



Studienabschlussarbeiten

Faculty of Social Sciences

UNSPECIFIED

Schulze, Jonas:

Qualität aus Sicht der Nutzer: Welchen Einfluss haben Nachrichtenmarken und Verbreitungswege auf die Wahrnehmung journalistischer Beiträge?

Eine quantitative Studie

Master of Arts, Summer Semester 2020

Gutachter*in: Hanitzsch, Thomas

Faculty of Social Sciences

Institut für Kommunikationswissenschaft und Medienforschung

Master Journalismus

Ludwig-Maximilians-Universität München

<https://doi.org/10.5282/ubm/epub.77662>



Qualität aus Sicht der Nutzer:

Welchen Einfluss haben Nachrichtenmarken und
Verbreitungswege auf die Wahrnehmung journalistischer
Beiträge?

Eine quantitative Studie.

Wissenschaftliche Arbeit zur Erlangung des Grades Master of Arts (M.A.)

Gutachter: Prof. Dr. Thomas Hanitzsch

Vorgelegt von: Jonas Bastian Schulze

Studiengang: M.A. Journalismus

München, Juli 2020

Kurzfassung

Diese Arbeit untersucht, in wie weit externe Faktoren wie die Marke oder der Verbreitungsweg die Qualitätswahrnehmungen von Nachrichtennutzern beeinflussen. Dazu wurde eine Serie von 2x2 Experimenten durchgeführt (N = 158), in denen die Teilnehmer fünf Nachrichtenbeiträge aus den Ressorts Politik, Wirtschaft, Gesundheit, Digitales und Medien bewerten sollten. Das Design der Beiträge wurde so manipuliert, dass von jedem Beitrag vier Versionen entstanden, die sich hinsichtlich der Nachrichtenmarke und des Verbreitungswegs (Soziale Medien oder Webseite eines Anbieters) unterschieden. Anschließend wurden die Teilnehmer mit zufällig ausgewählten Versionen der Texte und Videos konfrontiert. Bei der Datenerhebung wurden zwei Methoden eingesetzt: Erstens wurde die generische Qualitätsbewertung der Teilnehmer mit einer siebenstufigen Antwortskala abgefragt und zweitens wurden die Antworten auf fünf Teilfragen zu einer Likert-Skala kombiniert. Die Bewertungen der Teilnehmer wurden elektronisch erfasst und mit statistischen Methoden ausgewertet. Die Ergebnisse des Experiments deuten stark daraufhin, dass sowohl die Medienmarke als auch der Verbreitungsweg die normativen Bewertungen beeinflussen. Insbesondere die Nachrichtenmarke scheint für viele Menschen eine bedeutende Heuristik zu sein. Bei der Analyse der Zugangswege zeigte sich ein Zusammenhang zwischen den Mediennutzungsgewohnheiten und dem Qualitätsempfinden der Teilnehmer.

Abstract

This work investigates the extent to which external factors such as the brand or the distribution channel influence the quality perceptions of news users. To this end, a series of 2x2 experiments was conducted (N = 158) in which the participants were asked to rate five news items from the political, health, media, digital and economic sector. The design of the items was manipulated in such a way that four versions of each item were created, which differed in terms of news brand and distribution channel (social media or website of a provider). Subsequently, the participants were confronted with randomly selected versions of the texts and videos. Two methods were used to collect the data: firstly, the generic quality rating of the participants was queried using a seven-level answer scale and secondly, the answers to five sub-questions were combined to form a Likert scale. The participants' ratings were recorded electronically and evaluated using statistical methods. The results of the experiment strongly suggest that both the media brand and the distribution channel influence the normative ratings. Especially the news brand seems to be a significant heuristic for many people. The analysis of the access paths revealed a correlation between media usage habits and the participants' perception of quality.

Inhaltsverzeichnis

1 EINLEITUNG	5
2 WAS IST JOURNALISTISCHE QUALITÄT?	7
2.1 ÜBERALL QUALITÄTEN	7
2.2 EXPERTISE IST (NICHT) ALLES	9
2.3 QUALITÄT MESSEN – GEHT DAS ÜBERHAUPT?	11
2.4 JOURNALISTISCHE QUALITÄT: EINE SÄULE DER DEMOKRATIE	13
3 NACHRICHTENNUTZUNG IN DEUTSCHLAND	15
3.1 INTERNETNUTZUNG IN DEUTSCHLAND	15
3.2 NACHRICHTENINTERESSE UND NUTZUNGSHÄUFIGKEIT	15
3.3 FÜHREN ALLE WEGE ZUR NACHRICHT?	17
3.4 VERTRAUEN IN NACHRICHTEN	19
4 EMPIRISCHE QUALITÄTSFORSCHUNG	20
4.1 ANGEBOTSMODELLE	20
4.2 NACHFRAGEMODELLE	22
5 METHODIK	26
5.1 FORSCHUNGSFRAGEN UND HYPOTHESEN	26
5.2 STUDIENDESIGN	28
5.3 ÜBERBLICK ÜBER DIE TEILAUFGABEN	32
5.4 DATENERHEBUNG	37
5.4.1 <i>Likert-Skalen und Antwortskalen des Likert Typs</i>	37
5.4.2 <i>Singuläre und kombinierte Qualitätsbestimmung</i>	38
5.4.3 <i>Gestaltung des Fragebogens</i>	40
5.5 DURCHFÜHRUNG	42
5.6 DATENAUSWERTUNG UND ANALYSE	43
6 ERGEBNISSE	47
6.1 STICHPROBE	47
6.1.1 <i>Alter, Geschlecht und formale Bildung</i>	47
6.1.2 <i>Mediennutzungsverhalten</i>	48
6.1.3 <i>Vertrauen in Nachrichtenanbieter</i>	51

6.2 NACHRICHTENMARKEN	52
6.3 ZUGANGSWEGE	55
7 FAZIT.....	59
7.1 DIE NACHRICHTENMARKE: EINE BEDEUTENDE HEURISTIK	59
7.2 NUTZUNGSVERHALTEN PRÄGT DIE WAHRNEHMUNG	60
7.3 LIMITATIONEN	61
I TABELLENVERZEICHNIS	63
II ABBILDUNGSVERZEICHNIS.....	64
III LITERATURVERZEICHNIS	65

1 Einleitung

Medienhäuser und Verlage stehen seit Einführung der Online-Nachrichten unter enormem Druck (Neuberger et al., 2009, S. 9). Mehr Konkurrenz, sinkende Einnahmen und die daraus resultierenden Sparmaßnahmen führen häufig zu Zeit- und Personalknappheit in den Redaktionen. Gleichzeitig drängen neue private und semi-professionelle Anbieter auf den Markt (McNair, 2012, S. 75 f). Mit der Vergrößerung des Medienangebots steigt auch die Wahrscheinlichkeit, mit falschen, unvollständigen oder tendenziösen Nachrichten konfrontiert zu werden. Um selbstbestimmt am politischen Entscheidungsprozess teilnehmen zu können, müssen die Bürger aber richtig und umfassend informiert werden. Die Heterogenität des Medienangebots macht es für demokratische Gesellschaften zunehmend zur Herausforderung, ihre Bürger mit hochwertigen Nachrichten zu versorgen (Herzig & Martin, 2017, S. 127). Da jeder Bürger selbst entscheiden kann, welche Nachrichtenangebote er nutzen möchte, kommt der subjektiven Wahrnehmung bei der Auswahl der Nachrichtenquelle eine große Bedeutung zu. Studien zeigen, dass die Auswahl eines Artikels oder einer Nachrichtensendung maßgeblich davon abhängt, wie die Rezipienten die inhaltliche Qualität der Beiträge bewerten (Wolling, 2002). Die subjektiven Qualitätswahrnehmungen können folglich darüber entscheiden, ob die Bürger gut informiert oder indoktriniert werden. Langezeit wurde darüber diskutiert, ob Nicht-Experten in der Lage sind, adäquate Qualitätsbewertungen vorzunehmen. In den vergangenen Jahren konnte aber mehrfach gezeigt werden, dass die Rezipienten zumindest in begrenztem Maß hochwertige Beiträge von minderwertigen Beiträgen unterscheiden können (Jungnickel, 2011; Urban & Schweiger, 2013). Allerdings können die Einschätzungen anfällig für Störeinflüsse sein. So gibt es Hinweise darauf, dass externe Faktoren, wie die politische Einstellung, das Vertrauen in den Anbieter oder das Design, Auswirkungen auf die Qualitätsbewertungen haben können. In dieser Arbeit soll der Einfluss zweier Störfaktoren untersucht werden. Das Ziel ist es herauszufinden, inwieweit der Verbreitungsweg und die Nachrichtenmarke die Wahrnehmung journalistischer Beiträge beeinflussen. Bei der Untersuchung müssen vorab zwei Einschränkungen gemacht werden. Erstens konzentriert sich diese Arbeit auf Bewertungsunterschiede, die sich aus der Variation externer Faktoren ergeben. Welchen Einfluss der Nachrichteninhalt auf die Bewertung hat, wird nicht untersucht. Zweitens beschränkt sich die Analyse auf Online-Nachrichten. Traditionelle Verbreitungswege wie Zeitungen, Radio und Fernsehen bleiben unberücksichtigt.

Diese Arbeit gliedert sich in einen Theorie- und einen Praxisteil. Im Theorieteil wird zunächst die Frage erörtert, ob die journalistische Qualität eine objektiv messbare Größe ist. Anschließend wird eine Auswahl weitverbreiteter Qualitätsmodelle vorgestellt. Der nachfolgende Abschnitt vermittelt einen Überblick über die aktuellen Entwicklungen der Nachrichtennutzung und beschreibt den Aufstieg der sozialen Netzwerke zum wichtigsten Zugangsweg der jüngeren Generationen. Zum Abschluss des Theorieteils werden aktuelle Studienergebnisse aus dem Bereich der empirischen Qualitätsforschung vorgestellt.

Im Praxisteil wird eine Serie von fünf 2x2 Experimenten geplant und durchgeführt, mit dem der Einfluss der Nachrichtenmarke und des Zugangswegs auf die Qualitätswahrnehmung bestimmt wird. Jedes Telexperiment besteht aus einem Text oder einem Videobeitrag, der von den Versuchsteilnehmern bewertet werden soll. Von jedem Beitrag werden vorab vier Versionen erstellt, die inhaltlich identisch sind, sich aber in Bezug auf die Medienmarke und den Verbreitungsweg unterscheiden. Ein Zufallsgenerator entscheidet darüber, welche Versionen eines Beitrags ein Teilnehmer zu sehen bekommt. Das Experiment wird online mit Hilfe einer Umfragesoftware durchgeführt und die Daten elektronisch erfasst. Die Antworten der Teilnehmer werden anschließend mit statistischen Methoden wie dem t-Test und dem Mann-Whitney-U-Test auf Abweichungen untersucht. Das Ziel des Versuchsaufbaus ist es, valide, objektive und reliabel Ergebnisse zu bekommen und damit zum Verständnis beizutragen, wie externe Faktoren die Qualitätswahrnehmung von Nachrichten beeinflussen. Wichtige Quellen dieser Arbeit sind die Metanalyse „Defining and measuring quality journalism“ von Stephen Lacy und Tom Rosenstiel und die experimentelle Studie „News quality from the recipients`perspective“ von Urban und Schweiger.

2 Was ist journalistische Qualität?

2.1 Überall Qualitäten

Der Qualitätsbegriff nimmt in der Journalismusforschung einen zentralen Platz ein. Ob eine Medienmarke mit hoher oder niedriger Qualität assoziiert wird, bestimmt über die öffentliche Wahrnehmung, den Einfluss und den wirtschaftlichen Erfolg eines Medienhauses. Auch Wissenschaftler unterscheiden häufig zwischen Boulevard- und Qualitätsmedien. In Deutschland werden meist überregionale Zeitungen wie die FAZ und die Süddeutsche Zeitung und Nachrichtenmagazine wie der Spiegel und die Zeit der „Qualitätspresse“ zugeordnet (vgl. Meyen & Jandura, 2011, S. 211 f). Auch bei der Bewertung journalistischer Beiträge ist der Qualitätsbegriff allgegenwärtig. So gibt es im deutschsprachigen Raum aktuell 550 Journalistenpreise, die es sich zum Ziel gesetzt haben, besonders hochwertige Beiträge auszuzeichnen (Oberauer, 2020). So weitverbreitet das Wort Qualität ist, so wenig definiert ist hingegen seine Bedeutung. Im allgemeinen Sprachgebrauch ist der Qualitätsbegriff meist positiv konnotiert. So wirbt eine Vermarktungsgesellschaft für Lebensmittel mit „Qualitätskäse aus der Schweiz“ und Fußballtrainer, fordern regelmäßig mehr „Qualitätsspieler“ für ihre Mannschaft. In seiner ursprünglichen Bedeutung ist der Begriff allerdings wertfrei. Qualität kommt vom lateinische Wort *qualitas* und bedeutet so viel wie die Beschaffenheit oder „das Wesen“ einer Sache (Hermes, 2006, S. 17). Die Qualität ist ein neutraler Sammelbegriff für die Summe der Eigenschaften eines Objekts. Auf die Frage, welche Qualität ein Objekt hat, gibt es keine allgemeingültige Antwort. Ein Qualitätsurteil ist nur dann sinnvoll, wenn es sich auf einen konkreten Anwendungsfall bezieht (vgl. Nohr, 2001). Da die Menge der möglichen Anwendungen und Bezugsrahmen aber unendlich groß ist, gilt Gleiches auch für die Menge der möglichen Qualitätsurteile. Die Abhängigkeit vom Kontext gilt für immaterielle Güter wie journalistische Beiträge ebenso wie für physische Objekte. So hat beispielsweise ein Metallring mit einem Durchmesser von 15 Millimetern, keine Qualität, die dem „Ring an sich“ innewohnt. Wurde der Ring als Maschinenteil für einen Dieselmotor konstruiert, lassen sich aus diesem Anwendungsfall Eigenschaften wie Härte, Gewicht oder eine glatte Oberfläche ableiten, über die der Ring verfügen sollte. Das Vorhandensein oder die Abwesenheit dieser Eigenschaften bilden dann eine objektive Grundlage für ein Qualitätsurteil. Wird der Ring jedoch als Schmuckstück getragen, ändert sich auch der Bewertungsmaßstab. Eigenschaften wie die Farbe oder der Glanz

erhalten bei der Bewertung eine verstärkte Bedeutung, während die Härte oder das Gewicht des Rings in den Hintergrund treten. Auch können Eigenschaften wie die einfache Formbarkeit des Materials in einem Fall von Vorteil sein (Anpassung des Schmuckstücks an die Hand), während sie sich in einem anderen Fall (Einsatz im Verbrennungsmotor) negativ auswirken. Je nachdem wie der Anwendungsbereich eines Objekts definiert wird, ergeben sich unterschiedliche Bewertungskriterien (Arnold, 2008, S. 491). Auch bei Texten, Videos und Radiobeiträgen ist die Qualität vom Kontext abhängig. Viele Qualitätsstudien beginnen damit, die Funktion des Journalismus in Staat und Gesellschaft zu definieren. Ausgehend von dieser Definition werden Kriterien abgeleitet, denen ein qualitativ hochwertiges Medienangebot oder ein hochwertiger Beitrag entsprechen sollten, um seine staatstheoretische Aufgabe zu erfüllen. Unterschiedliche Staatsformen und politische Systeme führen zwangsläufig zu unterschiedlichen Auffassungen über die Aufgabe von Medien und Journalisten und damit zu abweichenden Vorstellungen darüber, was einen hochwertigen Beitrag ausmacht. So wurden etwa in autokratisch regierten Staaten wie der DDR von der Presse regierungsfreundliche Beiträge erwartet, während die Medien in liberalen Demokratien gerne als „vierte Gewalt im Staat“ bezeichnet werden (Hermes, 2006, S. 17) und die Regierung durch eine kritische und analytische Berichterstattung kontrollieren sollen (Prinzing, 2016, S. 106). Die Gütekriterien, die in dieser Arbeit berücksichtigt werden, beziehen sich allerdings ausnahmslos auf die Arbeit von Journalisten in liberalen Demokratien. In der Theorie ist die Zahl der möglichen Bewertungsmaßstäbe und der damit einhergehenden Qualitätsurteile unbegrenzt. In der Praxis gibt es aber meist Konventionen darüber, welche Eigenschaften ein hochwertiges von einem minderwertigen Objekt unterscheidet. Die Qualität ist ein normativer Begriff, d.h. es gibt keine Güter, denen eine hohe Qualität inhärent ist (Vlasic, 2004, S. 15). Die Kennzeichnung „hohe, mittlere oder niedrige Qualität“ erhalten Objekte stattdessen durch die Zuschreibung von außen. Die Aufgabe Bewertungskriterien zu definieren, wurde im Bereich des Journalismus von Politik- und Kommunikationswissenschaftlern, sowie Soziologen übernommen. Ob Kriterien, die von Experten festgelegt wurden, eine höhere Relevanz haben als diejenigen, die aus Publikumsbefragungen ermittelt wurden, ist in der Forschung umstritten.

2.2 Expertise ist (nicht) alles

Neben dem Anwendungskontext hat noch ein zweiter Faktor maßgeblichen Einfluss auf die Qualitätsbewertung. Jeder Mensch hat individuelle Vorlieben, Abneigungen und Bedürfnisse, die in dieser speziellen Zusammensetzung einzigartig sind. In der Folge unterscheiden sich auch die Bewertungsmaßstäbe, die einem Qualitätsurteil zu Grunde liegen, von Individuum zu Individuum. So können selbst Menschen, die im selben Land leben, ähnliche politische Ansichten haben und die gleichen Medien konsumieren, diametral unterschiedliche Auffassungen darüber haben, was einen gelungenen journalistischen Beitrag ausmacht. Die individuellen Lebensumstände spielen dabei ebenso eine Rolle, wie genetische Veranlagungen und soziökonomische Faktoren. Auch das persönliche Verhältnis zum Objekt der Bewertung kann ein bedeutender Einflussfaktor sein. So haben Journalisten und Verleger meist eine andere Qualitätsauffassung als die Leser, die ihre Beiträge konsumieren. Beispielsweise könnten Journalisten einen Zeitungsartikel eher danach bewerten, wie viel Zeit in die Recherche investiert wurde, während die Rezipienten vor allem darauf achten, wie gerne sie den Beitrag gelesen haben (Urban & Schweiger, 2013, S. 2). In der Regel werden Qualitätsbewertungen von Experten vorgenommen, die entweder an der Produktion des Objekts beteiligt sind (1), im Alltag regelmäßig damit konfrontiert werden (2) oder in einem relevanten Fachbereich über langjährige Erfahrungen verfügen (3). Im Fall des Metallrings wären dies zum Beispiel Industriemechaniker und Ingenieure. Die Tatsache, dass sich Qualitätsstudien bevorzugt auf die Meinung von Experten beziehen, ist einfach zu erklären: Menschen, die in engem Kontakt zu einem relevanten Anwendungsbereich stehen, haben in der Regel auch ein gesteigertes praktisches und theoretisches Wissen über das betreffende Objekt. Die Qualitätsurteile dieser Menschen haben eine höhere praktische Relevanz, weil sie eng mit dem beabsichtigten Einsatzzweck verknüpft sind. Natürlich könnten auch Krankenpfleger oder Grundschullehrer, die in ihrem Alltag nicht mit der Produktion von Maschinenteilen beschäftigt sind, eine Qualitätsbewertung über den Metallring abgeben. Allerdings ist davon auszugehen, dass diese Urteile nur in begrenztem Maß mit dem relevanten Anwendungsfall korrelieren. Bei fehlender Expertise können Qualitätsurteile einfacher von externen Einflussfaktoren beeinflusst werden. Für Objekte wie den Metallring, die für einen engen Anwendungsbereich ausgelegt sind, werden Untersuchungen zur Qualität daher meist auf kleine Expertengruppen zugeschnitten. Diese Beschränkungen sind aber nicht für alle Objekte sinnvoll.

