspersonen wegen einer vermuteten Ähnlichkeit zwischen der dimensionsübergreifenden Umweltfreundlichkeit und der Produktqualität eine Korrelation zwischen diesen beiden Variablen entstehen ließen.

Die Werbeanzeigen scheinen folglich vor allem im Rahmen eines Gesamteindrucks bzw. *General Impression Halo* zu wirken. Ein *Interdimensional Similarity Halo* erscheint unter Bedingungen geringer Informationen weniger relevant zu sein. Insofern sprechen die Ergebnisse gegen die Argumentation von Boatwright et al. (2008), die den *Interdimensional Similarity Halo* im Marketing für bedeutender als den *General Impression Halo* halten (s. Kap. 2.2). Dass die Korrelationen zwischen Halo-Index und erwarteter Produktqualität bei den beiden vagen Experimentalgruppen deutlich stärker ausfielen als bei ihren konkreten Vergleichsgruppen, wurde auf das besondere Potenzial vager Botschaften, einen Gesamteindruck bzw. *General Impression Halo* zu erzielen, der auf thematisch entfernte Konstrukte abstrahlt, zurückgeführt. Im Rückblick auf die zugehörige Theorie (s. Kap. 2.2) kann dies mit den bei vagen Argumenten verstärkten psychologischen Mechanismen des *Undersampling* sowie *Engulfing* (Cooper, 1981, S. 220) erklärt werden.

Auch wenn diese Forschungsarbeit den aktuellen Forschungsstand somit weiterführen konnte, sind einige Limitationen festzuhalten, die die Generalisierbarkeit bzw. Validität der Studie beeinträchtigen. Die Ergebnisse sind insofern im Lichte dieser Limitationen zu bewerten. Die wohl größte Limitation dieser Forschungsarbeit stellt die anfallende Stichprobe dar, die aufgrund der unsystematischen Rekrutierung (s. Kap. 3.3) übermäßig jung, gebildet und umweltbesorgt bei geringem Einkommen war (s. Kap. 3.3.2). Insofern ist eine Generalisierbarkeit der Ergebnisse, beispielsweise auf Verbraucher in Deutschland als Grundgesamtheit, nicht möglich. Zwar ist dies für den Vergleich zwischen zwei Experimentalgruppen grundsätzlich zweitrangig (Koch et al., 2019, S. 118–119), jedoch schlossen nicht alle Hypothesen einen solchen Vergleich ein. Letztlich sind die Ergebnisse dieser Forschungsarbeit also eher als Untersuchung einer Zielgruppe von Umweltbotschaften zu verstehen und gewinnen vor diesem Hintergrund eher Gehalt als bei Versuchen einer Generalisierbarkeit.

Mit einiger Vorsicht nahm diese Arbeit interpretative Vergleiche zwischen zwei Experimenten vor (Produktebene und Unternehmensebene). Trotz dieser Vorsicht kann die Vergleichbarkeit der beiden Stimulusgruppen einschließlich der anschließenden Messungen in Zweifel gezogen werden. Unklar bleibt außerdem, inwiefern zahlenbasierte Argumentationen in konkreten Umweltbotschaften gewisse Kompetenzen zur Einordnung erfordern, sodass diese nur dann von Verbrauchern als

belastbarer und eindeutiger verstanden werden können. Eine Quantifizierung ist nicht zwangsläufig vorteilhaft für Verbraucher, wenn diese ein Umweltargument dadurch nicht besser einordnen können. Den Jahresstromverbrauch eines Fernsehers in kWh können vermutlich nur wenige ohne Referenz einordnen; gleiches gilt für Energieeinsparungen in der Produktion als Prozentangabe, da unklar ist, ob nicht vielleicht jedes Unternehmen solche Einsparungen aufgrund von Auflagen und technischem Fortschritt leisten muss bzw. kann oder ob die Energieeffizienz des Unternehmens nicht vielleicht vorher so schlecht war, dass Energieeinsparungen schnell und mit einfachen Mitteln erzielt werden konnten. Wenn sie eine noch bessere Einordnung ihrer Kennwerte im Vergleich mit dem Markt leisten, könnten konkrete Anzeigen womöglich eine verbesserte Wirkung erzeugen (bei dem hier verwendeten Beispiel auf Produktebene wäre die Energieverbrauchskennzeichnung mit Angabe der Energieeffizienzklasse ohnehin verpflichtend gewesen). Dies ist eine vielversprechende Richtung für die kommende Forschung. Allerdings besteht aus praktischer Sicht dabei die Gefahr, dass man mit einer derartigen Informationsdichte nur noch Personen mit hoher Kaufabsicht und positiven Einstellungen gegenüber grünen Produkten erreicht. Zumindest legt dies eine Untersuchung von Matthes, Wonneberger und Schmuck (2014), die „funktionale Ads“ mit sechs Argumenten zum Gegenstand hatte, nahe.