Insbesondere journalistische Erzeugnisse bilden in dieser Hinsicht eine wichtige Ausnahme. Journalistische Beiträge werden in der Regel von einer sehr großen Personenzahl konsumiert. Nachrichten richten sich gerade nicht an einen klar definierten Personenkreis, sondern wollen eine möglichst große Zahl von Menschen erreichen. Auch wenn sich der Nachrichtenmarkt in den vergangenen Jahren stark ausdifferenziert hat und Wissenschaftler seit Jahren über Echokammern und Filterblasen diskutieren (vgl. Flaxman et al., 2016; Geschke et al., 2019), haben Nachrichtenbeiträge nach wie vor den Anspruch eine Gesamtöffentlichkeit herzustellen. In einer parlamentarischen Demokratie ist es eine wichtige Aufgabe der Presse die Staatsbürger über aktuelle Ereignisse und Entwicklungen zu informieren und den Meinungsaustausch politischer Strömungen öffentlich abzubilden (Fengler, 2016, S. 237). In jeder Demokratie kommt der Presse die Aufgabe zu, die Menschen über relevante Themen aufzuklären und dadurch zur Teilhabe am kollektiven Entscheidungsprozess zu befähigen (Uwe Hasebrink, 2011, S. 324). Kurz zusammengefasst lässt sich festhalten: Alle Staatsbürger sind Zielgruppe und Rezipienten journalistischer Beiträge. Untersuchungen über die Qualität des Nachrichtenangebots, die ausschließlich die Ansichten von Medienexperten berücksichtigen, können daher nur einen Ausschnitt der Wirklichkeit abbilden. Um ein vollständiges Bild des Einflusses journalistischer Beiträge auf das politische System zu bekommen, muss auch die Perspektive des Publikums miteinbezogen werden. Bei genauerer Betrachtung steht diese Erkenntnis auch nicht im Widerspruch zur Beobachtung, dass Experten in der Regel relevantere Qualitätsurteile abgeben. Da nahezu alle Menschen im Alltag mit Nachrichten konfrontiert sind, fallen sie unter die zweite obengenannte Expertengruppe („Menschen, die im Alltag regelmäßig mit dem Objekt konfrontiert werden“). Laut den aktuellen Zahlen des Digital News Reports konsumieren beispielsweise 94 Prozent der deutschen Onliner mehrmals wöchentlich Nachrichten in der ein oder anderen Form (Hölig & Hasebrink, 2020, S. 13). Zusätzlich sind alle volljährigen und mündigen Staatsbürger qua Wahlrecht Experten für das politische System, da sie als Kollektiv die politische Entscheidungsgewalt innehaben. Die Gruppe der Nicht-Experten ist in Bezug auf journalistische Beiträge folglich sehr klein oder möglicherweise gar nicht existent. Neben der Einsicht, dass das Publikum eine andere Sichtweise auf die Qualität der Nachrichtenmedien eröffnet, bietet dessen Einbezug noch einen weiteren Vorteil. Der Ansatz, die Meinung aller Bürger zu berücksichtigen folgt der Idee, dass Menschen, die sich zu einem Kollektiv zusammenschließen in der Lage sind, eine genauere Bewertung abzugeben als einzelne Experten. Die Voraussetzung dafür ist, dass alle Teilnehmer über

geringes Grundwissen über das Objekt verfügen. Wie oben gezeigt, ist diese Bedingung für journalistische Beiträge erfüllt. In einer viel zitierten Studie aus dem Jahr 1907 konnte Galton zeigen, dass eine Gruppe von Menschen in der Lage ist, das Gewicht eines Ochsen genauer einzuschätzen als ein Experte (Galton, 1907, S. 451). In der Auswertung des Experiments zeigte sich, dass diejenigen die das Gewicht überschätzten und diejenigen, die das Gewicht unterschätzten sich im Endresultat ausglich. Die kollektive Qualitätsbewertung fußt damit auf denselben Grundpfeilern, wie die Marktwirtschaft und das politische System liberaler Demokratien, die die Summe der Marktteilnehmer bzw. der Wähler für eine bessere Entscheidungsfindung und Allokation von Gütern nutzen.

2.3 Qualität messen – geht das überhaupt?

Der Begriff Qualitätsmessung ist irreführend, impliziert doch das Wort Messung, dass Qualität eine objektive Größe sei. Wie oben ausgeführt ist die Qualität aber normativ und wird durch individuelle Wertvorstellungen und den gesellschaftlichen Diskurs beeinflusst (Beck et al., 2010, S. 16). Der deutsche Kommunikationswissenschaftler Ruß-Mohl formulierte es im Jahr 1992 so:

"Qualität im Journalismus definieren zu wollen, gleicht dem Versuch, einen Pudding an die Wand zu nageln" (Ruß-Mohl, 1992, S. 85).

Im Gegensatz zu physikalischen Größen wie Masse, Spannung oder Stromstärke ist eine exakte Messung von Qualität nicht möglich. Um nicht von einer Messung zu sprechen, wird in der Literatur daher häufig der Begriff Qualitätsbestimmung verwendet (vgl. Bilke, 2008, S. 25). Der Begriff Bestimmung beinhaltet die gleiche Untersuchungs- und Definitionsabsicht, ohne dabei einen Anspruch auf Exaktheit oder Allgemeingültigkeit zu erheben. Um ein ganzheitliches Qualitätsmodell zu entwickeln, müssen alle Faktoren berücksichtigt werden, die sich auf die Zusammensetzung des Medienangebots auswirken können. Shoemaker und Reese präsentierten 1996 ein „Hierarchy of influences“-Modell, das die Einflussfaktoren nach ihrem Rang in der journalistischen Wertschöpfungskette ordnet (Shoemaker & Reese, 1996, S. 251 ff). Das Modell berücksichtigt dabei die Presse- und Rundfunkgesetze, soziokulturelle Einflüsse und organisations- und redaktionsinterne Entscheidungshierarchien. Auch auf der Ebene des einzelnen Reporters müssen unterschiedliche Faktoren, wie das Medium, das Ressort, die Zielgruppe und der

Veröffentlichungsrhythmus berücksichtigt werden. Vor Beginn einer Qualitätsstudie muss zunächst die Entscheidung getroffen werden, welches Objekt vor welchem Hintergrund untersucht werden soll. Diese Feststellung erscheint zunächst trivial. In der Praxis sind aber meist einige Vorüberlegungen nötig. Dies gilt insbesondere für abstrakte Objekte wie „den Journalismus“ oder „das Mediensystem“. Möchte man die Qualität der journalistischen Berichterstattung in Deutschland bestimmen, ist beispielsweise die Frage der Perspektive hoch relevant. Viele Studien, die sich mit diesem Thema auseinandersetzen, wählen eine Makroperspektive (vgl. Leidenberger, 2015). Diese Untersuchungen beschäftigen sich mit dem Mediensystem als Ganzes. Im Digital News Report der Universität Oxford werden die Teilnehmer beispielsweise gefragt, wie sehr sie dem Medienangebot insgesamt vertrauen (vgl. Hölig & Hasebrink, 2020, S. 30). Studien, die eine Meso-Perspektive einnehmen, konzentrieren sich dagegen auf die Berichterstattung einzelner Medienmarken und Verlagshäuser (vgl. Eisenegger et al., 2018). Auf der Mikroebene befasst sich die Forschung mit Einzelbeiträgen. Auch dabei kommen häufig qualitative Methoden zum Einsatz. Die Wahl der Perspektive beeinflusst die Kategoriensysteme, die den Bewertungen zu Grunde liegen. So kann etwa die Unterhaltsamkeit auf der Mikroebene als relevante Qualitätseigenschaft definiert werden, während auf der Makroebene mehr Wert auf die Meinungspluralität gelegt wird. Da sich das Nachrichtenangebot in den vergangenen Jahren stark ausdifferenziert hat, wird eine Gesamtbetrachtung zunehmend schwieriger. Während bis zum Beginn der 90er Jahre Zeitungen und öffentlich-rechtliche Rundfunksender den Nachrichtenmarkt in Deutschland dominierten, gleicht das Mediensystem 30 Jahre später einem Mosaik aus Einzelbausteinen, die durch das Internet in einem losen Verband zusammengehalten werden. Zusammen mit den Abonnenten- und Zuschauerzahlen der Zeitungen und Rundfunksender ist auch die Markentreue des Publikums gesunken. Im Jahr 2020 bauen viele Menschen ihren Nachrichtenmix eklektisch aus dem stetig wachsenden Informationspool zusammen. Nachrichtenaggregatoren wie Flipboard, Google- und Apple-News verstärken diesen Trend (Lee & Chyi, 2015). Vor diesem Hintergrund gewinnen Qualitätsuntersuchungen, die sich auf Einzelbeiträge beziehen, an Bedeutung. Studien auf der Makroebene sind dagegen sehr ressourcenintensiv und müssen sich mit dem Problem auseinandersetzen, dass Einteilungen des Markts in Mediengattungen und Anbieter immer weniger sinnvoll erscheinen.

2.4 Journalistische Qualität: eine Säule der Demokratie

Die journalistische Qualitätsforschung hat drei wesentlichen Anwendungsbereiche:

1) Wirtschaft

Um auf dem Medienmarkt konkurrenzfähig zu sein, sind Verleger und Medienunternehmen daran interessiert die Qualität ihrer Produkte zu bestimmen, da Unternehmen, die eine höhere Qualität als ihre Mitbewerber bieten über einen Wettbewerbsvorteil verfügen (Lacy & Rosenstiel, 2015, S. 9).

2) Kommunikationswissenschaft

Ein wesentlicher Anteil der Kommunikation moderner Gesellschaften findet über die traditionellen bzw. die sozialen Medien statt. Um Trends und Veränderungen adäquat beschreiben zu können, bedarf es etablierter Methoden, mit deren Hilfe der Umfang und die Qualität des Medienangebots bestimmt werden kann.

3) Demokratietheorie

In einer Demokratie hat die Presse einen zentralen Platz im politischen System. Ihre Aufgaben bestehen unter anderem darin, die Bevölkerung schnell und richtig zu informieren und die Mächtigen zu kontrollieren. Die journalistische Arbeit wird vom deutschen Grundgesetz daher unter besonderen Schutz gestellt. In Artikel 5 heißt es: *„Jeder hat das Recht, seine Meinung in Wort, Schrift und Bild frei zu äußern und zu verbreiten und sich aus allgemein zugänglichen Quellen ungehindert zu unterrichten“* (Grundgesetz, 1949).

Dem besonderen Schutz der Pressefreiheit liegt die Annahme zu Grunde, dass Staatsbürger durch richtige und umfassende Informationen in die Lage versetzt werden, bessere politische Entscheidungen zu treffen. Fehlinformationen führen im Umkehrschluss dazu, dass Bürger Entscheidungen treffen, die von ihren eigentlichen Präferenzen abweichen. So vertreten einige Wissenschaftler die Ansicht, dass Falschmeldungen und die Abwesenheit konträrer Meinungen in den sozialen Netzwerken entscheidend zur Wahl Donald Trumps zum US-Präsidenten beigetragen haben (Guo et al., 2020). Folglich können Änderungen des Medienangebots bedeutende politische Veränderungen zur Folge haben und die politische Stabilität eines Landes beeinflussen. Um zu verstehen wie sich der Journalismus auf die öffentliche Meinung auswirken kann,

muss man sich mit den Mediennutzungsgewohnheiten der Bevölkerung auseinandersetzen. Zwischen dem Inhalt journalistischer Beiträge und der öffentlichen Meinung besteht kein linearer Zusammenhang. Stattdessen ist der Wirkmechanismus nichtlinear und multivariabel. So beeinflussen neben dem Inhalt und der Aufbereitung der Nachrichten auch die Nutzungshäufigkeit, die verwendeten Empfangsgeräte oder die Zugangswege (soziale Medien, Radio, Zeitung, ...) das Vertrauen der Bevölkerung in Politik und Medien. Das nächste Kapitel soll daher einen Überblick über Mediennutzungsgewohnheiten in Deutschland bieten.

3 Nachrichtennutzung in Deutschland

3.1 Internetnutzung in Deutschland

Laut dem aktuellen D21 Digitalindex nutzen 86 Prozent der deutschen Bevölkerung das Internet (Prinzing, 2016, S. 106). Die Onlinestudie von ARD und ZDF ermittelt mit 90 Prozent sogar einen noch höheren Wert (Beisch et al., 2019, S. 374). Die Nutzungswahrscheinlichkeit sinkt dabei mit dem Alter und fällt bei Männern etwas höher aus als bei Frauen. Deutliche Unterschiede zeigen sich beim Vergleich der Altersgruppen. Während die Generation der unter 50-Jährigen nahezu vollständig online aktiv ist, sind es bei den über 70-Jährigen nur 52 Prozent. Menschen mit niedriger formaler Bildung und Nicht-Berufstätige sind deutlich seltener online. Der Stadt-Land Vergleich zeigt: Je urbaner die Umgebung und je mehr Menschen in einem Haushalt leben, desto eher wird das Internet genutzt (Schwaderer, 2020, S. 15). Die Zahl derer, die täglich mindestens einmal das Internet nutzen, ist in den vergangenen Jahren deutlich angestiegen und liegt mittlerweile bei etwa 50,3 Millionen Menschen (Beisch et al., 2019, S. 376). Hohe Zuwachsraten gibt es auch bei der Nutzung von Online-Videos und sozialen Medien. Der Kommunikationsdienst WhatsApp liegt mit einer Tagesreichweite von 64 Prozent auf dem ersten Platz. Facebook und Instagram liegen mit 21 und 13 Prozent dahinter auf den Plätzen zwei und drei. Instagram ist vor allem bei den 14- bis 29-Jährigen sehr beliebt und wird in dieser Altersgruppe täglich von 47 Prozent der Menschen genutzt (Beisch et al., 2019, S. 383). Mehr als drei Viertel der Deutschen besitzen ein Smartphone (Schwaderer, 2020, S. 20). Damit ist es das meistverbreitete Endgerät zur Nutzung digitaler Medien. Seine Verbreitung hat sich seit 2013 nahezu verdoppelt. Die Verbreitung von Laptops, Desktop-PCs und Tablets stagniert dagegen auf einem hohen Niveau. Obwohl sprachgesteuerte Assistenten, smarte Fitnessarmbänder und VR-Brillen medial viel Aufmerksamkeit bekommen, werden sie nur von einem kleinen Teil der Bevölkerung zu beruflichen oder privaten Zwecken genutzt (Schwaderer, 2020, S. 20).

3.2 Nachrichteninteresse und Nutzungshäufigkeit

Da in dieser Arbeit ausschließlich Online-Nachrichten untersucht werden, beziehen sich auch die folgenden Abschnitte auf Studien, die die digitale Nachrichtennutzung in den Mittelpunkt stellen. Eine der größten und umfassendsten Nutzungsstudien, ist der Digital

News Report, der jährlich vom „Reuters Institute of Journalism“ der Universität Oxford durchgeführt wird. Der Report gibt umfassende Einblicke in die Verhaltensmuster und die Einstellungen der Nutzer von Online-Nachrichten. Ein Schwerpunkt der diesjährigen Ausgabe liegt auf der Nachrichtennutzung über die sozialen Medien. Da 14 Prozent der Deutschen nicht im Internet aktiv sind, ist der Report allerdings nicht repräsentativ für die Gesamtbevölkerung.

Die große Mehrheit der deutschen Onliner interessiert sich für Nachrichten und aktuelle Ereignisse. 71 Prozent der Befragten des Digital News Reports gaben an, sich sehr oder sogar überaus stark für Nachrichten zu interessieren (Schwaderer, 2020, S. 8). Das Interesse steigt dabei mit dem Alter und ist in der Altersgruppe der über 55-Jährigen am höchsten. Auch die Zahl der Menschen, die mehrmals pro Woche Nachrichten konsumieren, steigt mit dem Alter an. Sie liegt aber in allen Altersgruppen seit 2013 auf einem konstant hohen Niveau. Insgesamt betrachtet nutzen 94 Prozent der Onliner mehrmals in der Woche nachrichtliche Angebote (Hölig & Hasebrink, 2020, S. 13). Bei jungen Erwachsenen ist das Interesse an Nachrichten im Vergleich zum Vorjahr deutlich gestiegen. In der Altersgruppe der 18- bis 24-Jährigen stieg das Interesse um sieben Prozentpunkt auf 50 Prozent und in der Gruppe der 25- bis 34-Jährigen sogar um neun Prozentpunkte auf 66 Prozent. Traditionelle Medien wie Zeitungen, Radio und Fernsehen haben in den vergangenen Jahren bei der Nachrichtennutzung an Bedeutung verloren. Allerdings vollzieht sich der Wandel zur Online-Nachrichtennutzung in Deutschland langsamer als in anderen Ländern. Newman et al. ordnen Deutschland den „Traditionalisten“ zu, also denjenigen Ländern, in denen Tageszeitungen und Fernsehnachrichten nach wie vor eine große Rolle spielen (Newman et al., 2016, S. 86 f). So ist das Fernsehen auch für die meisten deutschen Onliner (42 Prozent) die wichtigste Nachrichtenquelle. Zusätzlich gaben 70 Prozent der Befragten an, mindestens einmal pro Woche TV- Nachrichten zu konsumieren (Hölig & Hasebrink, 2020, S. 18). Die zweitwichtigste Nachrichtenquelle ist das Internet (38 Prozent), gefolgt von Radio- und Printnachrichten. Bei der Nutzungshäufigkeit liegt das Internet im Jahr 2020 erstmals gleichauf mit dem Fernsehen. Die verhältnismäßig hohe Bedeutung traditioneller Medien lässt sich auf die Altersstruktur der Bevölkerung zurückzuführen. In der Generation der unter 45-Jährigen hat das Internet das Fernsehen als meistgenutzte Nachrichtenquelle abgelöst. Aufgrund des hohen Altersdurchschnitts der deutschen Bevölkerung, prägen die Nutzungsgewohnheiten älterer Menschen aber noch immer die Gesamtstatistik.

3.3 Führen alle Wege zur Nachricht?

Die zunehmende Verbreitung von Smartphones und der damit einhergehende Wandel des Mediennutzungsverhaltens, hat auch die Nachrichtennutzung fundamental verändert (Struckmann & Karnowski, 2016). Traditionelle Anbieter wie Zeitungen und Rundfunksender haben ihre herausragende Stellung auf dem Medienmarkt verloren und kämpfen mit privaten und semiprofessionellen Anbietern um Aufmerksamkeit (Peters & Broersma, 2013, S. 6). Das Wissen darum, wie Nutzer zu ihren Inhalten gelangen, ist für Journalisten und Verlage von großer Bedeutung. Die Zugangswege zu Online-Nachrichten haben sich stark ausdifferenziert. Kommunikation und Nachrichtenkonsum finden heute zum Großteil über Intermediäre wie Suchmaschinen oder soziale Netzwerke statt. Auch Newsreader Apps wie Google News, Apple News oder Flipboard gewinnen an Bedeutung (Lee & Chyi, 2015). Der Digital News Report unterscheidet aktuell sieben Zugangswege zu Online-Nachrichten (Newman, 2019, S. 13 f). Die Zugangswege lassen sich anhand der Fragen, wie hoch das Informationsbedürfnis der Nutzer ist und ob sie hauptsächlich von Algorithmen geprägt werden, in vier Kategorien aufteilen:

	algorithmenbasiert	nicht-algorithmenbasiert
Hohes Informationsbedürfnis	Newsreader-Apps Suche nach einem Thema	Direktzugriff Suche nach einer Webseite
Geringes Informationsbedürfnis	Soziale Medien	Email-Newsletter Push-Nachrichten

Tabelle 1: Kategorisierung der Online-Zugangswege

Ausgehend von dieser Einteilung können Rückschlüsse auf Absichten und Bedürfnisse der Nutzer gezogen werden, die der Wahl eines bestimmten Zugangswegs zugrunde liegen. Die Nutzung eines nicht-algorithmenbasierten Zugangswegs deutet darauf hin, dass ein hohes Interesse daran besteht, die Nachrichtenquelle selbst zu wählen. Jeder Direktzugriff auf eine Webseite oder eine App erfordert eine aktive Entscheidung des Nutzers für einen bestimmten Anbieter. Ähnliches gilt für das Abonnement eines Newsletters oder einer Push-Benachrichtigung, wenn auch in abgeschwächter Form. Der Direktzugriff auf eine Webseite ist somit das digitale Äquivalent eines Zeitungskaufs am Kiosk. Die Anmeldung zu einem Newsletter ist mit dem Abonnement einer Tageszeitung

vergleichbar. Bei algorithmenbasierten Zugangswegen entscheidet hingegen ein Computerprogramm darüber, welche Beiträge die Nutzer zu sehen bekommen. Diese Zugangswege erfordern ein geringeres Maß an Selbständigkeit. Während die Nutzung von Suchmaschinen und Newsreader-Apps auf ein hohes Informationsbedürfnis schließen lässt, werden die sozialen Medien häufig auch zu Kommunikations- und Unterhaltungszwecken genutzt. Die Nachrichtennutzung in den sozialen Netzwerken setzt dagegen kein akutes Informationsinteresse voraus. In den sozialen Medien sind die Nachrichten nur ein Teil eines bunten Beitragsstroms, dessen Zusammensetzung sich mehrmals pro Minute ändern kann.

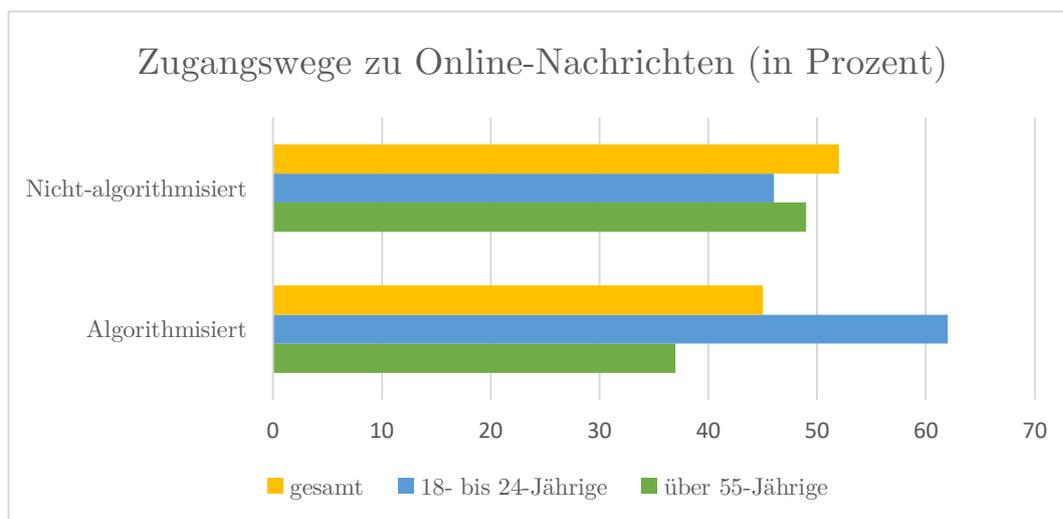


Abbildung 1: Algorithmenbasierte und Nicht-algorithmenbasierte Zugangswege (Hölig & Hasebrink, S.45)

Für den größten Teil der deutschen Onliner (34 Prozent) ist der Direktzugriff der wichtigste Zugangsweg zu nachrichtlichen Inhalten (Hölig & Hasebrink, 2020, S. 45). Berücksichtigt man auch die Möglichkeit, über eine Suchmaschine nach einer bestimmten Webseite zu suchen erhöht sich der Anteil auf 46 Prozent. Die sozialen Medien liegen in der Rangliste der beliebtesten Zugangswege mit 14 Prozent auf dem zweiten Platz. Beim Vergleich der Altersgruppen zeigen sich aber erneut deutliche Unterschiede. In der Gruppe der 18- bis 24-Jährigen stehen die sozialen Medien auf Platz eins, während sie für die über 55-Jährigen nur der fünftwichtigste Zugangsweg sind. Ein ähnliches Bild zeigt sich, wenn alle regelmäßig genutzten Zugangswege abgefragt werden. Auch in dieser Kategorie führt der Direktzugriff die Gesamtbetrachtung an. In der jüngsten Altersgruppe sind es hingegen die sozialen Medien. Insgesamt spielen algorithmenbasierte Zugangswege bei den 18- bis 24-Jährigen eine deutlich größere Rolle als in allen anderen Altersgruppen (Hölig & Hasebrink, 2020, S. 43).

3.4 Vertrauen in Nachrichten

Das Vertrauen in die mediale Berichterstattung war in Deutschland über viele Jahre hinweg stabil (Reinemann et al., 2017). In den vergangenen Jahren ist es allerdings messbar gesunken. Während im Jahr 2015 noch 60 Prozent der Onliner erklärten, den Nachrichten im Allgemeinen zu vertrauen, waren es im Jahr 2020 nur noch 45 Prozent (Hölig & Hasebrink, 2020, S. 30 f). Auch der Anteil derjenigen, die den Medien ein offenes Misstrauen aussprechen ist im selben Zeitraum von 15 auf 23 Prozent gestiegen. Wenn die Menschen nur zu den Medien befragt werden, die sie selbst nutzen, ist der Vertrauenswert mit 59 Prozent etwas höher. Allerdings ist auch dieser Wert seit 2015 um 9 Prozentpunkte zurückgegangen. Trotz des negativen Trends der vergangenen Jahre befindet sich Deutschland im internationalen Vergleich noch im gehobenen Mittelfeld. Insgesamt betrachtet ist das Vertrauen in fast allen Ländern rückläufig. Der Verdacht liegt nahe, dass die sinkenden Werte mit den veränderten Nutzungsgewohnheiten in Zusammenhang stehen könnten. In allen Altersgruppen herrscht gegenüber algorithmenbasierten Zugangswegen ein großes Misstrauen. 50 Prozent der Befragten misstrauen Nachrichten in den sozialen Netzwerken (Hölig & Hasebrink, 2020, S. 33). Das Misstrauen in Nachrichten, die über die Suchmaschinen abgerufen werden, fällt etwas geringer aus, ist mit 29 Prozent aber noch verhältnismäßig hoch. Berücksichtigt man die Nutzungshäufigkeiten, so ist es wenig erstaunlich, dass das allgemeine Vertrauen der jüngeren Altersgruppen besonders gering ausfällt. Bei den 18- bis 24-Jährigen stimmten im Jahr 2020 nur noch 31 Prozent der Aussage zu, dass sie den Nachrichten im Allgemeinen vertrauen. Das entspricht einen Rückgang von 9 Prozentpunkten im Vergleich zum Vorjahr (Hölig & Hasebrink, 2020, S. 29). Auch die Zahl der Menschen, die aktiv nachrichtliche Inhalte vermeiden, ist seit der erstmaligen Erhebung im Jahr 2017 leicht angestiegen. Mehr als die Hälfte der Befragten erklärte, häufig oder zumindest vereinzelt Nachrichten aus dem Weg zu gehen. Ob die veränderten Zugangswege aber wirklich zu einem Vertrauensverlust beigetragen haben, lässt sich anhand der Vielzahl möglicher Einflussfaktoren nur schwer abschätzen. Eventuell könnten auch politische Großereignisse wie der Syrienkrieg, der Brexit oder die Wahl Donald Trumps zum US-Präsidenten für das gewachsene Misstrauen verantwortlich sein.

4 Empirische Qualitätsforschung

In der Wissenschaft wurde die Qualität sowohl von der Angebots- als auch von der Nachfrageseite untersucht (Lacy & Rosenstiel, 2015, S. 11). Angebotsorientierte Studien sind stark mit der Theorie verknüpft und zeichnen sich durch ein deduktives Vorgehen aus. Ausgehend von der Theorie wird zunächst die Aufgabe des Journalismus in Staat und Gesellschaft definiert. Ausgehend von dieser Definition werden anschließend Eigenschaften abgeleitet, über die ein hochwertiger Beitrag verfügen sollte. Angebotsorientierte Studien beschränken sich dabei meist auf das Qualitätsverständnis von Journalisten, Wissenschaftlern und anderen Medienexperten. Im Gegensatz dazu befassen sich nachfrageorientierte Studien mit den Ansichten der Rezipienten. Vielen nachfrageorientierten Studien verfolgen einen empirischen Forschungsansatz und beginnen ihre Untersuchungen, ohne im Vorhinein feste Qualitätskriterien festgelegt zu haben (Lacy & Rosenstiel, 2015, S. 11 f). Stattdessen generieren sie ihr Qualitätsverständnis aus Befragungen der Rezipienten. Ziel der nachfrageorientierten Forschung ist es, die Bedürfnisse der Nutzer in den Mittelpunkt zu stellen. Unabhängig vom Forschungsansatz herrscht in der Wissenschaft weitgehend Einigkeit darüber, dass die Qualität keine binäre Eigenschaft ist. Stattdessen werden Beiträge meist in ein multivariablen Qualitätsspektrum eingeordnet (Spurk & Lublinski, o. J., S. 2).

4.1 Angebotsmodelle

Die Angebotsmodelle beruhen auf der Annahme, dass bestimmte Qualitätsmerkmale von den Bewertungen der Rezipienten unabhängig sind. Das bedeutet nicht, dass die Wünsche der Rezipienten keine Rolle spielen, sondern lediglich, dass die Qualität eines Beitrags auch ohne Befragung des Publikums ermittelt werden kann. In angebotsorientierten Studien wird die journalistische Qualität von Kommunikationswissenschaftlern oder anderen Medienexperten bestimmt. Dazu muss zunächst die Funktion definiert werden, die der Journalismus in einer demokratischen Gesellschaft erfüllen soll (vgl. McQuail, 1992). Die wesentliche Aufgabe des Journalismus ist es, Informationsungleichgewichte in der Gesellschaft zu kompensieren (Urban & Schweiger, 2013, S. 2). Die Nachrichten sollen zu einem besseren Verständnis von politischen Fragen und Problemen beitragen (Uwe Hasebrink, 2011, S. 324) und dafür

sorgen, dass eine Medienöffentlichkeit entsteht, in der ein Meinungsaustausch zwischen unterschiedlichen gesellschaftlichen Strömungen stattfinden kann (Habermas, 2006, S. 412ff). Außerdem sollen Journalisten als „vierte Gewalt im Staat“ durch eine aufmerksame Berichterstattung die politischen und wirtschaftlichen Entscheidungsträger kontrollieren (Urban & Schweiger, 2013, S. 3). Eine qualitativ hochwertige Berichterstattung versucht diesen Anforderungen gerecht zu werden.

Im Jahr 2008 versuchte Arnold die unterschiedlichen theoretischen Herleitungen zu systematisieren und in ein Gesamtkonzept einzubinden. In seiner Arbeit konnte er zeigen, dass sich das angebotsorientierte Qualitätsverständnis auf drei Säulen stützt (Arnold, 2008):

- (1) Funktional-systemorientierte Ebene (FS-Ebene): Grundlage der FS-Ebene sind die Vorgaben der journalistischen Leitcodes, die historisch gewachsen sind, und sich im Austausch von Anbietern und Abnehmern entwickelt haben.
- (2) Normativ-demokratieorientierte Ebene (ND-Ebene): Die ND-Ebene umfasst alle Kriterien, die sich aus der gesellschaftlichen Moral und dem Ethikverständnis ableiten lassen. Die Kriterien der ND-Ebene stehen in Zusammenhang zum politischen System.
- (3) Publikumsbezogene-handlungsorientierte Ebene (PH-Ebene): Die PH-Ebene bezieht sich auf die Anforderungen und Bedürfnisse der Rezipienten. Teilaspekte sind etwa die Verständlichkeit oder der Anwendungsbezug journalistischer Beiträge.

In den vergangenen Jahrzehnten haben sich insbesondere deutschsprachige Wissenschaftler mit der angebotsorientierten Qualitätsforschung befasst. Die Politikwissenschaftler Schatz und Schulz veröffentlichten im Jahr 1992 eine Arbeit, in der sie mit der Vorstellung aufräumten, dass Qualität eine singuläre Größe sei. Stattdessen schlugen sie vor, die journalistische Qualität in fünf Unterkategorien aufzuteilen: Vielfalt, Relevanz, Professionalität, Akzeptanz und Rechtmäßigkeit (Schatz & Schulz, 1992 ff). Zwei Jahre später nahm Rager in seinem Qualitätsmodell eine ähnliche Aufteilung vor, setzte dabei aber andere Schwerpunkte (Rager, 1994). Professionalität und Rechtmäßigkeit sind für Rager keine Qualitätskriterien, sondern Voraussetzungen der journalistischen Arbeit. Rager definiert im Gegensatz zu Schatz und Schulz vier Qualitätskriterien: Aktualität, Relevanz, Korrektheit und publikumsgerechte Aufbereitung. Die Vielfalt ist bei Rager kein Teilkriterium, sondern

stattdessen das allumfassende Ziel der journalistischen Berichterstattung. Ein weiteres Modell geht auf Ruß-Mohl zurück, der eine Aufteilung der Qualität in die Kategorien Aktualität, Relevanz, Korrektheit, Originalität und Komplexitätsreduktion vorschlägt (Beck et al., 2010, S. 24). Alle obengenannten Modelle sind vielschichtig und komplex. Dieser Abschnitt soll einen Überblick über einige wichtige Kategoriensysteme geben, die als Grundlage für empirische Arbeiten herangezogen werden.

Im Vergleich zeigt sich, dass es große Überschneidungen zwischen den Modellen gibt (Spurk & Lublinski, S. 3). Im Methodenteil einer inhaltsanalytischen Arbeit gelang es den Forschern, die Ergebnisse unterschiedlicher Qualitätsstudien in zehn Kriterien zusammenzufassen (Spurk & Lublinski, o. J., S. 3): Vielfalt, Aktualität, Relevanz, Glaubwürdigkeit, Korrektheit, Verständlichkeit, Unabhängigkeit, Kritikfähigkeit, Unparteilichkeit und Ausgewogenheit. Schweiger und Urban konnten die Zahl der Kriterien sogar auf sechs reduzieren. Ein Beispiel für eine große Qualitätsstudie im deutschsprachigen Raum ist das Schweizer Medienqualitätsranking (MQR). Das MQR kombiniert bei der Bewertung Publikumsbefragungen mit der Methode der Inhaltsanalyse. Im MQR wird die Qualität des Nachrichtenangebots anhand von vier Kriterien ermittelt: Relevanz, Vielfalt, Professionalität und Eigenleistung. Für jedes Kriterium wird in einem komplexen Verfahren ein Punktwert ermittelt und anschließend wird aus den Einzelbewertungen eine Gesamtpunktzahl errechnet (Eisenegger et al., 2018). Kurz zusammengefasst lässt sich festhalten: Die angebotsorientierte Qualitätsforschung hat eine jahrzehntelange Tradition und ist die Grundlage vieler empirische Arbeiten. Im Gegensatz zur nachfrageorientierten Forschung wird die Qualität von Wissenschaftlern definiert. Obwohl sich die theoretischen Ansätze der Forscher stark unterscheiden, gibt es große Überschneidungen zwischen den Kategoriensystemen.

4.2 Nachfragemodelle

Die Beweggründe, die der individuellen Nachrichtennutzung zu Grunde liegen, können sehr unterschiedlich sein. So könnte beispielsweise ein Angestellter, der jeden Tag mit dem Auto zu seiner Arbeitsstelle pendelt, sehr an Nachrichten zur aktuellen Verkehrslage interessiert sein, während jemand, der die öffentlichen Verkehrsmittel nutzt, diese Informationen als unwichtig empfindet (Lacy & Rosenstiel, 2015). Das Forschungsfeld,

das die Bedürfnisse Konsumenten untersucht, wird in der Literatur als „uses and gratifications“- Forschung bezeichnet (vgl. Ruggiero, 2000). Obwohl sich die ermittelten Bedürfnisse von Studie zu Studie unterscheiden, konnte Lacy in einer Metastudie fünf Hauptmotive der Nachrichtennutzung ausmachen (Lacy, 2000, S. 38):

- (1) Beobachtung (surveillance) → Informationen über aktuelle Ereignisse („Auf dem Laufenden bleiben“)
- (2) Unterhaltung (entertainment)
- (3) Soziale Teilhabe (social interaction)
- (4) Aufklärung (personal understanding)
- (5) Entscheidungsfindung (decision making)

Diese obengenannten Motive werden von allen Menschen geteilt. Daneben gibt es Spezialbedürfnisse, die nur bei Mitgliedern bestimmter gesellschaftlicher Gruppen zu finden sind. Auch wenn diese Bedürfnisse bei Menschen auftreten, die gerade nicht in direktem Kontakt mit anderen Gleichgesinnten stehen, sind sie immer auf soziale Interaktionen zurückzuführen. Laut Lacy und Rosenstiel lassen sich auch diese in fünf Kategorien einteilen (Lacy & Rosenstiel, 2015, S. 13):

- (1) Politik (political/ government) → Informationen über politische Entscheidungen und Gesetzgebungsverfahren
- (2) Wirtschaft (economic/ business) → Informationen über Unternehmen, aktuelle Konjunkturdaten, die Aktienmärkte und den Arbeitsmarkt
- (3) Soziale Aktivitäten (social activities) → Informationen über Freizeitaktivitäten (Sportveranstaltungen, Restaurantsbesuche)
- (4) Nutzwert-Nachrichten (news you can use) → Lösungsstrategien für Herausforderungen des Alltags (Heimwerken, Geldanlage, Gesundheit)
- (5) Kultur (cultural activities) → Informationen über kulturelle Ereignisse (Konzerte, Ausstellungen)

Eine Organisation, die sich eingehend mit den Nutzerbedürfnissen auseinandergesetzt hat, ist das US-amerikanische „Project for excellence in journalism“ (Lacy & Rosenstiel, 2015, S. 13). Seit 1997 wurden dort regelmäßig Befragungen durchgeführt, in denen ebenfalls fünf Bedürfniskategorien unterschieden wurden (Purcell et al., 2010). In den vergangenen Jahren wurden die Fragebögen noch weiter ausdifferenziert und enthalten

mittlerweile sieben Oberkategorien, die teilweise in mehrere Unterkategorien aufgeteilt wurden. Demnach erfüllen Nachrichten für das Publikum folgende Funktionen (Lacy & Rosenstiel, 2015, S. 14):

- (1) Teilhabe an Gesprächen über aktuelle Ereignisse
- (2) Politische Aufklärung
 - a. Ermächtigung zur politischen Teilhabe
 - b. Motivation sich in die Gemeinschaft einzubringen
- (3) Praktische Lebenshilfe
 - a. Finanzen (Geld sparen)
 - b. Gesundheit
 - c. Alltägliche Probleme
 - d. Kindererziehung
 - e. Konsumententscheidungen
- (4) Unterhaltung und Entspannung
- (5) Unterstützung im Berufsleben
- (6) Meinungsbildung
- (7) Kontakt zur Gesellschaft

Obwohl die individuellen Bedürfnisse sehr unterschiedlich sind, zeigen langjährige Untersuchungen, dass viele Menschen ähnliche Qualitätsvorstellungen teilen. Die meisten Motive für den Nachrichtenkonsum können in wenigen Kategorien zusammengefasst werden. Die nachfrageorientierte, empirische Qualitätsforschung basiert auf der Annahme, dass einige Wünsche und Verhaltensmuster integrale und unveränderliche Bestandteile der menschlichen Psyche sind (Lacy & Rosenstiel, 2015, S. 15). In den vergangenen Jahren haben sich einige Studien mit dem Qualitätsverständnis des Publikums auseinandergesetzt. Diese lassen sich in drei Kategorien einteilen (Urban & Schweiger, 2013, S. 4):

1) Explorative Studien zur Definition von Qualitätskriterien

Explorative Studien werden meist durchgeführt, um neue Qualitätskriterien zu ermitteln oder bestehende Modelle zu erweitern. Diese Studien bestätigen, dass die Qualitätsvorstellungen von Laien und Experten größtenteils übereinstimmen (Urban & Schweiger, 2013, S. 4). So legen Rezipienten beispielsweise Wert auf neutrale und

ausgewogene Berichterstattung und eine möglichst große Meinungsvielfalt (Gladney, 1996; Willnata & Weaver, 1998).

2) Verhaltensorientierte Studien

Verhaltensorientierte Studien beschäftigen sich mit der Frage, welchen Einfluss das Vorhandensein oder die Abwesenheit bestimmter Qualitätskriterien auf das Verhalten der Rezipienten hat (Urban & Schweiger, 2013, S. 4). Diese Untersuchungen zeigen, dass die Rezipienten bevorzugt Beiträge auswählen, die sie als relevant, ausgeglichen und verständlich empfinden (Wolling, 2002).

3) Evaluationsstudien

Evaluationsorientierte Studien werden durchgeführt, um die Qualität des Medienangebots zu bewerten (Urban & Schweiger, 2013, S. 4). Diese Studien deuten darauf hin, dass die Rezipienten die Qualität von Leitmedien in der Regel höher bewerten als die von Boulevardmedien (Gscheidle & Geese, 2017, S. 317ff).

4) Untersuchungen externer Einflussfaktoren

Es gibt zahlreiche Faktoren, die im Verdacht stehen, die Qualitätswahrnehmung der Rezipienten zu beeinflussen, wie beispielsweise das Vertrauen in einen Medienanbieter, persönliche Erfahrungen oder Nutzungsgewohnheiten. So konnten etwa Urban und Schweiger in einem 2x2 Experiment aus dem Jahr 2013 zeigen, dass das Vertrauen in einen Nachrichtenanbieter die Bewertung eines Textbeitrags beeinflussen kann (Urban & Schweiger, 2013).

5 Methodik

5.1 Forschungsfragen und Hypothesen

Wie in Abschnitt 2.1 dargelegt wurde, entzieht sich die journalistische Qualität einer allgemeingültigen Definition. Gleichwohl gibt es Konventionen darüber, welche Eigenschaften hochwertige Nachrichtenbeiträge ausmachen (vgl. Haas & Lojka, 1998, S. 117 ff). Empirische Studien zeigen, dass Wissenschaftler und Nicht-Medienexperten meist in der Beurteilung übereinstimmen, ob ein Nachrichtenbeitrag qualitativ hoch- oder minderwertig ist (Dahinden et al., 2004; Gladney, 1996). Im Gegensatz zu Wissenschaftlern haben Laien in der Regel aber keine exakten Vorstellungen davon, was journalistische Qualität ausmacht. Ein qualitative Studie aus dem Jahr 2012 zeigt, dass nur wenige der Lage sind konkrete Qualitätsmerkmale zu benennen (Fiedler, 2013). Stattdessen verlassen sie sich bei der Bewertung auf Erfahrungen und Intuitionen. Da die Bewertungen von Laien nicht auf einem theoretischen Fundament aufbauen, können sie einfacher durch externe Faktoren beeinflusst werden. Diese Arbeit nimmt die Perspektive der Rezipienten ein und untersucht zwei mögliche Einflussfaktoren auf das Qualitätsempfinden. Es gibt Hinweise darauf, dass die Nachrichtenmarke die Wahrnehmung eines Beitrags beeinflussen kann (Urban & Schweiger, 2013, S. 5). In der Betriebswirtschaft wird die Wechselwirkung von Marke und Produkt schon seit vielen Jahren erforscht (Aaker, 1996; Aaker & Keller, 1990). Auch für viele Unternehmer scheint der Einfluss der Marke auf das Produkt eine Gewissheit zu sein. Der deutsche Werbemarkt hatte im Jahr 2019 ein Gesamtvolumen von 32,5 Milliarden Euro (Nielsen, 2020). Einen nicht unerheblichen Teil dieser Summe geben die Unternehmen dafür aus, ihr Markenbild zu verbessern. Die Betriebswirtschaftslehre geht davon aus, dass bei der Güterbewertung ein Imagetransfer von der Marke auf das Produkt stattfindet. Ein Ziel dieser Arbeit ist es daher herauszufinden, ob sich dieser Zusammenhang auch bei journalistischen Erzeugnissen beobachten lässt. Neben der Nachrichtenmarke könnte auch der Verbreitungsweg die Qualitätswahrnehmung beeinflussen. Seit der Einführung des Internets haben sich die Verbreitungswege von Nachrichten vervielfältigt. Besonders die Nachrichtennutzung in den sozialen Netzwerken hat in den vergangenen Jahren einen Boom erlebt. Für die unter 35-Jährigen in Deutschland sind die sozialen Medien heute schon die wichtigste Online-Nachrichtenquelle (Hölig & Hasebrink, 2020, S.23). Einige Wissenschaftler vermuten, dass zwischen der Verbreitung der sozialen Netzwerke und

dem weltweit sinkenden Vertrauen in die Medien ein Zusammenhang besteht. Facebook, YouTube und Instagram stehen häufig in der Kritik, weil sie durch ein stark personalisiertes Nutzererlebnis zur Bildung von Filterblasen beitragen (Spohr, 2017). Der Nachrichtenkonsum über die sozialen Netzwerke wird oft als Ursache einer zunehmenden Polarisierung gesehen, die sich in den vergangenen Jahren in vielen westlichen Gesellschaften beobachten ließ (Groshek & Koc-Michalska, 2017). Die Vermutung liegt also nahe, dass der Verbreitungsweg „soziale Medien“ einen negativen Einfluss auf das Qualitätsempfinden haben könnte. Diese Arbeit konzentriert sich daher auf zwei Forschungsfragen:

Forschungsfrage 1: Welchen Einfluss hat die Nachrichtenmarke auf die Wahrnehmung eines journalistischen Beitrags?