Zu diesen tiefgreifenden Limitationen treten weitere, die mit den Entscheidungen im Rahmen der Ausgestaltung der Stimuli und des Erhebungsinstrumentes zusammenhängen. Darunter fällt der Umstand, dass nicht kontrolliert wurde, inwiefern es Unterschiede in der Glaubwürdigkeit der konkreten und vagen Umweltbotschaften gab. Eine solche Erhebung hätte die Versuchspersonen zu einer übermäßig kritischen Reflexion der Werbeanzeigen veranlasst, weswegen darauf verzichtet wurde. In einem Design, in dem nur ein einzelnes Produkt oder Unternehmen untersucht worden wäre, hätte eine solche Erhebung nachgestellt stattfinden können. Zum Beispiel gab es bei der konkreten Produktanzeige von „Zensoric“ keine Angabe einer EU-Energieverbrauchskennzeichnung, an die sich die Verbraucher womöglich schon gewöhnt haben; dies könnte die Glaubwürdigkeit der Anzeige belastet haben. Da zwei Stimuli in randomisierter Reihenfolge im Fragebogen enthalten waren, ist es darüber hinaus möglich, dass Positioneffekte unentdeckt blieben, die sich danach bestimmen, welche der beiden Bezugsebenen der Kommunikation im Experiment zuerst thematisiert wurden (s. Koch et al., 2019, S. 69–72). Die zusätzliche Erhebung und die anschließende Auswertung solcher Positioneffekte hätten den Rahmen dieser Forschungsarbeit gesprengt. Nicht zu vernachlässigen ist außerdem die Tatsache, dass der Fragebogen die Versuchspersonen notwendigerweise dazu aufforderte, Bewertungen bezüglich der weiteren Umweltbewertungsdimensionen vorzunehmen. Unklar bleibt somit, ob durch das Messinstrument selbst ein bewusster Meinungsbildungsprozess eingeleitet wurde, der in

der Realität ausgeblieben wäre. Insbesondere ist es wahrscheinlich, dass Verbraucher sich gar keine Gedanken über weitere Umwelteigenschaften der Produkte und Unternehmen machen, sondern ihren Gesamteindruck in unbewusster und automatischer Weise akzeptieren. Der Stellenwert von CSR-Halo-Effekten für Kaufentscheidungen oder für die Entwicklung von Einstellungen bezüglich Marken oder Produkten bleibt auch nach dieser Forschungsarbeit unbestimmt.

Im vorliegenden Experimentaldesign konnte jede Abweichung vom Halo-Indexmittelpunkt – der die absolute Ungewissheit einer Versuchsperson hinsichtlich der zu bewertenden Umwelteigenschaft bzw. -initiative kennzeichnete – nur deshalb als irrational eingestuft werden, weil die Unternehmen und Produkte rein fiktiv waren und somit eine etwaige „echte Korrelation“ der Bewertungsdimensionen vernachlässigt werden konnte. In der Realität, in der Verbraucher notwendigerweise etliche Entscheidungen für oder gegen Produkte und Unternehmen treffen müssen, können Halo-Effekte über Kaufentscheidungen hinweg durchaus effektive Bewertungsinstrumente sein (Boatwright et al., 2008). Insofern konnte diese Arbeit nicht beantworten, inwiefern Halo-Effekte (über das hier behandelte, isolierte Beispiel hinaus) die Stellung der Verbraucher als informierte Marktteilnehmer behindern.

Letztlich wird die Generalisierbarkeit der Studie auch durch den Fokus auf einen Produkttyp innerhalb einer Branche limitiert. Torelli et al. (2020) legen eindrücklich dar, dass Greenwashing in Branchen mit hoher Umweltrelevanz nachteiliger für Werbetreibende sein kann; insofern sollte von der teilweisen Wirksamkeit der Greenwashing-Kondition in dieser Studie nicht auf eine entsprechende branchenübergreifende Wirksamkeit geschlossen werden. Dieser Umstand verdeutlicht, dass auch hinsichtlich Halo-Effekten systematische Branchenvergleiche in weiteren Forschungsarbeiten wünschenswert sind.

Die erhobenen Daten sprechen dafür, dass sich konkrete Umweltargumente – sei es für Unternehmen oder für ihre Produkte im Speziellen – unter Bedingungen geringer Information nicht gegen vage Botschaften durchsetzen können, wenn es darum geht, einen umweltfreundlichen Gesamteindruck zu erzielen. Unter „geringer Information“ verstand diese Forschungsarbeit eine einzige Werbeanzeige inkl. Slogan. Für kommunizierende Unternehmen bedeutet dies, dass sie sich in dieser Hinsicht, sofern sie konkrete Maßnahmen vorweisen können, umfangreicher von der Konkurrenz abgrenzen müssen. Das Ergebnis deutet also darauf hin, dass größer angelegte Kampagnen mit mehreren Customer Touchpoints und mehreren beworbenen Umweltmaßnahmen bzw. Produkteigenschaften nötig sind, um mit konkreten Umweltargumenten einen Vorteil bei der Erzeugung eines umweltfreundlichen Gesamteindrucks erzielen zu können. Diese Schlussfolgerung soll aber nicht darüber hinwegtäuschen,

dass konkrete Argumente hinsichtlich anderer Variablen wie z.B. der Glaubwürdigkeit vorteilhaft sein können (bspw. Spack et al., 2012).