Die Literatur legt nahe, dass das Image einer Medienmarke die Bewertung eines Beitrags beeinflusst. In einer Studie aus dem Jahr 2013 bewerteten die Probanden einen Artikel mit dem Logo der Bildzeitung schlechter als einen wortgleichen Text mit dem Logo der Süddeutschen Zeitung (Urban & Schweiger, 2013). Es gibt eine Debatte darüber wie stark der Effekt ausfällt. Für die Mediengattungen Ton und Video ist der Zusammenhang weniger gut erforscht. Es ist unklar, wie groß der Vertrauensunterschied der Medienmarken sein muss, damit ein Effekt messbar wird. Für den Vergleich einer bekannten Marke mit einem unbekanntem Anbieter, gibt es ebenfalls keine Daten.

Forschungsfrage 2: Welchen Einfluss hat der Verbreitungsweg auf die Wahrnehmung eines journalistischen Beitrags?

In den vergangenen Jahren ist das Vertrauen in die Nachrichten weltweit zurückgegangen. In Deutschland wurden bis zum Jahr 2015 relativ stabile Vertrauenswerte gemessen. Aktuelle Zahlen belegen aber auch hierzulande einen leichten Vertrauensrückgang (Hölig et al., 2019b, S. 25). Einige Wissenschaftler vermuten, dass der Vertrauensverlust mit der Nachrichtennutzung über die sozialen Medien in Zusammenhang stehen könnte. Die Studien, die sich mit dem Thema befasst haben, kommen zu unterschiedlichen und teils widersprüchlichen Ergebnissen. Publikumsbefragungen zeigen, dass ein Großteil der Menschen den sozialen Netzwerken ein grundsätzliches Misstrauen entgegenbringt (Hölig & Hasebrink, 2020, S. 33). Ob

dieses Misstrauen auch die Bewertung einzelner Beiträge beeinflusst, wurde bislang nicht erforscht.

Diese Arbeit steht in der Tradition der experimentellen Qualitätsforschung. Im Gegensatz zu vorherigen Studien werden ausschließlich Online-Nachrichten untersucht. Die Untersuchung findet auf der Mikroebene statt und stellt Einzelbeiträge in den Fokus. Ausgehend von der Theorie und den empirischen Studien können folgende Hypothesen aufgestellt werden:

Hypothese 1: Nachrichtenbeiträge, die das Logo einer vertrauenswürdigen Medienmarke tragen, werden besser bewertet, als inhaltlich identische Beiträge, die das Logo einer weniger vertrauenswürdigen Marke tragen.

Hypothese 2: Nachrichtenbeiträge, die über die sozialen Medien verbreitet werden, werden schlechter bewertet als inhaltlich identische Beiträge, die ausschließlich über die Webseite des jeweiligen Anbieters verbreitet werden.

5.2 Studiendesign

Das Ziel des Studiendesigns war es, ein Experiment zu entwerfen, mit dem der Einfluss der Nachrichtenmarke und des Verbreitungswegs auf die Qualitätswahrnehmung gemessen werden konnte. Um Korrelationen aufzeigen zu können und eine Vergleichbarkeit der Ergebnisse zu garantieren, wurde ein quantitativer Forschungsansatz gewählt. Das Experiment sollte dabei den wissenschaftlichen Gütekriterien der Validität, Objektivität und Reliabilität genügen. Für den Versuch sollte eine möglichst repräsentative Stichprobe der deutschsprachigen Nutzer von Online-Nachrichten gezogen werden. Im nächsten Schritt sollten die Versuchsteilnehmer die Qualität journalistischer Beiträge bewerten, die sich in Bezug auf die Nachrichtenmarke und den Verbreitungsweg unterscheiden. Abschließend sollten die Antworten mit statistischen Methoden ausgewertet werden. Um sicherzugehen, dass mögliche Bewertungsunterschiede eindeutig auf einen der untersuchten Faktoren zurückzuführen sind, wurde beim Versuchsplan ein 2x2 Design gewählt. Diese Anordnung erlaubt es, mit einem Experiment den Einfluss zweier unabhängiger Variablen zu untersuchen (→ Validität). Die grundlegende Idee ist einfach umzusetzen und vielfach erprobt: Jeder

Teilnehmer wird per Zufallsprinzip in eine von vier Versuchsgruppen eingeteilt. Auf diese Weise soll verhindert werden, dass die Ergebnisse durch externe Einflussnahmen verzerrt werden (→ Objektivität). Während des Experiments bekommen alle Teilnehmer dieselben Bewertungsaufgaben, aber jeder Gruppe wird ein anderer Stimulus präsentiert. Zeigen sich bei der Auswertung der Antworten signifikante Unterschiede zwischen den Gruppen, kann davon ausgegangen werden, dass die Abweichungen auf die Stimuli zurückzuführen sind. Um ein umfassendes Verständnis davon zu bekommen, wie Verbreitungswege und Nachrichtenmarken die Qualitätswahrnehmung beeinflussen, wurde die Entscheidung getroffen, die Hypothesen an möglichst vielen unterschiedlichen Beiträgen zu testen. Dadurch sollte sichergestellt werden, dass die Ergebnisse reproduzierbar sind (→ Reliabilität). Darüber, welche Beiträge in das Experiment aufgenommen wurden, wurde in einem mehrstufigen Auswahlverfahren entschieden. Im ersten Auswahlschritt mussten die Mediengattungen der Beiträge bestimmt werden. Viele experimentelle Studien haben sich in der Vergangenheit auf Textbeiträge wie Reportagen, Features oder Berichte konzentriert. Texte haben den Vorteil, dass sie im Gegensatz zu Bildern und Tonaufnahmen weniger Interpretationsspielraum zulassen, da missverständliche Elemente wie Aussprache, Mimik und Gestik ausgeblendet werden. Außerdem können der Inhalt und das Design von Texten relativ einfach verändert werden, was bei der Erstellung unterschiedlicher Versionen von Vorteil ist. Dazu kommt, dass Texte auch im Online-Zeitalter die zentrale Säule des Nachrichtenangebots sind. Allerdings hat besonders die Nutzung von Videonachrichten in den vergangenen Jahren stark zugenommen. Für 26 Prozent der deutschen Onliner sind Videos die bevorzugte Nachrichtenquelle (Hölig & Hasebrink, 2020, S. 49) und etwa 40 Prozent der Deutschen nutzen regelmäßig die Videoplattform YouTube (Schwaderer, 2020, S. 24). Daher wurde die Entscheidung getroffen neben Nachrichtentexten auch Videos in das Experiment aufzunehmen. Zwar erleben auch Audioinhalte mit der Verbreitung von Podcast einen Boom, allerdings haben diese nur selten einen Nachrichtenbezug und sind mit einer durchschnittlichen Dauer zwischen 30 und 90 Minuten zu lang für das Experiment. Nachdem die Mediengattungen festgelegt waren, wurden im zweiten Schritt die Themen der Beiträge bestimmt. Das Ziel war es, für eine möglichst große inhaltliche Vielfalt zu sorgen. Außerdem sollten mindestens die Hälfte der Beiträge einen Bezug zum aktuellen Nachrichtengeschehen haben. Die beiden Ressorts Politik und Wirtschaft waren aufgrund ihrer herausragenden nachrichtlichen Bedeutung gesetzt. Zusätzlich wurden mit den Ressorts Gesundheit, Digitales und Medien drei Themenbereiche ausgewählt, die im

Online-Journalismus auf ein großes Nutzerinteresse stoßen. So hat etwa der Spiegel-Verlag das neue Ressort „Leben“ geschaffen, um so dem wachsenden Interesse der Leser an Gesundheitsthemen Rechnung zu tragen (Dürr & Lübke-Narberhaus, o. J.). Um wirklichkeitsgetreue Versuchsbedingungen zu schaffen, wurde die Entscheidung getroffen, anstelle von selbsterstellten Texten und Videos, veröffentlichte Originalbeiträge zu verwenden. Dieses Vorgehen erfordert etwas mehr Zeit beim Auswahlprozess, sorgt aber für eine hohe inhaltliche und sprachliche Diversität. Um die Konzentration der Studienteilnehmer zu erhalten, wurden Beiträge mit einer maximalen Lese- bzw. Abspieldauer von vier Minuten gesucht. Das Ziel war es, eine Gesamtbearbeitungszeit von 20 Minuten nicht zu überschreiten. Mehrere Studien zeigen, dass ein Großteil der Teilnehmer kürzere Fragebögen bevorzugt (Galesic & Bosnjak, 2009; Hoerger, 2010). Neben der Dauer und der inhaltlichen Vielfalt wurde auch darauf geachtet, dass alle Texte von unterschiedlichen Verlagen und Autoren stammten. Außerdem sollte die inhaltliche Qualität der Beiträge verschieden hoch sein. Nachdem drei Texte und zwei Videos als Testobjekte festgelegt waren, wurden von jedem Beitrag vier Versionen erstellt. Alle Versionen waren inhaltlich identisch, aber deren Design wurde so manipuliert, dass sie sich mit Blick auf die Nachrichtenmarke und den Verbreitungsweg unterschieden. Die folgende Tabelle gibt einen Überblick über die Versionen:

Qualitätsmarke & klassischer Verbreitungsweg	Qualitätsmarke & soziale Medien
Boulevardmarke & klassischer Verbreitungsweg	Boulevardmarke & soziale Medien

Tabelle 2: Versuchsgruppen des 2x2 Experiments

Die in der Tabelle verwendete Bezeichnung „Qualitätsmarke“ dient als Sammelbegriff für Medienmarken, die ein überdurchschnittliches Publikumsvertrauen genießen. Als Referenzwerte wurden die Ergebnisse des Digital News Reports herangezogen (Hölig & Hasebrink, 2020, S. 34f). In dieser Arbeit werden die Süddeutsche Zeitung, der Bayerische Rundfunk, die Zeit und der Spiegel als Qualitätsmarken bezeichnet. Die Apotheken Umschau wurde zusätzlich in die Kategorie aufgenommen, weil sie das meistbesuchte Gesundheitsportal in Deutschland ist und in Leserbefragungen in den Kategorien Glaubwürdigkeit und Verständlichkeit Bestnoten erzielt (Kreutner, 2019). Im Umkehrschluss werden Anbieter, die unterdurchschnittliche Vertrauenswerte genießen

oder einer breiten Öffentlichkeit unbekannt sind, in dieser Arbeit als Boulevardmarken bezeichnet. Tabelle 3 zeigt die ausgewählten Medienmarken in alphabetischer Reihenfolge.

Qualitätsmarken	Boulevardmarken/ alternative Medien
Apotheken Umschau	Bildzeitung
Bayerischer Rundfunk (BR24)	Bunte
Der Spiegel	Nachdenkseiten
Die Zeit	Technikfaultier
Süddeutsche Zeitung	Web.de

Tabelle 3: Übersicht über die ausgewählten Medienmarken

Der zweite Faktor, der beim Versuchsaufbau variiert wurde, ist der Verbreitungsweg. Wie in Abschnitt 3.3 gezeigt wurde, haben sich die Zugangswege zu Nachrichten in den vergangenen Jahren vervielfältigt. In Tabelle 2 werden die Zugangswege in vier Kategorien eingeteilt. Die beiden konträren Pole bilden die sozialen Medien und der direkte Zugangsweg. Die Nachrichtennutzung über die sozialen Netzwerke basiert auf Algorithmen und lässt auf ein unterdurchschnittliches Informationsinteresse schließen, während der Direktzugriff Selbstbestimmtheit und ein hohes Informationsbedürfnis voraussetzt. Um bei der Entwicklung der Testaufgaben einen möglichst großen Kontrast zu schaffen, wurden daher die beiden Extreme ausgewählt. Der Direktzugang über die Webseite des Anbieters wird von nun an als klassischer Verbreitungsweg bezeichnet. Die Bezeichnung „soziale Medien“ ist selbsterklärend. Im Bereich der sozialen Netzwerke wird der deutsche Markt von zwei US-amerikanischen Unternehmen dominiert. Laut den Zahlen der Onlinestudie von ARD und ZDF sind YouTube, Facebook und Instagram mit großem Abstand die meistgenutzten Plattformen. 53 Prozent der Erwachsenen Onliner in Deutschland sehen sich mindestens einmal pro Monat ein Video auf YouTube an. Facebook und Instagram werden monatlich von 35 Prozent bzw. 21 Prozent der Bundesbürger genutzt (Beisch et al., 2019, S. 8,12). Die Nutzerzahlen anderer Netzwerke wie Snapchat, Twitch oder Twitter sind deutlich niedriger (siehe Tabelle 4).

Soziales Netzwerk	Monatsreichweite
YouTube	≈ 37,3 Millionen
Facebook	≈ 24,7 Millionen
Instagram	≈ 14,8 Millionen
Snapchat	≈ 5,6 Millionen
Twitter	≈ 3,5 Millionen

Tabelle 4: Monatsreichweiten wichtiger sozialer Netzwerke (vgl. Beisch et al., 2019, S. 383)

Um den Verbreitungsweg soziale Medien möglichst realitätsnah abzubilden, wurden die drei großen Netzwerke als Vorlage für das Design der Testaufgaben verwendet. Jeweils zwei Beiträge wurde im Stil von YouTube und Facebook gestaltet. Als Vorlage des fünften Beitrags diente das Design von Instagram. Die graphische Umgestaltung erfolgte mit Hilfe der Bild- und Videobearbeitungsprogramme Adobe Photoshop und Adobe Premiere Pro. Inhaltlich waren alle vier Versionen identisch. Manipuliert wurden ausschließlich das Markenlogo und das Layout der Beiträge. Bei den Versionen, die im Stil der sozialen Medien gestaltet wurden, wurde darauf geachtet, dass im Bild keine weiteren Links oder Beitragsempfehlungen zu sehen sind, die die Ergebnisse verzerren könnten. Auch die Fotos, die zur Illustration der Artikel verwendet wurden, waren in allen vier Versionen identisch. Der nächste Abschnitt gibt einen Überblick über die gewählten Beiträge.

5.3 Überblick über die Teilaufgaben

Teilexperiment 1: Boulevard- und Verbraucherjournalismus?

Im Themenbereich Gesundheit wurde ein Video der Bildzeitung ausgewählt, das am 18. April auf der Plattform YouTube veröffentlicht wurde. Das Originalvideo ist fünf Minuten lang. Für das Experiment wurde es auf eine Länge von drei Minuten gekürzt. Das Video zeigt einen Ausschnitt aus einer Sondersendung zur Corona-Epidemie. Darin wird über ein neues Medikament berichtet, das in den USA zur Virusbekämpfung eingesetzt wird. Der Beitrag kann in drei Szenen unterteilt werden. Zu Beginn berichtet ein Live-Reporter davon, dass der Wirkstoff Remdesivir, der ursprünglich zur Behandlung des Ebola-Virus entwickelt wurde, in den USA auch an Covid-19 Patienten getestet wird. Anschließend kommt ein Patient aus Deutschland zu Wort, der das Medikament im Rahmen einer experimentellen Studie erhalten hat und seitdem wieder

genesen ist. Das Video endet mit einem Interview, in der eine Moderatorin einen Arzt zur Wirksamkeit des Medikaments befragt. Die Machart des Videos ist typisch für eine Boulevardzeitung. In der Unterzeile des Videos wird Remdesivir als neues „Wundermittel“ bezeichnet, und im Off-Text heißt es: „Eine Studie aus den USA zeigt: Mit Remdesivir könne Corona geheilt werden.“ Obwohl inklusive der Moderatorin fünf Personen zu Wort kommen, ist die Berichterstattung wenig ausgewogen. Das Video endet mit dem Satz der Moderatorin: „Das macht auf jeden Fall Hoffnung.“ Um einen Kontrast zur schrillen Aufmachung der Bildzeitung zu schaffen, wurden zwei der vier Versionen im aufgeräumten Design der Apotheken Umschau modelliert. Abbildung 2 zeigt alle vier Versionen im Überblick. Bei der Bearbeitung wurden die Logos ausgetauscht und die Farben der graphischen Elemente angepasst. Außerdem wurde das Mikrofon des Bildreporters hinter einer Bauchbinde versteckt.

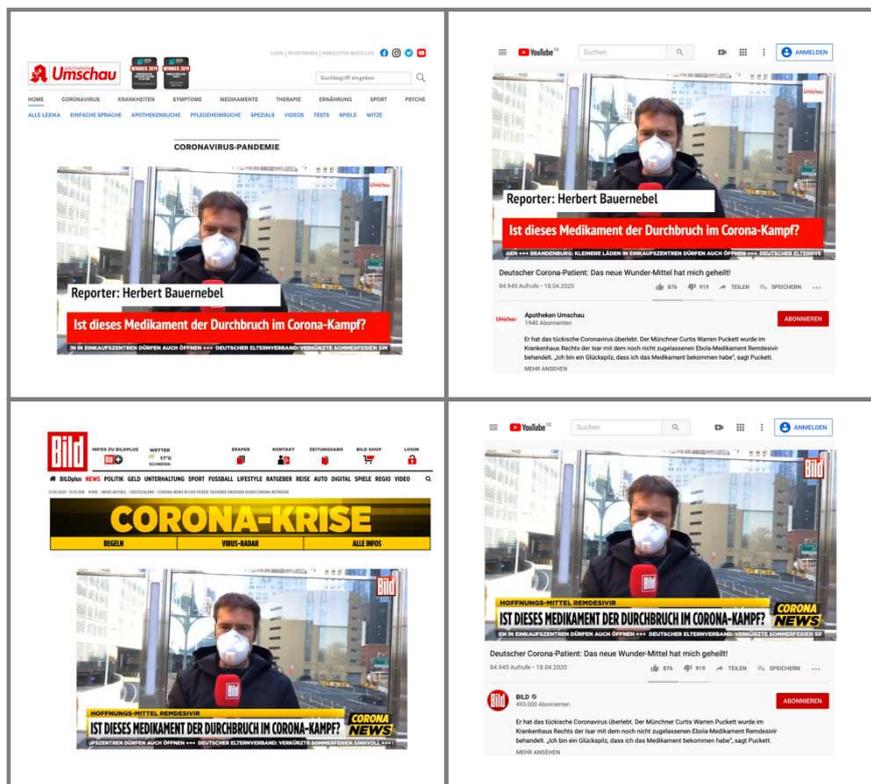


Abbildung 2: Überblick über die Beitragsversionen des ersten Teilerperiments

Teilerperiment 2: Investigative Recherche oder Agenturmaterial?

Im Themenbereich Wirtschaft/ Politik wurde ein Artikel gewählt, der von zwei Autoren des Investigativ-Ressorts der Süddeutschen Zeitung verfasst wurde. Der Artikel wurde am 27.05.2020 veröffentlicht und berichtet darüber, dass die Lufthansa AG

möglicherweise Geld in Steueroasen versteckt. Der Artikel ist ein Musterbeispiel eines aufwendig recherchierten Beitrags. Im Text werden fünf unabhängige Quellen zitiert. Außerdem wird deutlich gemacht, dass die Autoren der Lufthansa mehrmals die Gelegenheit gegeben haben, auf die Vorwürfe zu reagieren. Zur Gegenprobe wurden zwei Varianten im Stil des Nachrichtenportals Web.de umgestaltet. Web.de greift bei der Berichterstattung häufig auf das Material von Nachrichtenagenturen zurück. Im Gegensatz dazu ist die Süddeutsche Zeitung als Autorenblatt für investigative Recherchen bekannt (Panama Papers, Ibiza-Affäre). In der Gruppe der sozialen Netzwerke, dient das Facebook Design als Vorlage. Im Vergleich zum Originalbeitrag wurde der Text um zwei Absätze gekürzt. Außerdem wurde das Agenturbild gegen ein lizenzfreies Foto ausgetauscht.