Ein zweites aus praktischer Sicht bemerkenswertes Ergebnis dieser Forschungsarbeit war, dass die Einschätzung der Probanden, ein Unternehmen oder Produkt habe noch weitere Umwelteigenschaften, positiv mit der Einschätzung der Produktqualität korrelierte. Die Befürchtung, dass Verbraucher von technischen Produkten eine geringere Leistungsfähigkeit erwarten, wenn sie auf Basis eines ersten Eindruckes viele Umwelteigenschaften vermuten, war in diesen Kontext und Datensatz unbegründet. Insofern sprechen die Daten dafür, dass Anbieter technischer Produkte Umweltargumente problemlos auch in isolierter Form verwenden können und auf diese Weise sogar mittelbar positive Erwartungen der Qualität der Produkte erzeugen können, wenn durch die Anzeige ein umweltfreundlicher Eindruck bei den Verbrauchern entsteht. Dies gilt zumindest bei der Ansprache von Zielgruppen, die den Versuchspersonen dieser Studie ähneln, d.h. die jung, gebildet und umweltbesorgt sind. Trotz anhaltender Unklarheiten ließen sich aus den Ergebnissen dieser Studie somit Ratschläge für Marketingfachleute und PR-Praktiker sowie Empfehlungen für die kommende Forschung ableiten.

Anhangsverzeichnis

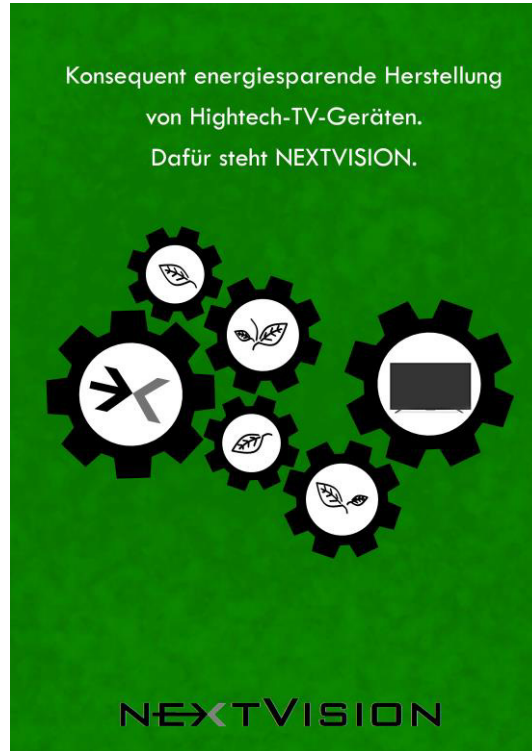
Anhang 1: Stimuli für das Experiment auf Produktebene	41
Anhang 2: Stimuli für das Experiment auf Unternehmensebene	42
Anhang 3: Fragebogen ohne Stimuli	43

Anhang 1: Stimuli für das Experiment auf Produktebene



Anmerkung. Links: konkrete Produktanzeige mit quantifizierter Umweltbotschaft; rechts: vage Produktanzeige.

Anhang 2: Stimuli für das Experiment auf Unternehmensebene



Anmerkung. Links: konkrete Unternehmensanzeige mit quantifizierter Umweltbotschaft; rechts: vage Umweltanzeige.

Anhang 3: Fragebogen ohne Stimuli

0% ausgefüllt

Herzlich Willkommen...

...und vielen Dank für Ihr Interesse an dieser Studie!

Im nachfolgenden Fragebogen werden Sie gebeten, junge Unternehmen und ihre Produkte nach einem ersten Eindruck zu bewerten. Hierzu werden Ihnen einige Werbeanzeigen vorgestellt. Ihre Stimme als Verbraucher*in ist für die Unternehmen sehr wichtig.

Das Ausfüllen des Fragebogens dauert ca. 10 Minuten. Selbstverständlich werden Ihre Angaben vollständig anonym erfasst.

Bitte nutzen Sie währenddessen nicht den "Zurück-Button" Ihres Browsers. Mit Klick auf "Weiter" startet die Befragung.

Weiter

[B.A. Marc Möller-Boldt](#), Student an der Hochschule Bonn-Rhein-Sieg, Grantham-Allee 20,
53757 Sankt Augustin – 2020

6% ausgefüllt

Nachfolgend sehen Sie eine Werbeanzeige der Firma NEXTVISION.
Anschließend werden Ihnen einige Fragen zu Ihrem ersten Eindruck von diesem Unternehmen gestellt.