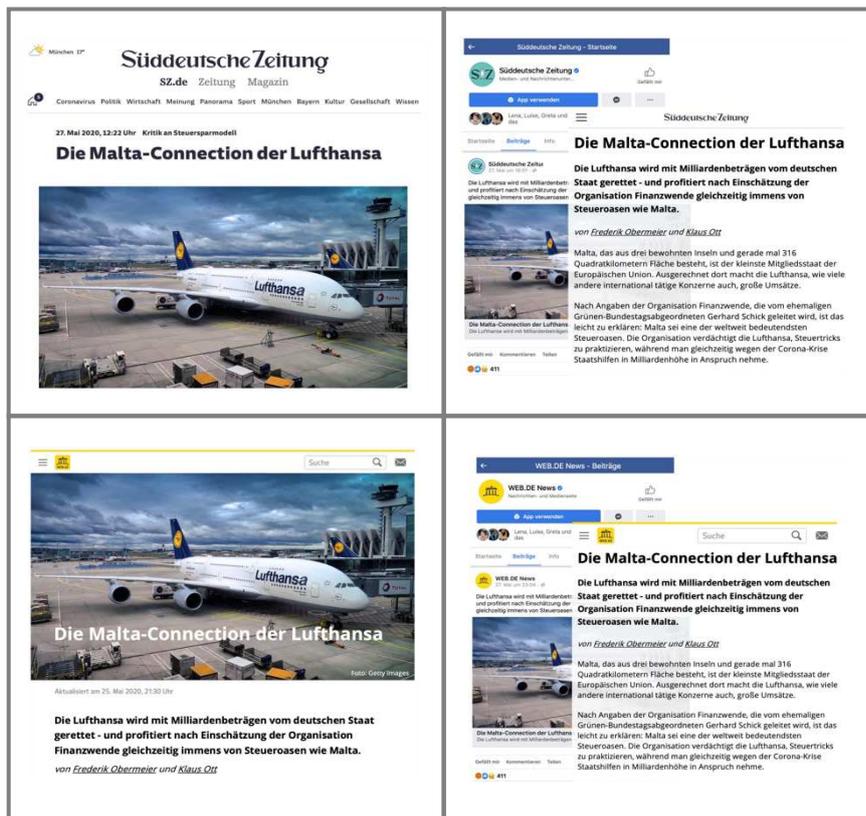


Abbildung 3: Überblick über die Beitragsversionen des zweiten Telexperiments

Telexperiment 3: Journalist oder Influencer?

In der Kategorie Technik und Digitales wurde ein Video des Spiegelverlags gewählt, das am 17.09.2019 in der Spiegel Netzwelt veröffentlicht wurde. Darin stellt der Redakteur Matthias Kremp die neuen Smartphones der Apple Inc. vor. Das ursprünglich viereinhalb

Minuten lange Video, wurde für das Experiment um etwa 40 Sekunden gekürzt. In dem Video wird die Leistungsfähigkeit der neuen iPhone-Kamera beworben und Unterschiede zwischen den Modellen besprochen. Kritische Betrachter könnten den Testbericht als Werbebotschaft fehlinterpretieren, da er ein überwiegend positives Bild des Geräts zeichnet. Die Überschrift des Artikels lautet: „iPhone 11 ist ein No-Brainer“. Zur Gegenprobe wurden zwei Varianten im Design des relativ unbekanntes Technikblogs „Technikfaultier“ gestaltet. Als Schablone im Bereich der sozialen Medien diente der Instagram-Kanal des Blogs. Influencer und Blogger werden von Unternehmen häufig mit kostenlosen Geräten ausgestattet und erhoffen sich im Gegenzug einen positiven Testbericht. Die Grenze zwischen Journalismus und Werbung ist in diesem Bereich fließend.

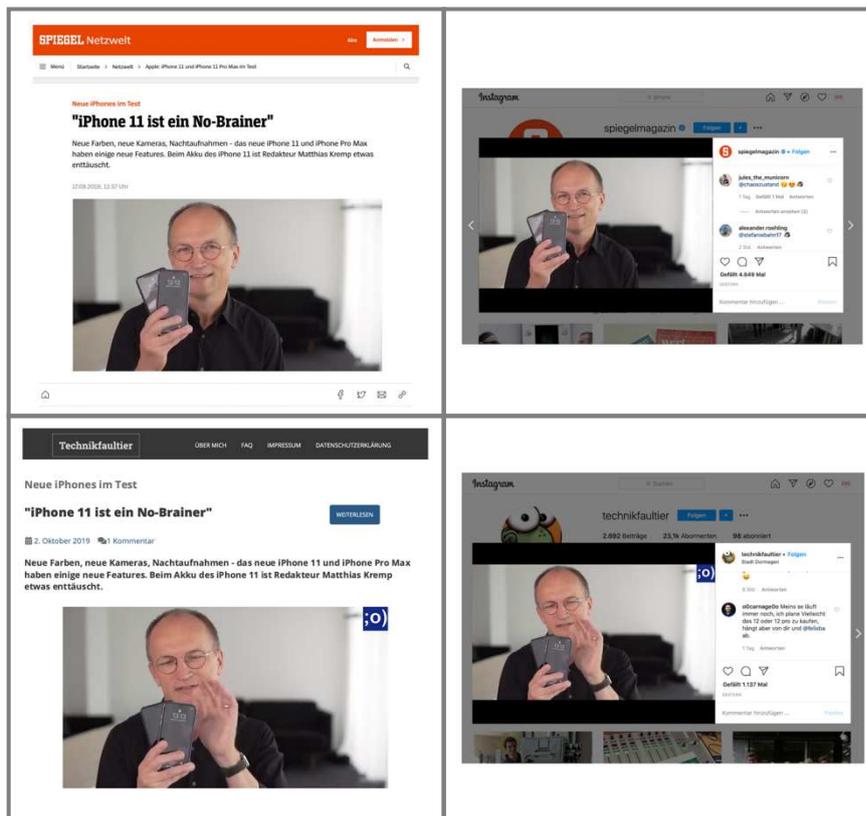


Abbildung 4: Überblick über die Beitragsversionen des dritten Teilerperiments

Teilerperiment 4: Öffentliches Interesse oder Voyeurismus?

Der Beitrag aus dem Themenbereich Politik/ Zeitgeschehen wurde im Gegensatz zu den anderen Teilaufgaben 1 bis 3 selbst erstellt. Allerdings ist er eng an die Berichterstattung der Buntten und der Bildzeitung vom 23.05.2017 angelehnt. Der Text beschreibt das

Chaos nach dem Terroranschlag auf ein Popkonzert in Manchester. Die Bildzeitung war für einen ähnlichen Beitrag vom Deutschen Presserat gerügt worden. Genau wie im Originalbeitrag wurde auch die Testaufgabe mit dem Foto eines minderjährigen Mädchens illustriert. Außerdem wurde eine fast identische Überschrift gewählt: „Ihre Mutter weiß noch nicht, dass Haley tot ist“. Die Versionen wurden im Design der Illustrierte „Bunte“ und der Nachrichtenseite des Bayerischen Rundfunks gestaltet. Als Vorlage für die soziale-Medien-Version diente erneut das YouTube Layout. YouTube wird häufig als reine als Videoplattform wahrgenommen. Phil Benson argumentiert aber in seinem Buch „The discourse of YouTube“, dass auch Texte ein integraler Bestandteil des Netzwerks sind (Benson, 2017, S. 79 f). Der Text wurde deshalb im Stil einer Videobeschreibung aufbereitet.

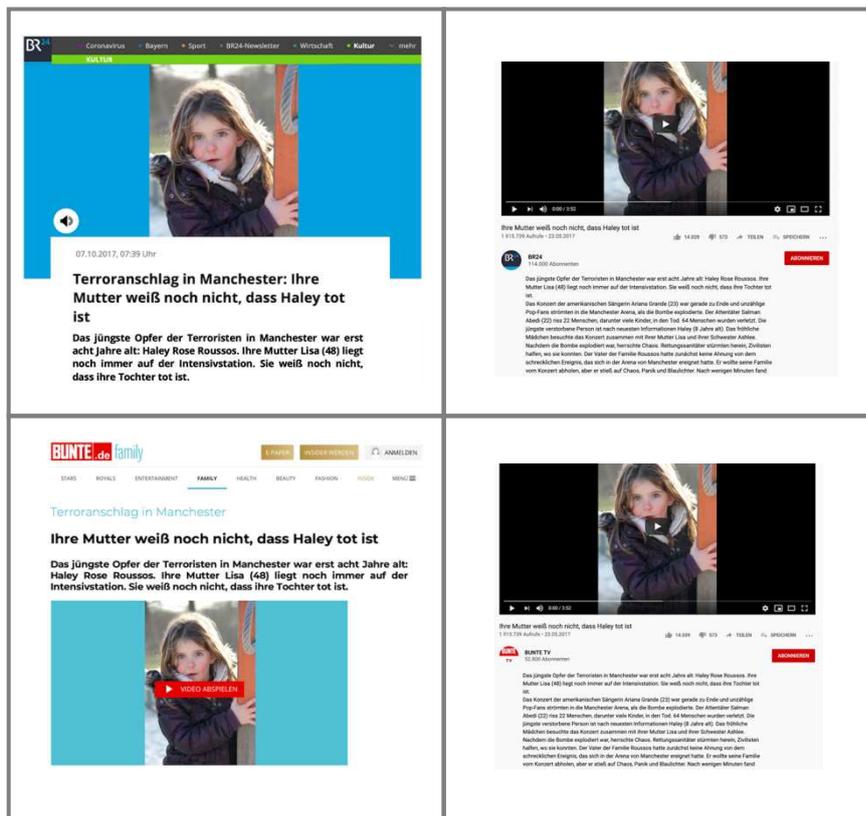


Abbildung 5: Überblick über die Beitragsversionen des vierten Teilexperiments

Teilexperiment 5: Kritischer Einwand oder Verschwörungstheorie?

Im Themenbereich Medien wurde ein Beitrag gewählt, der am 28.07.2017 auf den Nachdenkseiten beworben wurde. Ursprünglich stammt der Artikel von einem Autor des Media-Magazins. Für das Experiment wurde der Text an einigen Stellen umgeschrieben

und gekürzt. In dem Text berichtet der Autor darüber, dass die Journalistin Claudia Zimmermann keine Aufträge mehr bekommt, weil sie sich kritisch über die Staatsnähe des öffentlich-rechtlichen Rundfunks geäußert habe. Der Text räumt der Position von Frau Zimmermann viel Raum ein. Andere Quellen werden nicht zitiert. Allerdings ist der Autor um einen sachlichen Sprachstil bemüht. Die Varianten wurden im Stil der Nachdenkseiten und von Zeit-Online gestaltet. Im Bereich der sozialen Medien wurde das Facebook Design gewählt.

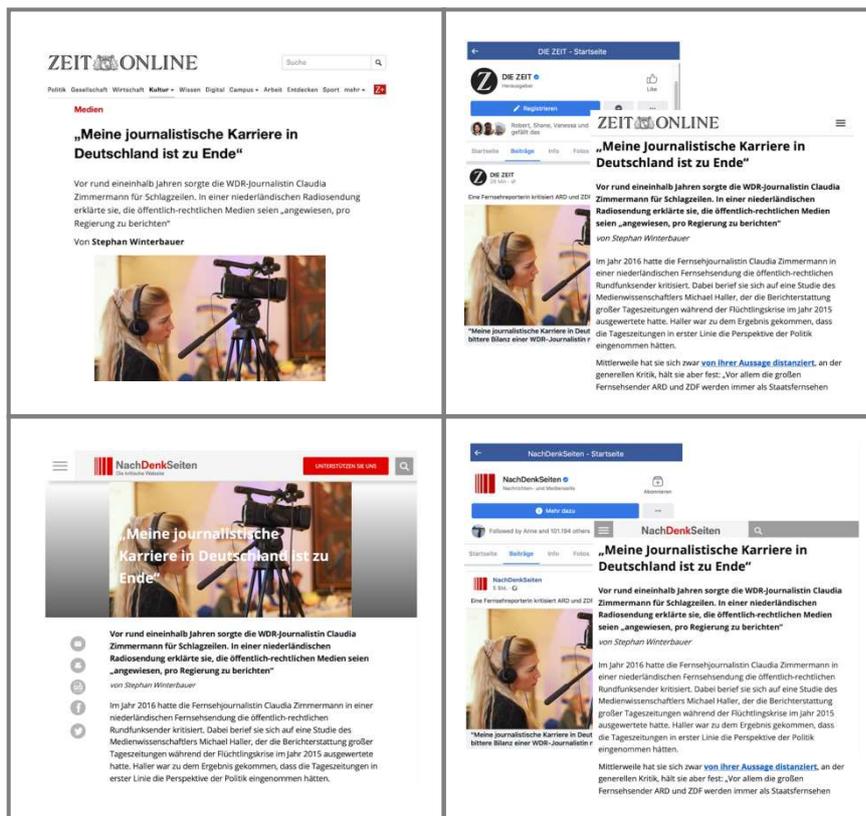


Abbildung 6: Überblick über die Beitragsversionen des fünften Telexperiments

5.4 Datenerhebung

5.4.1 Likert-Skalen und Antwortskalen des Likert Typs

Um individuelle Meinungen und Bewertungen in quantitative Werte zu übertragen, werden in der Sozialwissenschaft meist Antwortskalen des Likert Typs verwendet. Skalen des Likert Typs begrenzen die Menge möglicher Antworten auf wenige diskrete Werte. Dabei wählen die Teilnehmer aus einer begrenzten Zahl an Antwortoptionen diejenige aus, die ihrer persönlichen Einstellung am meisten entspricht. Die Antwortoptionen sind

ordinalskaliert und entweder in aufsteigender oder absteigender Reihenfolge sortiert. Der Vorteil der Skalen des Likert Typs besteht darin, dass es für die Befragten sofort offensichtlich ist, wie die Abstufungen der Skala zu verstehen sind. Bei Intervallskalen und stufenlosen Schieberegler besteht dagegen die Gefahr, dass die Teilnehmer vom Umfang der möglichen Antworten überfordert werden und in der Folge willkürliche Antworten abgeben. Obwohl Skalen des Likert Typs nominell ordinalskaliert sind, können sie unter gewissen Voraussetzung mit parametrischen Testverfahren ausgewertet werden. Die Bezeichnung Likert-Skala wird oft synonym für alle Skalen des Likert-Typs verwendet. Diese Vereinfachung ist aber unzulässig und kann im Extremfall zu falschen Ergebnissen führen. Eine echte Likert-Skala ist ein Kompositum aus mindestens vier Einzelskalen und wird zur Bewertung facettenreicher Objekte verwendet, die über mehrere relevante Eigenschaften verfügen.

5.4.2 Singuläre und kombinierte Qualitätsbestimmung

Das Ziel der Datenerhebung war es, eine kompakte und gleichzeitig differenzierte Qualitätsbewertung zu bekommen. Da die journalistische Qualität sowohl als Einzelgröße als auch als zusammengesetztes Konstrukt betrachtet werden kann, wurden zwei unterschiedliche Ansätze kombiniert: die singuläre und die kombinierte Qualitätsbestimmung. Für die singuläre Qualitätsbestimmung wurden die Teilnehmer darum gebeten, eine generische Bewertung des gesehenen Beitrags abzugeben. Die Formulierung der ersten Frage wurde bewusst allgemein gehalten: „Insgesamt betrachtet, wie bewerten Sie die journalistische Qualität des gerade gesehenen Beitrags?“ Die Antworten wurden mit einer siebenstufigen Skala des Likert Typs aufgezeichnet. Die Antwortoptionen reichten dabei von „sehr niedrig“ bis „sehr hoch“. Rolf Porst empfiehlt in seinem Buch „Fragebogen: Ein Arbeitsbuch“ eine Skala in maximal sieben Schritte zu gliedern, da viele Teilnehmer bei einer feineren Untergliederung den Überblick verlieren und in der Folge zufällige Antworten auswählen würden (Porst, 2014, S. 91 f). Außerdem wurde die Skala um eine „kann ich nicht beurteilen“ Option erweitert. Empirische Untersuchungen zeigen, dass auf diese Weise die Datenqualität erhöht werden kann, da unentschlossene Teilnehmer, ansonsten die Ergebnisse durch willkürliche Antworten verzerren können (R. Pennington, Nancy A. Pachana, Sa, 2001). Um den Teilnehmer einen intuitiven Zugang zu ermöglichen wurden alle Antwortoptionen zusätzlich mit lächelnden oder schmollenden Gesichtern illustriert (siehe Abbildung 7).

1. Insgesamt betrachtet, wie bewerten Sie die journalistische Qualität des gerade gesehenen Beitrags?
Die Qualität des Beitrags ist:



Abbildung 7: Siebenstufige Antwortskala des Likert Typs

Für die kombinierte Qualitätsbestimmung wurden unterschiedliche Qualitätsdimensionen berücksichtigt. Als Vorlage diente das Modell von Urban und Schweiger, das eine Aufteilung der Qualität in sechs Teilkriterien vorsieht: Genauigkeit, Unparteilichkeit, Verständlichkeit, Relevanz, Ethik und Vielfalt. Das Modell bietet zwei Vorteile: Erstens stützt es sich auf ein breites theoretisches Fundament (Arnold, 2009; Pöttker, 2000; Schatz & Schulz, 1992) und zweitens wurde es schon in ähnlicher Form in einer empirischen Arbeit zur Qualitätsbestimmung eingesetzt. Zusätzlich wurde das Modell noch um das Kriterium der Glaubwürdigkeit erweitert. Aus jedem Teilkriterium wurde ein Aussagesatz abgeleitet, der dessen wesentliche Bedeutung in kompakter Form zusammenfasst.

	Der Text/ das Video ...
Vielfalt	... lässt unterschiedliche Meinungen zu Wort kommen.
Relevanz	... ist gut recherchiert und liefert alle wichtigen Informationen.
Ethik	... respektiert die Persönlichkeitsrechte.
Verständlichkeit	... ist leicht verständlich.
Genauigkeit	... ordnet das Geschehen gut ein.
Unparteilichkeit	... berichtet neutral und unvoreingenommen.
Glaubwürdigkeit	... ist glaubwürdig.

Tabelle 5: Teildimensionen des gewählten Qualitätsmodells

Die Antworten wurden mit fünfstufigen Skalen des Likert Typs aufgezeichnet. Die Antwortoptionen reichten dabei von „stimme gar nicht zu“ bis „stimme voll zu“. Zusätzlich wurde den Teilnehmern über eine „kann ich nicht beurteilen“ Option die Möglichkeit eingeräumt, die Antwort zu verweigern. Studien zeigen, dass auf diese Weise die Datenqualität erhöht werden kann. Die Skalen wurden so konstruiert, dass sie bei

der Datenauswertung zu einer Likert-Skala zusammengesetzt werden können: Alle sieben Teilaspekte wurden mit derselben Skala abgefragt. Die Skalen waren symmetrisch ausgelegt und die Antwortoptionen wurden von links nach rechts aufsteigend geordnet.

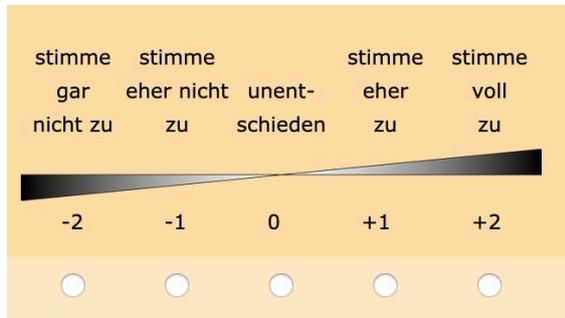


Abbildung 8: Fünfstufige Antwortskala

5.4.3 Gestaltung des Fragebogens

Rolf Porst nennt in seinem Arbeitsbuch drei zentrale Punkte, nach denen ein Fragebogen optimaler Weise gestaltet werden sollte, um möglichst unverzerrte Ergebnisse zu liefern:

- (1) Die Formulierung der Fragen sollte klar und einfach gehalten werden.
- (2) Der Fragebogen sollte eine eigene Dramaturgie haben.
- (3) Das Layout des Fragebogens sollte ansprechend gestaltet werden.

Bei der Zusammenstellung des Fragebogens wurde darauf geachtet, dass diese Vorgaben berücksichtigt werden. Zur Gestaltung und Durchführung des Fragebogens wurde die Umfragesoftware ScoSci-Survey verwendet. Die Software kann für studentische Arbeiten kostenlos verwendet werden und bietet einen großen Funktionsumfang und ein ausführliches Online-Benutzerhandbuch. Bei allen Telexperimenten wurden identische Testfragen eingesetzt. Änderungen des Fragebogens können das Antwortverhalten systematisch verzerren und somit den Vergleich zwischen den Telexperimenten erschweren. Auf eine Variation der Formulierungen oder der Reihenfolge wurde daher bewusst verzichtet. Der Fragebogen wurde in drei Abschnitte gegliedert: Im ersten Teil wurden soziodemografische Daten der Teilnehmer wie Alter, Geschlecht und Bildungsabschluss erhoben. Dieser Abschnitt sollte den Teilnehmern einen einfachen Start in den Fragebogen ermöglichen. Außerdem konnte auf diese Weise sichergestellt werden, dass jedem Interview soziodemografische Eckdaten zugeordnet werden können, auch wenn ein Experiment vorzeitig abgebrochen wurde. Im zweiten Teil beantworteten die Teilnehmer Fragen zu ihrer Mediennutzung. Dieser Teil orientierte sich eng an den

Fragen des Digital News Reports. Zunächst sollten die Teilnehmer angeben, welche Geräte (Smartphone, PC, Fernseher, Radio, Zeitung) sie am häufigsten zum Nachrichtenkonsum nutzen. Außerdem wurden sie gefragt, wie sie in der vergangenen Woche auf Online-Nachrichten aufmerksam geworden sind. Schließlich wurden die Probanden am Ende des zweiten Abschnitts noch darum gebeten, für die Medienmarken aus Tabelle 3 eine Bewertung auf einer Skala von 1 bis 5 abzugeben. Damit sollte überprüft werden, ob die Vorüberlegungen, die der Aufteilung in Qualitäts- und Boulevardmarken zu Grunde liegen, mit dem Empfinden der Teilnehmer übereinstimmen. Im dritten Teil wurde das eigentliche Experiment durchgeführt. Nacheinander wurden den Probanden die vorbereiteten Texte und Videos gezeigt. Jede Teilaufgabe bestand aus einem Stimulus und einem Antwortblock. Welche der vier Designvarianten als Stimulus angezeigt wurde, wurde per Zufall bestimmt.