Mit Klick auf "Weiter" erscheint die Anzeige, die Sie in beliebiger Dauer betrachten können.

Weiter

1. Bitte geben Sie Ihre Zustimmung zu den folgenden Aussagen an:

Bitte legen Sie nun Ihren ersten Eindruck des Unternehmens dar. Entscheiden Sie nach Ihrem Bauchgefühl, das Sie durch die Werbeanzeige gewonnen haben.

	stimme gar nicht zu						stimme voll zu
Die Fernseher von NEXTVISION haben eine hohe Qualität.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Fernseher von NEXTVISION haben eine hohe Bildqualität.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Fernseher von NEXTVISION bieten für verschiedene Anwendungen (Spielfilme, Sport, Gaming) eine hohe Leistung.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Fernseher von NEXTVISION können sich mit den leistungsfähigsten TV-Geräten auf dem Markt messen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Fernseher von NEXTVISION haben qualitative Mängel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Fernseher von NEXTVISION haben einen hohen Kontrastumfang.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Weiter](#)

24% ausgefüllt

2. Wie wahrscheinlich ist es, dass NEXTVISION Maßnahmen in den folgenden Bereichen ergreift?

Bitte legen Sie nun Ihren ersten Eindruck des Unternehmens dar. Entscheiden Sie nach Ihrem Bauchgefühl, das Sie durch die Werbeanzeige gewonnen haben.

	sehr unwahr- scheinlich	sehr wahrscheinlich
NEXTVISION produziert so, dass möglichst viele Materialien und Bauteile wiederverwendet werden können.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	
NEXTVISION stellt einen energieeffizienten Transport sicher, indem es nach Möglichkeit das Gewicht von Geräten und Bauteilen reduziert.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	
NEXTVISION stellt sicher, dass die Lebensdauer seiner Produkte möglichst lang anhält.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	
NEXTVISION minimiert Schadstoffe wie Quecksilber und halogenierte Flammschutzmittel.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	
NEXTVISION nutzt Verpackungen, die zum Teil aus recyceltem Material bestehen.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	

Weiter

29% ausgefüllt

Nachfolgend sehen Sie eine Werbeanzeige der Firma ZENSORIC. Anschließend werden Ihnen einige Fragen zu Ihrem ersten Eindruck von den Produkten dieses Unternehmens gestellt.

Mit Klick auf "Weiter" erscheint die Anzeige, die Sie in beliebiger Dauer betrachten können.

Weiter

3. Bitte geben Sie Ihre Zustimmung zu den folgenden Aussagen an:

Bitte legen Sie nun Ihren ersten Eindruck des Fernsehers dar. Entscheiden Sie nach Ihrem Bauchgefühl, das Sie durch die Werbeanzeige gewonnen haben.

	Stimme gar nicht zu						Stimme voll zu
Der Omro UHD-Fernseher hat einen hohen Kontrastumfang.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher kann sich mit den leistungsfähigsten TV-Geräten auf dem Markt messen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher hat eine hohe Qualität.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher hat eine hohe Bildqualität.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher bietet für verschiedene Anwendungen (Spielfilme, Sport, Gaming) eine hohe Leistung.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher hat qualitative Mängel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

[Weiter](#)

47% ausgefüllt

4. Wie wahrscheinlich ist es, dass der Omro UHD-Fernseher die folgenden Eigenschaften hat?

Bitte legen Sie nun Ihren ersten Eindruck des Fernsehers dar. Entscheiden Sie nach Ihrem Bauchgefühl, das Sie durch die Werbeanzeige gewonnen haben.

	sehr unwahr- scheinlich						sehr wahrscheinlich
Der Omro UHD-Fernseher wurde so hergestellt, dass möglichst viele Materialien und Bauteile wiederverwendet werden können.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher enthält nur minimale Mengen an Schadstoffen wie Quecksilber und halogenierte Flammschutzmittel.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Verpackung des Omro UHD-F Fernsehers besteht zum Teil aus recyceltem Material.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher hat ein niedriges Gewicht, sodass beim Transport weniger Energie verbraucht wird.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Omro UHD-Fernseher hat eine lange Lebensdauer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Weiter

53% ausgefüllt

5. Bitte geben Sie Ihre Zustimmung zu den folgenden Aussagen an:

	stimme gar nicht zu						stimme voll zu
Ich weiß viel über Fernseher.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich interessiere mich für verschiedene Marken von Fernsehern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin offen für neue Marken von Fernsehern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe nicht vor, in der Zukunft einen Fernseher zu kaufen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich interessiere mich nicht für Fernseher.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Weiter