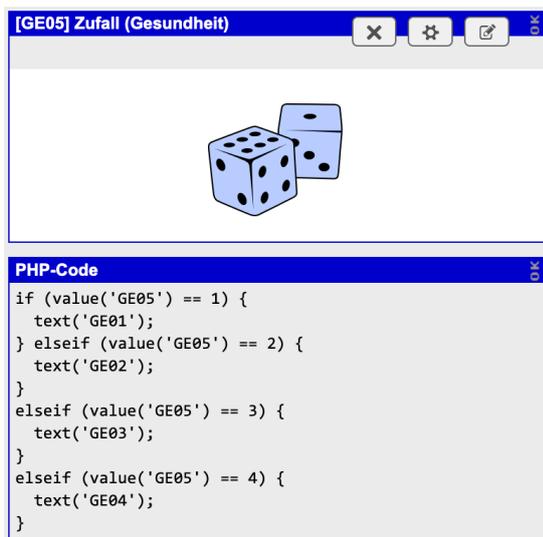


Abbildung 9: Ausschnitt aus dem Programmcode

Abbildung 9 zeigt den Programmcode, mit dem die Auswahl realisiert wurde. Die Ziehung der Zufallsvariablen waren so programmiert, dass eine Gleichverteilung in abgeschlossenen Fragebögen angestrebt wurde. Jeder Beitrag wurde mit einem Einleitungssatz vorgestellt, in dem das Thema und das Veröffentlichungsdatum genannt wurden. Unterhalb davon wurde der Stimulus als Bild oder Videodatei präsentiert, an den sich der Fragenblock anschloss. Die Bearbeitungszeit für den kompletten Fragebogen war auf 20 bis 25 Minuten ausgelegt. Das Programm wurde so eingerichtet, dass die Verweildauer der Teilnehmer auf den jeweiligen Frageseiten automatisch aufgezeichnet wurde.

5.5 Durchführung

Vor dem Start in die Feldphase wurde der vorläufige Fragebogen mit sechs Probanden getestet. Die Software ScoSci-Survey stellt Forschern zu diesem Zweck eine Pretest Funktion bereit. Dabei können die Probanden den Fragebogen unter Originalbedingungen bearbeiten und ihre Änderungsvorschläge über ein zusätzliches Eingabefeld mitteilen. Nach der Auswertung des Pretests wurden zwei Änderungen vorgenommen: Erstens wurde das Layout des Fragebogens für die mobile Nutzung per Smartphone oder Tablet optimiert. Zweitens wurden einige Formulierungen geändert, die zu Verständnisschwierigkeiten geführt hatten. Auf der ersten Seite des Fragebogens wurde den Teilnehmern nach einer kurzen Begrüßung mitgeteilt, dass sie mit der Beantwortung der Fragen zu einem besseren Verständnis des Online-Journalismus beitragen. Die Einführung war bewusst vage formuliert. Über die genaue Forschungsabsicht wurden die Teilnehmer im Unklaren gelassen. Dadurch sollte ein möglichst unvoreingenommenes Antwortverhalten sichergestellt werden. Die Probanden wurden lediglich darum gebeten, die folgenden Texte und Videos aufmerksam zu lesen und anzusehen und anschließend zu bewerten. Die Teilnehmer für das Experiment wurden auf drei Wegen rekrutiert: Über die Umfrageplattformen „pollpool“ und „SurveyCircle“ (1), über Werbung im Freundes- und Bekanntenkreis (2) und über Kontakte zu fünf Gymnasien und Realschulen aus Bayern und Nordrhein-Westfalen (3). Vor Beginn des Experiments wurden mehrere Gymnasial- und Realschullehrer mit der Bitte kontaktiert, den Link zur Umfrage an ihre aktuellen und ehemaligen Schüler weiterzuleiten. Da der Link zum Fragebogen (www.soscisurvey.de/medienkritik2020/) öffentlich zugänglich sein sollte, mussten die Urheber der ausgewählten Texte und Videos vorab um ihr Einverständnis gebeten werden. Nachdem die Rechteinhaber ihr Einverständnis gegeben hatten, begann die Feldphase am 10.06.2020. Der letzte Befragungstag war der 30.06.2020. Eine Woche nach dem Start der Feldphase wurde beim fünften Telexperiment ein Grafikfehler korrigiert, der dazu geführt hatte, dass sich die Abstände zwischen dem Bild und der Unterzeile des „Zeit & klassisch“ Beitrags verschoben hatten. Bei den übrigen 19 Versionen mussten keine Korrekturen vorgenommen werden.

5.6 Datenauswertung und Analyse

In der quantitativen Sozialforschung gibt es seit vielen Jahren eine Debatte darüber, ob Likert-Skalen und Skalen des Likert Typs mit parametrischen Verfahren ausgewertet werden können. Um parametrische Verfahren wie eine Varianzanalyse oder einen t-Test durchführen zu können, müssen vier Voraussetzungen erfüllt sein.

- 1) Die Messwerte müssen voneinander unabhängig sein.
- 2) Die Varianzen der abhängigen Variablen müssen in allen Gruppen homogen sein.
- 3) Die abhängige Variable muss mindestens intervallskaliert sein.
- 4) Die abhängige Variable muss normalverteilt sein.

Skalen des Likert Typs können diese Anforderung formal nicht erfüllen, da bei ordinalverteilten Messwerten die Bildung des arithmetischen Mittelwerts und der Standardabweichung nicht möglich sind. Eigentlich können Skalen des Likert Typs daher nur mit nicht-parametrischen Verfahren analysiert werden. Nicht-parametrische Verfahren verfügen aber über eine geringere Teststärke und können daher Unterschiede in den Daten weniger gut erkennen (Carifio & Perla, 2008, S. 1150). Unter bestimmten Voraussetzungen kann eine Skala des Likert-Typs aber trotzdem wie eine Intervallskala interpretiert werden: Erstens muss die Skala symmetrisch angelegt sein und über einen neutralen Mittelwert verfügen (Symmetrieforderung) und zweitens müssen die Intervalle zwischen allen Antwortoptionen gleich groß sein (Äquidistanzforderung). Dazu muss beispielsweise gezeigt werden, dass zwischen den Werten „sehr niedrig“ und „niedrig“ und „durchschnittlich“ und „eher hoch“ der gleiche Abstand besteht. Um diesen Forderungen so nahe wie möglich zu kommen, wurden die Skalen im Fragebogen symmetrisch ausgelegt. Außerdem wurden die Antwortoptionen mit einer ansteigenden Gerade graphisch unterlegt und um numerische Werte ergänzt. Statistiker und Sozialwissenschaftler diskutieren seit vielen Jahren darüber, ob eine solche Uminterpretation den Einsatz parametrischer Testverfahren rechtfertigen kann. Die amerikanischen Wissenschaftler Carifio und Perla hatten die Debatte im Jahr 2008 für beendet erklärt. In einer vielzitierten Arbeit waren sie zu dem Ergebnis gekommen, dass Mittelwerte, Standardabweichungen und parametrische Verfahren wie die Varianzanalyse problemlos zur Analyse von Likert-Skalen eingesetzt werden können (Carifio & Perla, 2008). Allerdings beziehen sie sich in ihrer Arbeit nur auf echte Likert-Skalen, die aus mehreren Antwortskalen zusammengesetzt werden. Bei der Analyse einzelner Skalen des Likert Typs sei dagegen Vorsicht geboten:

„Analyzing a single Likert item, it should also be noted, is a practice that should occur only rarely” (Carifio & Perla, 2008, S. 1152).

Der Kanadier Geoff Norman geht zwei Jahre später noch einen Schritt weiter als seine amerikanischen Kollegen. In einer Metaanalyse konnte er zeigen, dass parametrische Testverfahren auch bei der Analyse einzelner Skalen des Likert Typs konstant richtige Ergebnisse liefern. Auch bei heterogenen Varianzen, kleinen Stichproben und nicht normalverteilten Größen müsse man „keine Angst“ haben, zu einem falschen Ergebnis zu kommen. Seine Schlussfolgerung lautet:

„One of the beauties of statistical methods is that, although they often involve heroic assumptions about the data, it seems to matter very little even when these are violated“ (Norman, 2010, S. 2).

Auch die niederländischen Wissenschaftler De Winter und Dodou hatten im Jahr 2010 die Daten fünfstufiger Skalen des Likert Typs mit parametrischen Tests untersucht. Mit Hilfe eines Hochleistungsrechners hatten sie 10.000 zufällige Stichproben aus 14 Datensammlungen gezogen. In ihrem Fazit kommen sie zu dem Schluss, dass sowohl der parametrische t-Test als auch der nicht-parametrische Mann-Whitney-U-Test mit gleicher Wahrscheinlichkeit einen statistisch signifikanten Unterschied erkennen können (de Winter & Dodou, 2019, S. 2 ff).

Nach dem Ende der Feldphase am 30.06.2020 wurden die erhobenen Daten mit Hilfe des Statistikprogramms SPSS ausgewertet. Die in Abschnitt 5.1 aufgestellten Hypothesen wurden in zwei Schritten getestet: Die siebenstufigen Skalen des Likert Typs und die zusammengesetzten Likert-Skalen wurden zunächst getrennt untersucht und die Ergebnisse im Anschluss zu einem Gesamtbild zusammengefügt. Da laut der Forschungsliteratur die Bildung von arithmetischen Mittelwerten und Standardabweichungen bei Likert-Daten unproblematisch ist, wurden zur Analyse auch parametrische Testverfahren eingesetzt. Zunächst wurden für jedes Telexperiment die Antworten der siebenstufigen Skalen ausgewertet, auf der die Teilnehmer eine generische Qualitätsbewertung abgegeben hatten. Jeder Antwortoption wurde ein Zahlenwert von 1 bis 7 zugeordnet, wobei die Zahl 1 für den Wert „sehr niedrig“ und die Zahl 7 für den Wert „sehr hoch“ stand. Anschließend wurden die arithmetischen Mittelwerte der vier Teilgruppen mit einer Varianzanalyse (ANOVA) auf Abweichungen untersucht (vgl. Fisher, 1992). Statistisch signifikante Unterschiede zwischen zwei Gruppen wurden im Anschluss mit einem t-Test (vgl. Student, 1908) reevaluiert. Vor jedem t-Tests wurden die Stichproben mit einem Levene-Test (Levene, 1960) auf Varianzhomogenität

überprüft. Um die Wahrscheinlichkeit, einen Fehler erster oder zweiter Art zu begehen, weiter zu minimieren, wurden zusätzliche Mann-Whitney-U-Tests (Mann & Whitney, 1947) durchgeführt. Anschließend wurden für alle signifikanten Ergebnisse die Korrelationskoeffizienten nach Pearson berechnet.

$$r = \left| \sqrt{\frac{t^2}{t^2 + df}} \right| \quad r = \left| \frac{z}{\sqrt{n}} \right|$$

Abbildung 10: Formeln zur Berechnung des r-Werts (links: t-Test, rechts: Mann-Whitney-U-Test)

Um die Größe des Effekts beurteilen zu können wurde die Klassifizierung nach Cohen verwendet (Cohen, 1988, S. 79 f):

0,1 < r < 0,3 → schwacher Effekt

0,3 < r < 0,5 → mittlerer Effekt

0,5 < r → starker Effekt

Für den zweiten Analyseschritt wurden die fünfstufigen Antwortskalen zu Likert-Skalen zusammengesetzt. Dazu wurden die Werte der Teilskalen ungewichtet aufsummiert. Um die Reliabilität der Skalen zu testen wurden anschließend die Cronbachschen Alpha-Werte berechnet (Novick & Lewis, 1967). Dabei zeigte sich, dass sich die Alpha-Werte in allen Telexperimenten erhöhen, wenn die Skalen, mit denen die Verständlichkeit und die ethische Korrektheit der Beiträge abgefragt worden waren, bei der Zusammenstellung unberücksichtigt bleiben.

	mit Verständlichkeit und Ethik	ohne Verständlichkeit und Ethik	Differenz
Teilaufgabe 1	0,816	0,85	0,034
Teilaufgabe 2	0,844	0,865	0,021
Teilaufgabe 3	0,725	0,825	0,1
Teilaufgabe 4	0,809	0,848	0,039
Teilaufgabe 5	0,801	0,826	0,025

Tabelle 6: Werte des Cronbachschen Alphas vor und nach der Neuberechnung

In der Hoffnung, die Reliabilität zu erhöhen, wurden die Likert-Skalen in der Folge neu berechnet. Abweichend von der Theorie wurden nur noch fünf Teilkriterien berücksichtigt: Vielfalt, Relevanz, Unparteilichkeit, Genauigkeit und Glaubwürdigkeit. Die kombinierten Likert-Skalen wurden anschließend mit denselben Methoden untersucht, die auch zur Analyse der siebenstufigen Skalen eingesetzt worden waren.

6 Ergebnisse

6.1 Stichprobe

6.1.1 Alter, Geschlecht und formale Bildung

Das Ziel der Teilnehmerrekrutierung war es, eine möglichst repräsentative Stichprobe der 14- bis 34-jährigen deutschen Wohnbevölkerung zu bekommen. Wie in Abschnitt 3.3 dargelegt wurde, sind die sozialen Netzwerke für die unter 34-Jährigen in Deutschland der wichtigste Zugangsweg zu Online-Nachrichten. Die nach 1990-geborenen Menschen, die häufig als Generationen Y und Z bezeichnet werden, sind die ersten, die vom Kindesalter an mit Smartphones und sozialen Netzwerken konfrontiert waren. Da die meisten unter 34-Jährigen täglich soziale Medien nutzen, ist davon auszugehen, dass viele in dieser Altersgruppe mit der Funktionsweise der Netzwerke vertraut sind. Menschen, die über praktische Erfahrungen verfügen, können in der Regel differenziertere Bewertungen abgeben und sind weniger anfällig für Vorurteile. Um die Größe der Stichprobe und damit auch die Validität der Studie zu erhöhen, wurde der Zugang zum Fragebogen aber nicht auf die unter 30-Jährigen beschränkt. Da im Fragebogen auch das Alter der Teilnehmer erhoben wurde, kann eine Trennung nach Altersgruppen auch im Nachhinein vorgenommen werden. Nach Ende der Befragungszeit wurden in der Rücklaufstatistik insgesamt 231 Datensätze registriert. 119 Datensätze wurden vollständig abgeschlossen. Das entspricht einer Rücklaufquote von 52 Prozent. Je nach Teilaufgabe bewegte sich die Zahl der gültigen Antworten zwischen $N = 158$ und $N = 116$. Das Ziel, eine möglichst junge und onlineaffine Stichprobe zu bekommen, wurde erreicht. Mehr als die Hälfte der Teilnehmer kam aus der Altersgruppe der 15- bis 18-Jährigen. Insgesamt waren 82 Prozent aller Teilnehmer jünger als 35 Jahre.

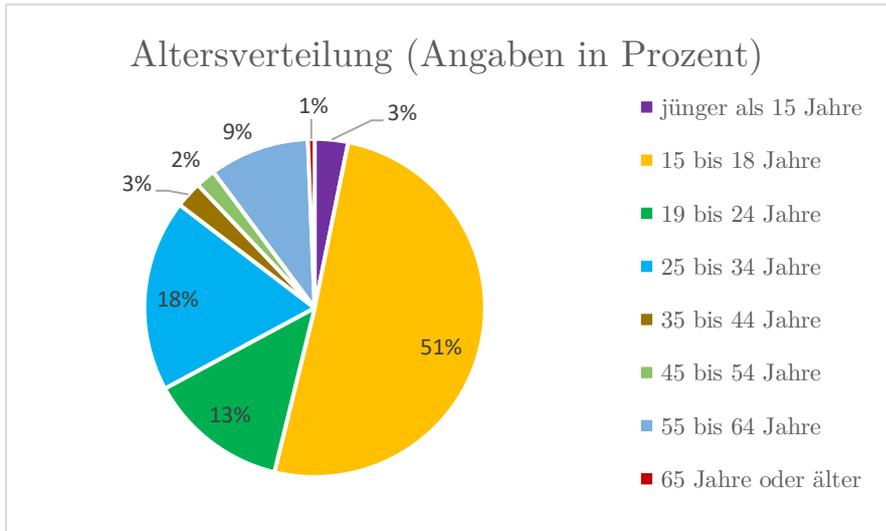


Abbildung 11: Altersverteilung der Stichprobe

Die Geschlechterverteilung der Stichprobe war ausgewogen: 54 Prozent der Befragten waren weiblich und 46 Prozent männlich. Mit Blick auf die formalen Bildungsabschlüsse kann keine eindeutige Aussage über die Repräsentativität getroffen werden, da die Schüler mit 35 Prozent die größte Gruppe unter den Befragten stellen. Da die Zahl der Menschen, die einen Hauptschulabschluss oder eine abgeschlossene Lehre als höchsten Abschluss angeben aber sehr gering ist, ist davon auszugehen, dass die höher gebildeten Teilnehmer zumindest leicht überrepräsentiert sind.

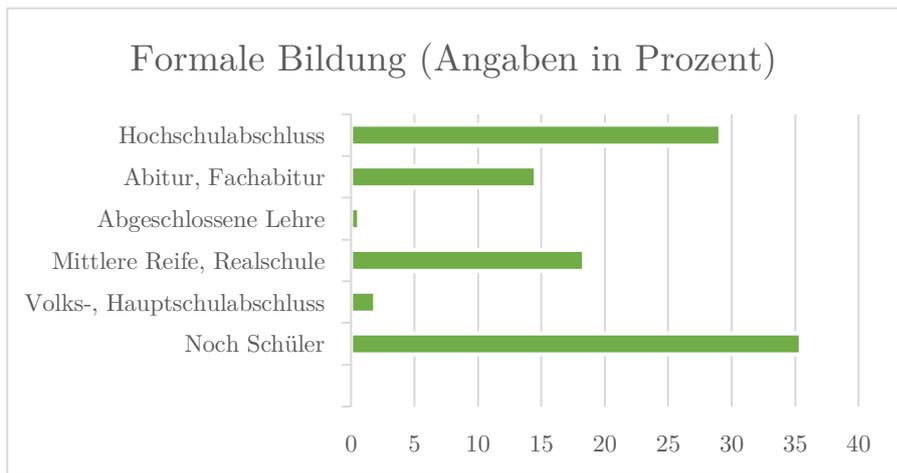


Abbildung 12: Formale Bildung der Stichprobe

6.1.2 Mediennutzungsverhalten

Um die Repräsentativität der Stichprobe besser abschätzen zu können, wurden auch Daten zur Nachrichtennutzung erhoben. Dabei sollte überprüft werden, ob sich die Nutzungsgewohnheiten der Teilnehmer, von denen der Gesamtbevölkerung

unterscheiden. Bei der statistischen Auswertung zeigten sich deutliche Unterschiede zwischen den Altersgruppen. In der Gruppe der unter 35-Jährigen hat das Smartphone eine herausragende Bedeutung. 68,4 Prozent der unter 35-Jährigen erklärten, das Smartphone sei das Gerät, mit dem sie am häufigsten auf Nachrichten zugreifen. Für 90,8 Prozent der unter 35-Jährigen ist das Smartphone zumindest unter den Top drei. Bei den über 35-Jährigen sind Fernseher und Zeitungen die meistverwendeten Geräte/Medien. Das Smartphone ist aber auch für die älteren Teilnehmern bedeutsam und wird von 45 Prozent in die Top drei gewählt. Desktop-Computer, Laptops und Tablets spielen beim Nachrichtenkonsum dieser Altersgruppe dagegen eine untergeordnete Rolle. Insgesamt wurden in der Gruppe der unter 35-Jährigen $N = 98$ und in der Gruppe der über 35-Jährigen $N = 20$ gültige Antworten aufgezeichnet. Die zahlenmäßige Stärke der jüngeren Teilnehmer hat einen deutlichen Einfluss auf die Gesamtstatistik. Das Smartphone führt die Liste der wichtigsten Geräte sowohl in der Einzelbetrachtung der ersten Rangplätze als auch in der Addition der Prozentwerte an. Das Fernsehen belegt in beiden Kategorien den zweiten Platz.

	Smartphone	PC/ Laptop	Tablet	Fernsehen	Radio	Zeitung/ Zeitschrift
unter 35-Jährige						
Rangplatz 1	68,4	9,2	1	11,2	6,1	4,1
Rangplatz 2	14,3	30,6	9,2	26,5	15,3	4,1
Rangplatz 3	8,2	13,3	7,1	27,6	18,4	25,5
Summe	90,8	53,1	17,3	65,3	39,8	33,7
über 35-Jährige						
Rangplatz 1	10	5	5	40	10	30
Rangplatz 2	15	5	5	20	15	40
Rangplatz 3	10	10	5	20	30	25
Summe	45	20	15	80	55	95
alle Altersgruppen						
Rangplatz 1	58	8,4	1,7	16	6,7	9,2
Rangplatz 2	14,3	26,1	8,4	25,2	16	10,1
Rangplatz 3	8,4	12,6	6,7	26,9	20,2	25,2
Summe	80,7	47,1	16,8	68,1	42,9	44,5

Tabelle 7: Wichtigste Geräte zur Nachrichtennutzung (Angaben in Prozent)

Frage: Welche Geräte nutzen Sie am häufigsten, um auf Nachrichten zuzugreifen?