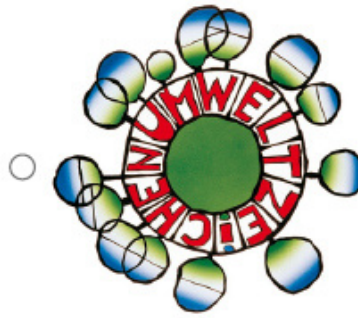
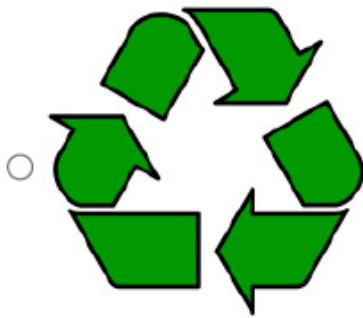
6. Bitte geben Sie Ihre Zustimmung zu den folgenden Aussagen an:

	stimme gar nicht zu	stimme voll zu
Ich bin bereit, Opfer für den Schutz der Umwelt zu bringen.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	
Ich mache mir Sorgen um die Umwelt.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	
Der Zustand der Umwelt beeinflusst meine Lebensqualität.	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	

Weiter

Als nächstes stellen wir Ihnen einige Fragen, in denen es um Ihr Wissen über Umweltthemen geht. Bei jeder Frage gibt es jeweils nur eine richtige Antwortmöglichkeit. **Wir möchten Sie bitten, währenddessen nicht nach den Lösungen zu recherchieren.** Nur so kann diese Studie gültige Ergebnisse erzielen. Die Lösungen können Sie am Ende der Befragung einsehen.

7. Welches der folgenden Symbole kennzeichnet ein Produkt als recycelbar?



Weiß nicht

Weiter

8. Was bedeutet es, wenn das folgende Symbol auf einer Verpackung erscheint?



- Dies ist kein offizielles Zeichen zur Kennzeichnung umweltfreundlicher Produkte.
- Dies ist ein offizielles Siegel, das umweltfreundliche Produkte kennzeichnet.
- Die Herstellung dieses Produktes ist umweltfreundlich.
- Die Herstellung dieses Produktes fügt Pflanzen keinen Schaden zu.

Weiß nicht

Weiter

9. Im Folgenden sehen Sie vier Siegel für organische Textilien. Welches von ihnen stellt eine Zertifizierung durch eine unabhängige Drittpartei dar?



Weiß nicht

Weiter

10. Was versteht man unter dem CO₂-Fußabdruck eines Produktes?

- Die Menge von CO₂, die ein Produkt freisetzt, wenn es sich zersetzt.
 - Die Menge aller Treibhausgasemissionen, die entlang des Lebenszyklus eines Produktes entstehen.
 - Die chemische Änderung, die durch CO₂ in der Atmosphäre hervorgerufen wird.
 - Die typische Färbung des Himmels, die durch eine hohe CO₂-Konzentration verursacht wird.
-
- Weiß nicht

11. Erneuerbare Energien werden aus erneuerbaren Ressourcen gewonnen, die sich innerhalb menschlicher Zeitmaßstäbe wieder auffüllen. Welche der folgenden Energiequellen bietet keine erneuerbare Energie?

- Deponiegas
- Erdgas
- Gasförmige Biomasse

Weiß nicht

12. Welches Leuchtmittel ist am energieeffizientesten (am meisten Lumen/Watt)?

- Halogenlampe
- LED (Licht emittierende Dioden)
- Herkömmliche Glühlampe
- Energiesparlampen

Weiß nicht

Weiter

Zuletzt möchten wir Sie um einige Angaben zu Ihrer Person bitten. Selbstverständlich werden Ihre Antworten weiterhin vollständig anonym erfasst.

13. Welches biologische Geschlecht haben Sie?

weiblich

männlich

anderes, und zwar:

14. Wie alt sind Sie?

Ich bin Jahre

Weiter

"5. Erneuerbare Energien werden aus erneuerbaren Ressourcen gewonnen, die sich innerhalb menschlicher Zeitmaßstäbe wieder auffüllen. Welche der folgenden Energiequellen bietet

94% ausgefüllt

15. Welchen Bildungsabschluss haben Sie?

Bitte wählen Sie den höchsten Bildungsabschluss, den Sie bisher erreicht haben.

- (Noch) kein Abschluss
- Volks-, Hauptschulabschluss, Quali
- Mittlere Reife, Realschul- oder gleichwertiger Abschluss
- Fachabitur, Fachhochschulreife
- Abitur, Hochschulreife
- Fachhochschul-/Hochschulabschluss
- Anderer Abschluss, und zwar:

16. Wie hoch ist ungefähr Ihr monatliches Nettoeinkommen?

Gemeint ist der Betrag, der sich aus allen Einkünften zusammensetzt und nach Abzug der Steuern und Sozialversicherungen übrig bleibt.

[Bitte auswählen] ▼

[Bitte auswählen]

Ich habe kein eigenes Einkommen
weniger als 250 €
250 € bis unter 500 €
500 € bis unter 1000 €
1000 € bis unter 1500 €
1500 € bis unter 2000 €
2000 € bis unter 2500 €
2500 € bis unter 3000 €
3000 € bis unter 3500 €
3500 € bis unter 4000 €
4000 € oder mehr
Ich will darauf nicht antworten

Weiter

Vielen Dank für Ihre Teilnahme!

Ich möchte mich ganz herzlich für Ihre Mithilfe bedanken. In dieser Studie wird untersucht, wie Umweltargumente auf Verbraucher*innen wirken.

Alle genutzten Unternehmen und Produkte sind rein fiktiv.

Dieser Fragebogen ist Teil einer Abschlussarbeit im Master-Studiengang "Technik- und Innovationskommunikation" an der Hochschule Bonn-Rhein-Sieg, die Betreuung obliegt Frau Prof. Dr. Susanne Keil. Wenden Sie sich bei Fragen oder Anregungen gerne an den Verfasser der Studie, Marc Möller-Boldt: marc.moeller2@smail.h-brs.de

Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen.

Lösungen zum Umweltwissenstest:

"1. Welches der folgenden Symbole kennzeichnet ein Produkt als recycelbar?"



"2. Was bedeutet es, wenn das oben abgebildete Symbol auf einer Verpackung erscheint?"
(schwarz-weißes Baumsymbol)

Lösung: Dies ist kein offizielles Zeichen zur Kennzeichnung umweltfreundlicher Produkte.

"3. Im Folgenden sehen Sie vier Siegel für organische Textilien. Welches von ihnen stellt eine Zertifizierung durch eine unabhängige Drittpartei dar?"



"4. Was versteht man unter dem CO₂-Fußabdruck eines Produktes?"

Lösung: Die Menge aller Treibhausgasemissionen, die entlang des Lebenszyklus eines Produktes entstehen.

keine erneuerbare Energie?"

Lösung: Erdgas

"6. Welches Leuchtmittel ist am energieeffizientesten (am meisten Lumen/Watt)?"

Lösung: LED (Licht emittierende Dioden)

Literaturverzeichnis

- Atkinson, L. & Kim, Y. (2015). "I Drink It Anyway and I Know I Shouldn't": Understanding Green Consumers' Positive Evaluations of Norm-violating Non-green Products and Misleading Green Advertising. *Environmental Communication*, 9(1), 37–57.
<https://doi.org/10.1080/17524032.2014.932817>
- Avcilar, M. Y. & Demirgünes, B. K. (2016). Developing Perceived Greenwash Index and Its Effect on Green Brand Equity: A Research on Gas Station Companies in Turkey. *International Business Research*, 10(1), 222. <https://doi.org/10.5539/ibr.v10n1p222>
- Banerjee, S., Gulas, C. S. & Iyer, E. (1995). Shades of Green: A Multidimensional Analysis of Environmental Advertising. *Journal of Advertising*, 24(2), 21–31.
<https://doi.org/10.1080/00913367.1995.10673473>
- Baum, L. M. (2012). It's Not Easy Being Green ... Or Is It? A Content Analysis of Environmental Claims in Magazine Advertisements from the United States and United Kingdom. *Environmental Communication*, 6(4), 423–440. <https://doi.org/10.1080/17524032.2012.724022>
- Boatwright, P., Kalra, A. & Zhang, W. (2008). Research Note —Should Consumers Use the Halo to Form Product Evaluations? *Management Science*, 54(1), 217–223.
<https://doi.org/10.1287/mnsc.1070.0742>
- Burke, P. F., Dowling, G. & Wei, E. (2018). The relative impact of corporate reputation on consumer choice: beyond a halo effect. *Journal of Marketing Management*, 34(13-14), 1227–1257.
<https://doi.org/10.1080/0267257X.2018.1546765>
- Chen, Y.-S. & Chang, C.-H. (2013). Greenwash and Green Trust: The Mediation Effects of Green Consumer Confusion and Green Perceived Risk. *Journal of Business Ethics*, 114(3), 489–500.
<https://doi.org/10.1007/s10551-012-1360-0>
- Chen, Y.-S., Lin, C.-L. & Chang, C.-H. (2014). The influence of greenwash on green word-of-mouth (green WOM): the mediation effects of green perceived quality and green satisfaction. *Quality & Quantity*, 48(5), 2411–2425. <https://doi.org/10.1007/s11135-013-9898-1>
- Chernev, A. & Blair, S. (2015). Doing Well by Doing Good: The Benevolent Halo of Corporate Social Responsibility. *Journal of Consumer Research*, 41(6), 1412–1425.
<https://doi.org/10.1086/680089>
- Chernev, A. & Carpenter, G. S. (2001). The Role of Market Efficiency Intuitions in Consumer Choice: A Case of Compensatory Inferences. *Journal of Marketing Research*, 38(3), 349–361.
<https://doi.org/10.1509/jmkr.38.3.349.18865>