Die zweite Frage zur Mediennutzung beschäftigte sich ausschließlich mit Online-Nachrichten. Die Teilnehmer sollten angeben, auf welche Weise sie in der letzten Woche auf Nachrichten gestoßen sind. Dabei hatten sie die laut Digital News Report sechs wichtigsten Online-Zugangswege zur Auswahl: Newsreader Apps wie Apple News oder Flipboard, Suchmaschinen, Soziale Netzwerke, Push-Benachrichtigungen, Email-Newsletter und den Direktzugriff über eine Webseite oder App (Hölig & Hasebrink, 2020, S. 44). Mehrfachnennungen waren möglich. Die Auswertung der Antworten bestätigte den Trend, der sich schon bei der Gerätenutzung abgezeichnet hatte. Bei Webseiten, Newsreader-Apps und Suchmaschinen war das Nutzungsverhalten in beiden Altersgruppen sehr ähnlich. Signifikante Unterschiede zeigten sich bei der Nutzung von sozialen Medien. In der jüngeren Altersgruppe sind die sozialen Medien der meistverbreitete Zugangsweg. Mehr als 60 Prozent der Befragten waren in der vergangenen Woche über die sozialen Netzwerke auf journalistische Beiträge gestoßen. Bei den über 35-Jährigen traf dies nur auf 20 Prozent der Befragten zu. Bei der Nutzung von Push-Benachrichtigungen zeigte sich ein ähnliches Bild: Während bei den unter 35-Jährigen 31,6 Prozent der Befragten in der vergangenen Woche Push-Nachrichten erhalten hatten, waren es bei den älteren Teilnehmern nur 10 Prozent. Der Email-Newsletter ist der einzige digitale Zugangsweg, der bei den älteren Teilnehmern weiterverbreitet ist.

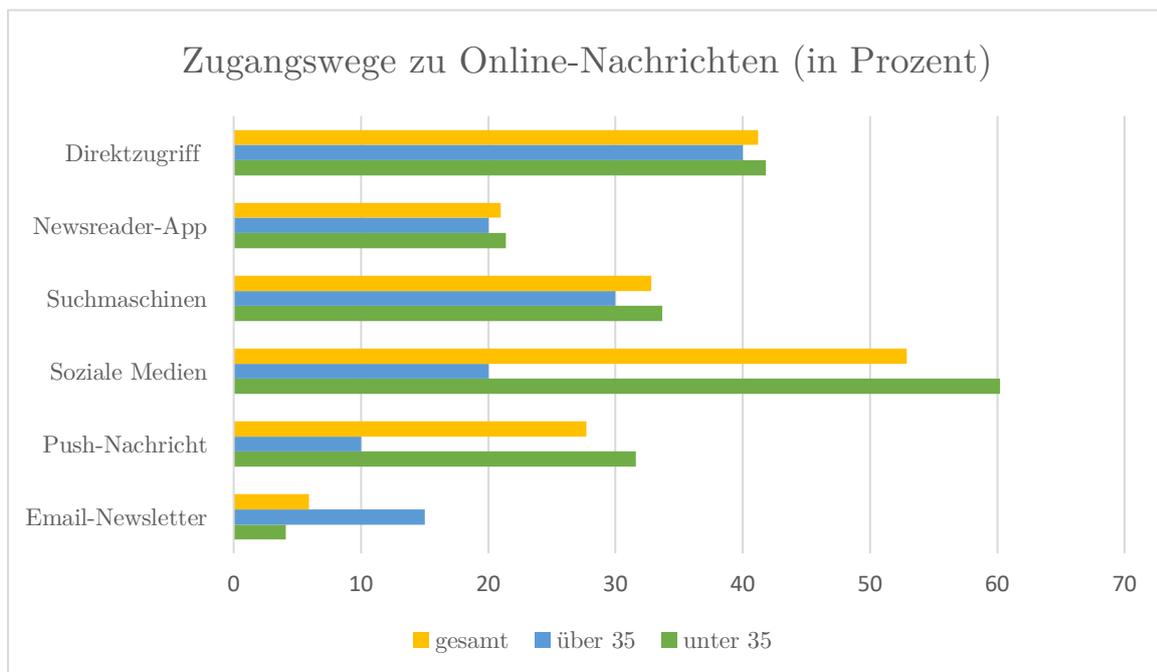


Abbildung 13: Zugangswege zu Online-Nachrichten (Angaben in Prozent)

Frage: Bitte denken Sie nun an die Nachrichten, die Sie vergangene Woche online abgerufen haben. Wie sind Sie auf bestimmte Artikel und Berichte gestoßen?

6.1.3 Vertrauen in Nachrichtenanbieter

Beim Studiendesign wurden je fünf Nachrichtenanbieter mit einem hohen Vertrauenswert und fünf Anbieter mit einem niedrigen Vertrauenswert ausgewählt. Um sicherzugehen, dass die Annahmen mit der Realität übereinstimmen, wurden die Teilnehmer vor Beginn des Experiments gefragt, wie viel Vertrauen sie den gewählten Anbietern entgegenbringen. Die Bewertung erfolgte auf einer Skala mit diskreten Werten von 1 bis 5, wobei die Ziffer 1 die schlechteste und die Ziffer 5 die bestmögliche Bewertung darstellte. Tabelle 8 präsentiert die Ergebnisse der Befragung und ordnet die Anbieter nach ihrer Vertrauenswürdigkeit in absteigender Reihenfolge.

	N	Mittelwert	SD
Süddeutsche Zeitung	108	4,29	0,786
Die Zeit/ Zeit-Online	99	4,14	0,926
Bayerischer Rundfunk	104	3,96	0,934
Der Spiegel	113	3,75	0,996
Apotheken Umschau	85	3,04	1,005
Web.de	82	2,51	0,724
Nachdenkseiten	53	2,34	0,831
Bunte	85	2,25	0,815
BILD	106	2,04	0,975

Tabelle 8: Vertrauen in Nachrichtenmarken

Frage: Wie vertrauenswürdig sind Ihrer Meinung nach die Nachrichten aus folgenden Quellen?

Zusammenfassend lässt sich festhalten, dass die Daten mit den Vorüberlegungen übereinstimmen. Die Anbieter, die beim Studiendesign in die Gruppe der Qualitätsmarken aufgenommen wurden, belegen in der Rangliste die Plätze eins bis fünf. Das höchste Vertrauen genoss die Süddeutsche Zeitung mit einem Durchschnittswert von 4,29. Die Bildzeitung lag mit einem arithmetischen Mittelwert von 2,04 auf dem letzten Platz. Der Nachrichtenanbieter Technikfaultier fehlt in der Liste, da er nur einem sehr kleinen Personenkreis bekannt war.

6.2 Nachrichtenmarken

Die erste Forschungsfrage zielt darauf ab, den Einfluss der Nachrichtenmarke auf das Qualitätsempfinden des Publikums zu untersuchen. In Hypothese 1 wird die Vermutung geäußert, dass die Qualität journalistischer Beiträge, die das Logo eines vertrauenswürdigen Anbieters tragen, überdurchschnittlich gut bewertet wird. Um diese Hypothese zu testen wurden alle Teilnehmer, die die gleiche Version eines Beitrags bearbeitet hatten, in einer Gruppe zusammengefasst. Alle fünf Telexperimente wurden getrennt voneinander ausgewertet. Die Zusammensetzung der vier Versuchsgruppen war bei jedem Telexperiment unterschiedlich. Welche der vier Versionen ein Teilnehmer als Stimulus präsentiert bekommen hatte, war von einem Zufallsgenerator bestimmt worden. Jede Versuchsgruppe bestand aus etwa 30 Personen. Für jede Gruppe wurden zunächst die arithmetischen Mittelwerte der singulären und der kombinierten Qualitätsbewertung berechnet. Dabei zeigte sich ein bemerkenswerter Trend: In 17 von 20 Fällen wurden die Versionen mit dem Logo einer Qualitätsmarke besser bewertet als ihre inhaltlich identischen Pendanten. Tabelle 9 gibt einen Überblick über die Ergebnisse. Bei der Auswertung wurden nur die Bewertungen von Teilnehmern berücksichtigt, die mehr 20 Sekunden auf der entsprechenden Seite des Fragebogens verbracht hatten und mindestens die Hälfte des gesamten Fragebogens bearbeitet hatten. Die Erfahrung aus der Testphase zeigen, dass auf diese Weise die Datenqualität erhöht werden kann.

	Qualitätsmarken			Boulevardmarken			
<i>Telexperiment 1</i>	Apotheken U. & klassisch			Bild & klassisch			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	30	4,40	0,89	27	3,41	1,25	0,99
Qualität (kombiniert)	24	14,50	4,03	23	11,09	3,42	3,41
<i>Telexperiment 1</i>	Apotheken U. & sozial			Bild & sozial			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	31	3,77	1,26	30	3,27	1,11	0,5
Qualität (kombiniert)	28	14,14	4,95	27	11,37	3,52	2,77
<i>Telexperiment 2</i>	SZ & klassisch			Web.de & klassisch			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	29	4,90	1,26	30	4,43	1,43	0,47
Qualität (kombiniert)	26	16,96	4,57	24	16,08	5,09	0,88

	Qualitätsmarken			Boulevardmarken			
<i>Teilexperiment 2</i>	SZ & sozial			Web.de & sozial			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	30	5,23	1,01	28	4,57	1,29	0,66
Qualität (kombiniert)	27	17,56	3,52	22	15,77	4,46	1,79
<i>Teilexperiment 3</i>	Spiegel & klassisch			Technikfaultier & klassisch			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	29	4,31	1,67	29	4,17	1,51	0,14
Qualität (kombiniert)	30	14,97	4,41	23	14,57	4,45	0,4
<i>Teilexperiment 3</i>	Spiegel & sozial			Technikfaultier & sozial			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	30	4,33	1,21	31	3,61	1,41	0,72
Qualität (kombiniert)	27	14,67	3,82	22	13,86	4,92	0,81
<i>Teilexperiment 4</i>	BR24 & klassisch			Bunte & klassisch			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	28	3,93	1,56	29	3,31	1,89	0,62
Qualität (kombiniert)	23	14,39	4,29	29	13,31	4,20	1,08
<i>Teilexperiment 4</i>	BR24 & sozial			Bunte & sozial			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	29	3,41	1,12	30	3,60	1,65	-0,19
Qualität (kombiniert)	21	14,05	3,99	23	13,78	4,56	0,27
<i>Teilexperiment 5</i>	Zeit & klassisch			Nachdenkseiten & klassisch			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	29	4,38	1,40	30	4,47	1,46	-0,09
Qualität (kombiniert)	21	16,86	3,93	24	17,29	4,11	-0,43
<i>Teilexperiment 5</i>	Zeit & sozial			Nachdenkseiten & sozial			
	N	M	SD	N	M	SD	Differenz
Qualität (singulär)	30	4,20	1,42	27	4,00	1,30	0,2
Qualität (kombiniert)	26	16,92	3,25	27	14,44	3,69	2,48

Tabelle 9: Mittelwertvergleich zwischen Qualitäts- und Boulevardmarken

In den ersten drei Teilexperimenten sind alle durchschnittlichen Bewertungen der Beiträge im Design der Qualitätsmarken höher, als die der Beiträge im Design der Boulevardmarken. Auch im vierten Teilexperiment gewinnen die Beiträge der Qualitätsmarken drei der vier Vergleiche zwischen den Gruppen. Die einzige Ausnahme ist das fünfte Teilexperiment: Beim Vergleich der klassischen Verbreitungswege schneidet der Artikel im Design der Nachdenkseiten sowohl bei der singulären als auch bei der

kombinierten Bewertung besser ab als der Artikel im Design von Zeit-Online. Die Abweichung überrascht vor allem deshalb, weil Zeit-Online in der Rangliste der vertrauenswürdigsten Nachrichtenanbieter (siehe Tabelle 8) den zweiten Platz belegt. Allerdings gibt es eine mögliche Erklärung für die Abweichung: Bei der Übertragung der fertig gestalteten Beiträge auf den Server der Umfragesoftware war das Design des Zeit Artikels leicht verzerrt worden. Die Abstände zwischen Unterzeile und Bild waren verkürzt und ein Teil des Texts war in horizontaler Richtung verschoben. Der Graphikfehler blieb in den ersten Tagen unbemerkt und konnte erst nach dem Ende der Feldphase korrigiert werden. Bei der Anzeige der übrigen 19 Versionen hatte es keine Probleme gegeben. Berücksichtigt man bei der Analyse nur die korrekt übertragenen Versionen, so haben die Beiträge im Design der Qualitätsmarken in 17 von 18 Vergleichen besser abgeschnitten haben. Um den Trend einer genaueren Analyse zu unterziehen, wurde zusätzlich für jedes Versuchsgruppenpaar ein t-Test und ein Mann-Whitney-U-Test durchgeführt. Dabei zeigte sich, dass die Unterschiede zwischen den Gruppen in sieben von zwanzig Fällen statistisch signifikant sind (Konfidenzniveau = 95 Prozent). Zwischen den Ergebnissen der t-Tests und der Mann-Whitney-U-Tests gibt es nur minimale Unterschiede. Die Kenngrößen der statistisch signifikanten Tests sind in Tabelle 10 abgebildet. Da Hypothese 1 eine Wirkungsvermutung beinhaltet, werden die einseitigen Signifikanzwerte (p-Werte) angezeigt.

	<i>Apotheken U. klassisch</i>	<i>Bild klassisch</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	4,4	3,41	0,0005	3,48	0,42	0,001	0,42
Qualität (kombiniert)	14,5	11,09	0,0015	3,12	0,42	0,004	0,42
	<i>Apotheken U. sozial</i>	<i>Bild sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	3,77	3,27	0,05	1,668	0,21	0,061	0,20
Qualität (kombiniert)	14,14	11,37	0,01	2,39	0,311	0,033	0,29
	<i>SZ sozial</i>	<i>Web.de sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	5,23	4,57	0,0165	2,19	0,28	0,032	0,28

	<i>Spiegel sozial</i>	<i>Technikfaultier sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	4,33	3,61	0,0185	2,14	0,27	0,027	0,28
	<i>Die Zeit sozial</i>	<i>Nachdenkseiten sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (kombiniert)	16,92	14,44	0,006	2,59	0,34	0,025	0,31

Tabelle 10: Signifikante Unterschiede beim Vergleich der Nachrichtenmarken

Um die Bedeutsamkeit der Ergebnisse besser beurteilen zu können, wurden anschließend die Effektstärken berechnet. Da die r-Werte im Bereich zwischen 0,2 und 0,42 liegen, sind die Effekte nach Cohen als schwach bis mittelstark einzustufen. Mit Ausnahme des vierten Telexperiments gab es bei allen Telexperimenten mindestens ein signifikantes Ergebnis. Auch beim vierten Telexperiment ist der Mittelwert der singulären Qualitätsbewertung in der Versuchsgruppe „BR24 & klassisch“ um 0,62 Punkte höher als in der Gruppe „Bunte & klassisch“. Allerdings liegt der errechnete p-Wert beim einseitigen t-Test mit 0,09 über dem Signifikanzniveau von 0,05. Fünf der signifikanten Ergebnisse stammen aus einem Vergleich von Versuchsgruppen, die im Stil der sozialen Medien gestaltet worden waren. Dies deutet darauf hin, dass der Zugangsweg „soziale Medien“ möglicherweise den positiven oder negativen Einfluss einer Medienmarke verstärkt. Allerdings könnte es sich dabei auch um eine zufällige Beobachtung handeln. Insgesamt lässt sich festhalten: Wenn zwischen zwei Medienmarken ein ausreichend großer Vertrauensunterschied besteht, wird das Qualitätsempfinden des Publikums messbar beeinflusst. Im Extremfall kann die Medienmarke einen mittelstarken Einfluss auf die Qualitätswahrnehmung haben. Die Hypothese, dass Beiträge von Medienmarken, die ein hohes Vertrauen genießen, im Durchschnitt besser bewertet werden, wird auf Grundlage der Daten angenommen.

6.3 Zugangswege

Die zweite Forschungsfrage beschäftigt sich damit, ob der Weg, über den eine Nachricht verbreitet wird, die Qualitätswahrnehmung beeinflusst. In Hypothese 2 wird die Vermutung geäußert, dass Nachrichten, die in den sozialen Netzwerken verbreitet werden, im Durchschnitt schlechter bewertet werden als solche, die ausschließlich über die Webseite eines Nachrichtenanbieters veröffentlicht werden. Um diese Hypothese zu

testen wurden zunächst wieder die arithmetischen Mittelwerte der Versuchsgruppen gebildet und anschließend auf Abweichungen untersucht. Dabei wurden die Versuchsgruppen „Qualität & klassisch“ mit den Gruppen „Qualität und sozial“ und die Versuchsgruppen „Boulevard & klassisch“ mit den Gruppen „Boulevard und sozial“ verglichen. Beim Vergleich der Mittelwerte konnte diesmal aber keine eindeutige Tendenz ausgemacht werden. In zwölf von 20 Fällen waren die Mittelwerte des klassischen Verbreitungswegs höher als die Mittelwerte des Verbreitungswegs „soziale Medien“. In acht Fällen waren die Verhältnisse umgekehrt. Um weitere Informationen zu erhalten wurden die neuen Versuchsgruppenpaare wieder mit t-Tests und Mann-Whitney-U-Tests auf Abweichungen untersucht. Bei der Analyse zeigte sich, dass die Unterschiede zwischen den Mittelwerten in zwei Fällen statistisch signifikant sind (siehe Tabelle 11). Die Effektstärken nach Cohen bewegen sich dabei im schwachen Bereich.

	<i>Apotheken U. klassisch</i>	<i>Apotheke sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	4,4	3,77	0,0145	2,233	0,28	0,0265	0,25
	<i>Nachdenkseiten klassisch</i>	<i>Nachdenkseiten sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (kombiniert)	17,29	14,44	0,006	2,605	0,34	0,0045	0,37

Tabelle 11: Signifikante Unterschiede beim Vergleich der Zugangswege

Insgesamt betrachtet lässt sich bei den Bewertungen eine leicht positive Tendenz zugunsten der Beiträge im Design des klassischen Verbreitungswegs beobachten. Die Zahl der statistisch signifikanten Ergebnisse ist aber relativ gering. Auf Grundlage der erhobenen Daten kann die Hypothese 2 weder angenommen noch verworfen werden. Um zu testen, ob es einen Zusammenhang zwischen den Nachrichtennutzungsgewohnheiten und der Qualitätswahrnehmung gibt, wurden zusätzlich einige Quertests durchgeführt. Für den ersten Quertest wurde die gesamte Stichprobe in zwei Gruppen aufgeteilt. Die erste Gruppe bildeten die Teilnehmer, die das Smartphone bei der Nachrichtennutzung auf den ersten Platz gewählt hatten (N = 69). Die übrigen Teilnehmer kamen in die zweite Gruppe (N = 50). Anschließend wurden die obenstehenden Analyseschritte für beide Gruppen wiederholt. In Bezug auf die Zugangswege konnten dabei keine neuen Erkenntnisse gewonnen werden. Allerdings zeigte sich ein interessanter Trend: Bei allen Teilerperimenten waren die durchschnittlichen singulären und kombinierten

Qualitätsbewertungen der ersten Gruppe höher als die der zweiten Gruppe. Die Abweichungen waren zwar nicht statistisch signifikant, aber die Tendenz war eindeutig. Ob zwischen der Smartphone-Nutzung und der Strenge der Bewertung tatsächlich ein Zusammenhang besteht, kann aber nur gemutmaßt werden.

		Smartphone auf Platz 1	Nicht auf Platz 1	
		M	M	Differenz
Experiment 1	<i>singulär</i>	3,71	3,71	0
	<i>kombiniert</i>	13,72	12,03	1,69
Experiment 2	<i>singulär</i>	4,82	4,77	0,05
	<i>kombiniert</i>	17,19	15,82	1,37
Experiment 3	<i>singulär</i>	4,26	3,88	0,38
	<i>kombiniert</i>	15,03	13,8	1,23
Experiment 4	<i>singulär</i>	3,71	3,28	0,43
	<i>kombiniert</i>	14,35	12,95	1,4
Experiment 5	<i>singulär</i>	4,28	4,25	0,03
	<i>kombiniert</i>	16,66	15,79	0,87

Tabelle 12: Vergleich zwischen den Gruppen "Smartphone auf Platz 1" und "nicht auf Platz 1"

Auch für den nächsten Quertest wurde die Stichprobe in zwei Gruppen aufgeteilt. Die erste Gruppe wurde aus den Teilnehmern gebildet, die bei der Befragung angegeben hatten, in der vergangenen Woche über die sozialen Netzwerke auf Nachrichten gestoßen zu sein (N = 63). Die übrigen Teilnehmer bildeten die zweite Gruppe (N = 56). Diese Aufteilung beruhte auf der Überlegung, dass Menschen, die regelmäßig soziale Medien zum Nachrichtenkonsum nutzen, möglicherweise eine andere Wahrnehmung dieses Zugangswegs haben. Die Analyse der Mittelwerte ergab ein ähnliches Bild wie der erste Quertest: Die Teilnehmer aus der soziale-Medien-Gruppe hatten bei allen Teilerperimenten im Durchschnitt eine höhere Bewertung abgegeben. Um zu testen, ob es sich bei den Smartphone- und den soziale-Medien-Nutzern um dieselben Teilnehmer handelt, wurde die Schnittmenge zwischen den Gruppen gebildet. Es zeigte sich, dass die Zusammensetzungen der Gruppen unterschiedlich genug sind, um eine getrennte Analyse zu rechtfertigen: 63 Prozent der Teilnehmer aus der Smartphone Gruppe waren auch in der soziale-Medien-Gruppe vertreten. Die im nächsten Schritt durchgeführten t-Tests lieferten sehr interessante Ergebnisse: In der ersten Gruppe war nur noch einer von zwanzig Vergleichen statistisch signifikant.