- Christophersen, T. & Grape, C. (2009). Die Erfassung latenter Konstrukte mit Hilfe formativer und reflektiver Messmodelle. In S. Albers, D. Klapper, U. Konradt, A. Walter & J. Wolf (Hrsg.), *Methodik der empirischen Forschung* (S. 103–118). Wiesbaden: Gabler Verlag. Zugriff am 10.09.2020. Verfügbar unter: https://link.springer.com/chapter/10.1007/978-3-322-96406-9_8
- Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin*, 90(2), 218–244.
<https://doi.org/10.1037/0033-2909.90.2.218>
- Delmas, M. A. & Burbano, V. C. (2011). The Drivers of Greenwashing. *California Management Review*, 54(1), 64–87. <https://doi.org/10.1525/cm.2011.54.1.64>
- Diekmann, A. (1997). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen* (3., durchgesehene Auflage). Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl.
- Döring, N. (2013). Zur Operationalisierung von Geschlecht im Fragebogen: Probleme und Lösungsansätze aus Sicht von Mess-, Umfrage-, Gender- und Queer-Theorie. *GENDER - Zeitschrift für Geschlecht, Kultur und Gesellschaft*, 5(2), 94–113. Zugriff am 29.08.2020. Verfügbar unter: <https://www.budrich-journals.de/index.php/gender/article/download/17854/15529>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th edition). Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE.
- Freitas Netto, S. V. de, Sobral, M. F. F., Ribeiro, A. R. B. & Soares, G. R. d. L. (2020). Concepts and forms of greenwashing: a systematic review. *Environmental Sciences Europe*, 32(1).
<https://doi.org/10.1186/s12302-020-0300-3>
- Geiger, S. & Holzhauer, B. (2020, Februar). *Weiterentwicklung einer Skala zur Messung von zentralen Kenngrößen des Umweltbewusstseins* (Texte 25). Dessau-Roßlau: Umweltbundesamt. Zugriff am 29.08.2020. Verfügbar unter: <https://www.umweltbundesamt.de/publikationen/weiterentwicklung-skala-umweltbewusstsein>
- Gosselt, J. F., van Rompay, T. & Haske, L. (2019). Won't Get Fooled Again: The Effects of Internal and External CSR ECO-Labeling. *Journal of Business Ethics*, 155(2), 413–424.
<https://doi.org/10.1007/s10551-017-3512-8>
- Guo, R., Tao, L., Li, C. B. & Wang, T. (2017). A Path Analysis of Greenwashing in a Trust Crisis Among Chinese Energy Companies: The Role of Brand Legitimacy and Brand Loyalty. *Journal of Business Ethics*, 140(3), 523–536. <https://doi.org/10.1007/s10551-015-2672-7>
- Jong, M. D. T. de, Harkink, K. M. & Barth, S. (2018). Making Green Stuff? Effects of Corporate Greenwashing on Consumers. *Journal of Business and Technical Communication*, 32(1), 77–112. <https://doi.org/10.1177/1050651917729863>

- Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin Books.
- Kardes, F. R., Posavac, S. S. & Cronley, M. L. (2004). Consumer Inference: A Review of Processes, Bases, and Judgment Contexts. *Journal of Consumer Psychology*, 14(3), 230–256.
- Koch, T., Peter, C. & Müller, P. (2019). *Das Experiment in der Kommunikations- und Medienwissenschaft*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-19754-4>
- Lee, J., Bhatt, S. & Suri, R. (2018). When consumers penalize not so green products. *Psychology & Marketing*, 35(1), 36–46. <https://doi.org/10.1002/mar.21069>
- Leiner, D. J. (2019). Too Fast, too Straight, too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys. 229-248 Pages / Survey Research Methods, Vol 13 No 3 (2019) / Survey Research Methods, Vol 13 No 3 (2019). <https://doi.org/10.18148/srm/2019.v13i3.7403>
- Leiner, D. J. (2020). SoSci Survey (Version 3.2.10) [Computer software]. Verfügbar unter: www.soscisurvey.de
- Luchs, M. G., Naylor, R. W., Irwin, J. R. & Raghunathan, R. (2010). The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference. *Journal of Marketing*, 74(5), 18–31.
- Lyon, T. P. & Montgomery, A. W. (2015). The Means and End of Greenwash. *Organization & Environment*, 28(2), 223–249. <https://doi.org/10.1177/1086026615575332>
- Matthes, J., Wonneberger, A. & Schmuck, D. (2014). Consumers' green involvement and the persuasive effects of emotional versus functional ads. *Journal of Business Research*, 67(9), 1885–1893. <https://doi.org/10.1016/j.jbusres.2013.11.054>
- Mohr, L. A., Eroğlu, D. & Ellen, P. S. (1998). The Development and Testing of a Measure of Skepticism Toward Environmental Claims in Marketers' Communications. *Journal of Consumer Affairs*, 32(1), 30–55. <https://doi.org/10.1111/j.1745-6606.1998.tb00399.x>
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38(4), 379–387. <https://doi.org/10.1037/0003-066X.38.4.379>
- Newman, G. E., Gorlin, M. & Dhar, R. (2014). When Going Green Backfires: How Firm Intentions Shape the Evaluation of Socially Beneficial Product Enhancements. *Journal of Consumer Research*, 41(3), 823–839. <https://doi.org/10.1086/677841>
- Nisbett, R. E. & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35(4), 250–256. <https://doi.org/10.1037/0022-3514.35.4.250>

- Nyilasy, G., Gangadharbatla, H. & Paladino, A. (2012). Greenwashing: a consumer perspective. *Economics & Sociology*, 5(2), 116–123.
- Nyilasy, G., Gangadharbatla, H. & Paladino, A. (2014). Perceived Greenwashing: The Interactive Effects of Green Advertising and Corporate Environmental Performance on Consumer Reactions. *Journal of Business Ethics*, 125(4), 693–707. <https://doi.org/10.1007/s10551-013-1944-3>
- Pancer, E., McShane, L. & Noseworthy, T. J. (2017). Isolated Environmental Cues and Product Efficacy Penalties: The Color Green and Eco-labels. *Journal of Business Ethics*, 143(1), 159–177. <https://doi.org/10.1007/s10551-015-2764-4>
- Parguel, B., Benoit-Moreau, F. & Russell, C. A. (2015). Can evoking nature in advertising mislead consumers? The power of 'executional greenwashing'. *International Journal of Advertising*, 34(1), 107–134. <https://doi.org/10.1080/02650487.2014.996116>
- Parguel, B., Benoît-Moreau, F. & Larceneux, F. (2011). How Sustainability Ratings Might Deter 'Greenwashing': A Closer Look at Ethical Corporate Communication. *Journal of Business Ethics*, 102(1), 15–28. <https://doi.org/10.1007/s10551-011-0901-2>
- Rahman, I., Park, J. & Chi, C. G. (2015). Consequences of "greenwashing". *International Journal of Contemporary Hospitality Management*, 27(6), 1054–1081. <https://doi.org/10.1108/IJCHM-04-2014-0202>
- Schmidt, A. & Donsbach, W. (2012). „Grüne“ Werbung als Instrument für „schwarze“ Zahlen. *Publizistik*, 57(1), 75–93. <https://doi.org/10.1007/s11616-012-0140-2>
- Schmuck, D., Matthes, J. & Naderer, B. (2018). Misleading Consumers with Green Advertising? An Affect–Reason–Involvement Account of Greenwashing Effects in Environmental Advertising. *Journal of Advertising*, 47(2), 127–145. <https://doi.org/10.1080/00913367.2018.1452652>
- Schuhwerk, M. E. & Lefkoff-Hagius, R. (1995). Green or Non-Green? Does Type of Appeal Matter When Advertising a Green Product? *Journal of Advertising*, 24(2), 45–54. <https://doi.org/10.1080/00913367.1995.10673475>
- Seele, P. & Gatti, L. (2017). Greenwashing Revisited: In Search of a Typology and Accusation-Based Definition Incorporating Legitimacy Strategies. *Business Strategy and the Environment*, 26(2), 239–252. <https://doi.org/10.1002/bse.1912>
- Segev, S., Fernandes, J. & Hong, C. (2016). Is Your Product Really Green? A Content Analysis to Reassess Green Advertising. *Journal of Advertising*, 45(1), 85–93. <https://doi.org/10.1080/00913367.2015.1083918>

- Smith, N. C., Read, D. & Lopez, S. (2018). CSR Halo: The Gift that Keeps on Giving? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3135132>
- SoSci Survey GmbH. (2019, 13. Februar). *Technischer Funktionstest*, SoSci Survey GmbH. Benutzerhandbuch. Zugriff am 11.10.2020. Verfügbar unter: <https://www.soscisurvey.de/help/doku.php/de:survey:function-test>
- Spack, J. A., Board, V. E., Crighton, L. M., Kostka, P. M. & Ivory, J. D. (2012). It's Easy Being Green: The Effects of Argument and Imagery on Consumer Responses to Green Product Packaging. *Environmental Communication*, 6(4), 441–458. <https://doi.org/10.1080/17524032.2012.706231>
- Szabo, S. & Webster, J. (2020). Perceived Greenwashing: The Effects of Green Marketing on Environmental and Product Perceptions. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-020-04461-0>
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4, 25–29.
- Torelli, R., Balluchi, F. & Lazzini, A. (2020). Greenwashing and environmental communication: Effects on stakeholders' perceptions. *Business Strategy and the Environment*, 29(2), 407–421. <https://doi.org/10.1002/bse.2373>
- Umweltbundesamt. (2020). *Fernseher*, Umweltbundesamt. Zugriff am 27.07.2020. Verfügbar unter: <https://www.umweltbundesamt.de/umwelttipps-fuer-den-alltag/elektrogeraete/fernseher>
- Wells, F. L. (1907). A statistical study of literary merit. *Archives of Psychology*, 1(7).