"Soziale Medien" ausgewählt	<i>Apotheken U. klassisch</i>	<i>Apotheken U. sozial</i>	t-Test			Mann-Whitney- U	
	M	M	p	t	r	p	r
Qualität (singulär)	4,8	3,95	0,0165	2,22	0,36	0,023	0,33

Tabelle 13: Signifikante Unterschiede beim Vergleich der Zugangswege (Quertest A)

Bei den Menschen, die in der Woche vor dem Test keine Nachrichten über die sozialen Netzwerke abgerufen hatten, erhöhte sich die Zahl der signifikanten Ergebnisse dagegen auf vier. Bei allen signifikanten Ergebnissen wurde die Version im Design des klassischen Verbreitungswegs besser bewertet, als die inhaltlich identische Version im Design eines sozialen Netzwerks.

"Soziale Medien" nicht ausgewählt	<i>Apotheken U. klassisch</i>	<i>Apotheken U. sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	4	3,22	0,03	1,99	0,40	0,043	0,36
Qualität (kombiniert)	13,64	10,38	0,0415	1,84	0,41	0,057	0,36
	<i>Technikfaultier klassisch</i>	<i>Technikfaultier sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	4,38	3,6	0,034	1,906	0,35	0,059	0,30
	<i>Nachdenkseiten klassisch</i>	<i>Nachdenkseiten sozial</i>	t-Test			Mann-Whitney-U	
	M	M	p	t	r	p	r
Qualität (singulär)	17,44	12,5	0,0045	2,96	0,58	0,0045	0,61

Tabelle 14: Signifikante Unterschiede beim Vergleich der Zugangswege (Quertest B)

Die Ergebnisse deuten darauf hin, dass der Zugangsweg tatsächlich einen messbaren Einfluss auf die Qualitätswahrnehmung zu hat. Wie stark der Einfluss ist, scheint zu einem großen Teil vom Nachrichtennutzungsverhalten abhängig zu sein. Interessant ist, dass alle statistisch signifikanten Ergebnisse (Konfidenzniveau = 0,95) von den Telexperimenten 1, 3 und 5 stammen. Die Beiträge dieser drei Experimente wurden im Design unterschiedlicher sozialer Netzwerke gestaltet (Experiment 1: YouTube, Experiment 3: Instagram, Experiment 5: Facebook). Daher liegt die Vermutung nahe, dass es keine Rolle spielt, in welchem sozialen Netzwerk ein Beitrag veröffentlicht wurde.

7 Fazit

7.1 Die Nachrichtenmarke: eine bedeutende Heuristik

Das Ziel dieser Arbeit war es herauszufinden, ob die Nachrichtenmarke und der Zugangsweg die Wahrnehmung journalistischer Beiträge beeinflussen. Die Ergebnisse des Experiments deuten stark darauf hin, dass ein solcher Einfluss existiert. Insbesondere die Nachrichtenmarke scheint für viele Menschen eine bedeutende Heuristik zu sein. In vier von fünf Telexperimenten konnten bei der singulären oder der kombinierten Bewertung statistisch signifikante Unterschiede zwischen mindestens zwei Versuchsgruppen ausgemacht werden. Dabei wurden jedes Mal die Versionen im Design derjenigen Marke besser bewertet, die über einen höheren Vertrauenswert verfügte. Dies gilt auch für das dritte Telexperiment, in dem eine unbekannte Marke (Technikfaultier) mit einer Qualitätsmarke verglichen wurde. Diese Ergebnisse decken sich mit den Erkenntnissen der psychologischen und betriebswirtschaftlichen Forschung der vergangenen Jahrzehnte. Viele Forscher gehen davon aus, dass bei der Güterbewertung ein Imagetransfer von der Marke auf das Produkt stattfindet. Insofern ist es wenig überraschend, dass dieser Imagetransfer auch bei journalistischen Erzeugnissen beobachtet werden kann. Diese Erkenntnisse bestätigen auch die Beobachtung von Urban und Schweiger, die im Jahr 2013 den Einfluss einer Nachrichtenmarke auf die Wahrnehmung untersucht hatten. Die Ergebnisse des Experiments zeigten eine eindeutige Tendenz: In 17 von 20 Fällen waren die durchschnittlichen Bewertungen bei den Qualitätsmarken höher als bei den Boulevardmarken. In sieben dieser Fälle waren die Unterschiede statistisch signifikant. Dabei schien es keine Rolle zu spielen, ob es sich bei dem Beitrag um einen Text oder ein Video handelte. Interessanterweise wurden sechs der sieben signifikanten Ergebnisse beim Vergleich von Versuchsgruppen aus den Kategorien „Qualität & sozial“ und „Boulevard & sozial“ ermittelt. Dies überrascht vor allem deshalb, weil das Design der sozialen Netzwerke streng genormt ist und beim Layout oder der Schriftart keine Abweichungen zulässt. Eigentlich könnte man annehmen, dass diese erzwungene Ähnlichkeit dazu führt, dass die Bewertungsunterschiede zwischen den Versuchsgruppen kleiner werden. Allerdings scheint das genaue Gegenteil der Fall zu sein. Offenbar bewerten die Menschen Beiträge von Marken, die nur wenig Vertrauen genießen, kritischer, wenn diese über die sozialen Netzwerke verbreitet werden. Natürlich wäre es auch denkbar, dass die Beiträge angesehener Medienmarken in den sozialen Netzwerken

einen positiven Imageschub erhalten. Da die sozialen Netzwerke aber regelmäßig für negative Schlagzeilen sorgen und für zahlreiche gesellschaftliche Probleme verantwortlich gemacht werden, erscheint dieser Erklärungsansatz wenig plausibel. Auf Grundlage der erhobenen Daten kann die Hypothese angenommen werden, dass Nachrichtenbeiträge von Anbietern, die ein hohes Vertrauen genießen, besser bewertet werden als solche, die von Anbietern mit einem niedrigen Vertrauenswert stammen. Ob ein linearer Zusammenhang zwischen der Qualitätswahrnehmung und dem Vertrauen in die Marke existiert, und wie groß der Vertrauensunterschied sein muss, damit ein Einfluss erkennbar wird, konnte im Rahmen dieser Studie allerdings nicht beantwortet werden. Dafür waren die Ergebnisse zu wenig einheitlich und die gemessenen Vertrauensunterschiede zwischen den Medienmarken zu ähnlich. Diese Studie konnte aber zeigen, dass es höchst wahrscheinlich einen Einfluss gibt und in welche Richtung er wirkt. Um die Wirkmechanismen im Detail zu analysieren, bedarf es zukünftiger Untersuchungen.

7.2 Nutzungsverhalten prägt die Wahrnehmung

Das Experiment lieferte auch zur Beantwortung der zweiten Forschungsfrage interessante Erkenntnisse. Die Daten deuten darauf hin, dass auch der Verbreitungsweg die Qualitätswahrnehmung beeinflusst. Die Hinweise auf einen möglichen Zusammenhang sind aber nicht eindeutig. Die t-Tests zwischen den Versuchsgruppen lieferten in nur zwei von 20 Fällen statistisch signifikante Ergebnisse. Um die zweite Hypothese annehmen zu können bedarf es weiterer Daten. Die beobachtete Tendenz reicht aber aus, um die Vermutung aufrecht zu erhalten, dass die sozialen Medien die Wahrnehmung von Nachrichtenbeiträgen negativ beeinflussen. Facebook, YouTube oder Instagram stehen seit Jahren in der Kritik und werden häufig mit Hetze und Falschmeldungen in Verbindung gebracht. Einige Wissenschaftler vertreten sogar die These, dass sie maßgeblich zum Auseinanderdriften der westlichen Gesellschaften beigetragen haben. Gleichwohl sind die sozialen Netzwerke bei den Nutzern enorm beliebt und sind für die unter 35-Jährigen heute der wichtigste Zugangsweg zu Nachrichten. Die Detailanalyse der erhobenen Daten deutet darauf hin, dass das Nutzungsverhalten der Teilnehmer im Zusammenhang mit der Qualitätswahrnehmung steht. Für die Versuchsteilnehmer, die in der Woche vor dem Experiment Nachrichten über die sozialen Medien konsumiert hatten, spielte der Zugangsweg bei der Bewertung nur eine untergeordnete Rolle. Die

übrigen Teilnehmer bewerteten die Beiträge im Design der Sozialen Medien dagegen in vier Fällen signifikant schlechter. Für die Ergebnisse gibt es zwei unterschiedliche Erklärungsansätze: Erstens könnte die Wahrnehmung der Menschen umso weniger von negativen Vorurteilen beeinflusst werden, je häufiger sie sich selbst in den sozialen Medien bewegen. Für diese Theorie spricht, dass Vorurteile häufig auf Unkenntnis beruhen und bei regelmäßigem Umgang mit einer Sache durch eigene Erfahrungen ersetzt werden. Eine zweite mögliche Erklärung ist, dass diejenigen, die in der Vergangenheit negative Erfahrungen mit den sozialen Medien gemacht haben, heute andere Zugangswege nutzen. Dafür spräche die Tatsache, dass jüngere Menschen den Großteil der soziale-Medien-Nutzer ausmachen und wegen ihres geringen Alters schlicht weniger Zeit hatten, negative Erfahrungen zu machen. Auch wenn der Einfluss des Verbreitungswegs auf die Qualitätswahrnehmung nicht eindeutig belegt werden kann, liefert diese Arbeit dennoch vielversprechende Hinweise. Um die Vermutung bestätigen zu können, dass die Mediennutzung mit der Bewertung der sozialen Netzwerke in Zusammenhang steht, bedarf es aber weiterer Studien, die sich noch intensiver mit dem Nutzungsverhalten der Teilnehmer auseinandersetzen. In dieser Arbeit wurde eine junge und onlineaffine Stichprobe untersucht. Obwohl die Vermutung naheliegt, dass die Nutzungsgewohnheiten der jüngeren Generationen, künftig die gesamte Gesellschaft prägen, müssen andere Studien noch den Beweis erbringen, dass sich die Ergebnisse auch bei älteren und weniger gebildeten Menschen replizieren lassen. Die Wahrscheinlichkeit ist aber groß, dass die beobachteten Effekte sogar noch größer ausfallen könnten, da weniger gebildete Menschen in der Regel auch eine geringere Medienkompetenz besitzen und daher leichter von externen Faktoren wie der Nachrichtenmarke oder dem Verbreitungsweg beeinflusst werden könnten.

7.3 Limitationen

Abschließend muss darauf hingewiesen werden, dass die Ergebnisse dieser Studie nicht losgelöst vom Versuchskontext interpretiert werden sollten. Auch wenn beim Studiendesign viel Mühe darauf verwandt wurde, die Medienmarken und die Zugangswege möglichst realitätsnah abzubilden, könnten einigen Teilnehmer die Manipulationen aufgefallen sein.

Außerdem muss berücksichtigt werden, dass das Experiment durch einen strengen Versuchsplan bestimmt war. Die Versuchsteilnehmer bewegten sich in einer künstlich geschaffenen Umgebung und könnten daher von ihrem normalen Rezeptionsverhalten abgewichen sein. Es ist möglich, dass die Teilnehmer die Beiträge wesentlich aufmerksamer gelesen/ angesehen haben als üblich, weil sie im Experiment explizit dazu aufgefordert wurden. Allerdings ist auch ein gegenteiliger Effekt denkbar: Die Teilnehmer werden während des Experiments zur Auseinandersetzung mit den Beiträgen gezwungen und entschieden sich nicht freiwillig für den Konsum. Daher könnten sie die Beiträge auch weniger gründlich rezipiert haben. Der Transfer der Ergebnisse auf alltägliche Nutzungssituationen sollte daher mit Einschränkungen erfolgen. Diese Arbeit beschäftigt sich mit der Qualitätswahrnehmung inhaltlich identischer Beiträge. Welchen Einfluss der Nachrichteninhalt auf die Wahrnehmung hat, wurde nicht untersucht. Zusätzlich muss bedacht werden, dass der Zugangsweg soziale Medien möglicherweise nicht alleine durch die Kopie des Designs von Facebook, YouTube oder Instagram nachgestellt werden kann. Die sozialen Netzwerke zeichnen sich insbesondere dadurch aus, dass sie ein personalisiertes Nutzererlebnis bieten. Algorithmen entscheiden auf Grundlage des vergangen Nutzungsverhaltens darüber, welche Inhalte ein Nutzer zu sehen bekommt. Die Vermutung liegt nahe, dass die individualisierte Umgebung zumindest einen Teil des Nutzererlebnisses ausmacht. Im Rahmen dieser Studie wurden die Teilnehmer aber mit denselben Beiträgen konfrontiert, da personalisierte Beiträge einen präzisen Vergleich zwischen den Versuchsgruppen unmöglich gemacht hätten.

I Tabellenverzeichnis

Tabelle 1: Kategorisierung der Online-Zugangswege.....	17
Tabelle 2: Versuchsgruppen des 2x2 Experiments	30
Tabelle 3: Übersicht über die ausgewählten Medienmarken	31
Tabelle 4: Monatsreichweiten wichtiger sozialer Netzwerke (vgl. Beisch et al., 2019, S. 383)	32
Tabelle 5: Teildimensionen des gewählten Qualitätsmodells.....	39
Tabelle 6: Werte des Cronbachschen Alphas vor und nach der Neuberechnung	45
Tabelle 7: Wichtigste Geräte zur Nachrichtennutzung (Angaben in Prozent) Frage: Welche Geräte nutzen Sie am häufigsten, um auf Nachrichten zuzugreifen?	49
Tabelle 8: Vertrauen in Nachrichtenmarken Frage: Wie vertrauenswürdig sind Ihrer Meinung nach die Nachrichten aus folgenden Quellen?.....	51
Tabelle 9: Mittelwertvergleich zwischen Qualitäts- und Boulevardmarken	53
Tabelle 10: Signifikante Unterschiede beim Vergleich der Nachrichtenmarken	55
Tabelle 11: Signifikante Unterschiede beim Vergleich der Zugangswege.....	56
Tabelle 12: Vergleich zwischen den Gruppen "Smartphone auf Platz 1" und "nicht auf Platz 1" ..	57
Tabelle 13: Signifikante Unterschiede beim Vergleich der Zugangswege (Quertest A)	58
Tabelle 14: Signifikante Unterschiede beim Vergleich der Zugangswege (Quertest B)	58

II Abbildungsverzeichnis

Abbildung 1: Algorithmenbasierte und Nicht-algorithmenbasierte Zugangswege (Hölig & Hasebrink, S.45)	18
Abbildung 2: Überblick über die Beitragsversionen des ersten Teilexperiments.....	33
Abbildung 3: Überblick über die Beitragsversionen des zweiten Teilexperiments	34
Abbildung 4: Überblick über die Beitragsversionen des dritten Teilexperiments	35
Abbildung 5: Überblick über die Beitragsversionen des vierten Teilexperiments	36
Abbildung 6: Überblick über die Beitragsversionen des fünften Teilexperiments.....	37
Abbildung 7: Siebenstufige Antwortskala des Likert Typs.....	39
Abbildung 8: Fünfstufige Antwortskala.....	40
Abbildung 9: Ausschnitt aus dem Programmcode.....	41
Abbildung 10: Formeln zur Berechnung des r-Werts (links: t-Test, rechts: Mann-Whitney-U-Test)45	
Abbildung 11: Altersverteilung der Stichprobe.....	48
Abbildung 12: Formale Bildung der Stichprobe	48
Abbildung 13: Zugangswege zu Online-Nachrichten (Angaben in Prozent) Frage: Bitte denken Sie nun an die Nachrichten, die Sie vergangene Woche online abgerufen haben. Wie sind Sie auf bestimmte Artikel und Berichte gestoßen?	50

III Literaturverzeichnis

- Aaker, D. A. (1996). Building strong brands. Free Press.
- Aaker, D. A., & Keller, K. L. (1990). Consumer Evaluations of Brand Extensions. *Journal of Marketing*, 54(1), 27–41. <https://doi.org/10.1177/002224299005400102>
- Arnold, K. (2008). Qualität im Journalismus – ein integratives Konzept. *Publizistik*, 53(4), 488–508. <https://doi.org/10.1007/PL00022233>
- Arnold, K. (2009). Qualitätsjournalismus: Die Zeitung und ihr Publikum. UVK Verlagsgesellschaft.
- Beck, K., Schubert, C., & Reineck, D. (2010). Journalistische Qualität in der Wirtschaftskrise. UVK Verlagsgesellschaft.
- Beisch, V. N., Koch, W., & Schäfer, C. (2019). ARD/ZDF-Onlinestudie 2019: Mediale Internetnutzung und Video-on- Demand gewinnen weiter an Bedeutung. 15.
- Benson, P. (2017). The discourse of YouTube: Multimodal text in a global context. Routledge, Taylor & Francis Group.
- Bilke, N. (2008). Qualität in der Krisen- und Kriegsberichterstattung: Ein Modell für einen konflikt sensitiven Journalismus (1. Aufl). VS Verlag für Sozialwissenschaften.
- Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, 42(12), 1150–1152. <https://doi.org/10.1111/j.1365-2923.2008.03172.x>
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed). L. Erlbaum Associates.
- Dahinden, U., Kaminski, P., & Niederreuther, R. (2004). Content is King?: Qualitätsbeurteilung aus Angebots- vs. Rezipientenperspektive. In K. Beck, W. Schweiger, & W. Wirth (Hrsg.), *Gute Seiten, schlechte Seiten: Qualität in der Onlinekommunikation* (S. 103–126). Verlag Reinhard Fischer.
- de Winter, J. F. C., & Dodou, D. (2019). Five-Point Likert Items: T test versus Mann-Whitney-Wilcoxon (Addendum added October 2012). *Practical Assessment, Research, and Evaluation*, 15(1). <https://doi.org/10.7275/bj1p-ts64>
- Dürr, A., & Lüpke-Narberhaus, F. (o. J.). *Neue Perspektiven für Ihren Alltag*. <https://www.spiegel.de/backstage/backstage-floete-leben-ressort-a-e8ff90f4-6f1e-467f-a01b-e2bf2c1c7c73>
- Eisenegger, M., Hauser, L., Schneider, J., Ingenhoff, D., & Bachmann, P. (2018). *Schwizer Medienqualitätsranking 2018*. <https://www.mqr-schweiz.ch/files/mqr/pdf/MQR-18.pdf>
- Fengler, S. (2016). Journalismus als rationales Handeln. In M. Löffelholz & L. Rothenberger (Hrsg.), *Handbuch Journalismustheorien* (S. 235–248). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-18966-6_13
- Fiedler, C. (2013). Nachrichtenqualität aus Rezipientensicht: Eine qualitative Studie. Bachelorarbeit: Technische Universität Ilmenau.

Fisher, R. A. (1992). Statistical Methods for Research Workers. In S. Kotz & N. L. Johnson (Hrsg.), *Breakthroughs in Statistics* (S. 66–70). Springer New York. https://doi.org/10.1007/978-1-4612-4380-9_6

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80(S1), 298–320. <https://doi.org/10.1093/poq/nfw006>

Galesic, M., & Bosnjak, M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *Public Opinion Quarterly*, 73(2), 349–360. <https://doi.org/10.1093/poq/nfp031>

Galton, F. (1907). *Vox Populi*. *Nature*, 75(1949), 450–451. <https://doi.org/10.1038/075450a0>

Geschke, D., Lorenz, J., & Holtz, P. (2019). The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, 58(1), 129–149. <https://doi.org/10.1111/bjso.12286>

Gladney, G. A. (1996). How Editors and Readers Rank and Rate the Importance of Eighteen Traditional Standards of Newspaper Excellence. *Journalism & Mass Communication Quarterly*, 73(2), 319–331. <https://doi.org/10.1177/107769909607300204>

Groshek, J., & Koc-Michalska, K. (2017). Helping populism win? Social media use, filter bubbles, and support for populist presidential candidates in the 2016 US election campaign. *Information, Communication & Society*, 20(9), 1389–1407. <https://doi.org/10.1080/1369118X.2017.1329334>

Grundgesetz, (1949). <https://www.gesetze-im-internet.de/gg/BJNR000010949.html>

Gscheidle, V. C., & Geese, S. (2017). Die Informationsqualität der Fernsehnachrichten aus Zuschauersicht. 15.

Guo, L., A. Rohde, J., & Wu, H. D. (2020). Who is responsible for Twitter’s echo chamber problem? Evidence from 2016 U.S. election networks. *Information, Communication & Society*, 23(2), 234–251. <https://doi.org/10.1080/1369118X.2018.1499793>

Haas, H., & Lojka, K. (1998). Qualität auf dem Prüfstand. In W. Duchkowitsch, F. Hausjell, W. Hömberg, A. Kutsch, & I. Neverla (Hrsg.), *Journalismus als Kultur* (S. 115–132). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-87316-3_10

Habermas, J. (2006). Political Communication in Media Society: Does Democracy Still Enjoy an Epistemic Dimension? The Impact of Normative Theory on Empirical Research. *Communication Theory*, 16(4), 411–426. <https://doi.org/10.1111/j.1468-2885.2006.00280.x>

Hermes, S. (2006). *Qualitätsmanagement in Nachrichtenredaktionen*. Halem.

Herzig, B., & Martin, A. (2017). Erfassung und Messbarkeit von Medienkompetenz als wichtige Voraussetzung für politische Bildung. *Schriftenreihe Medienkompetenz*, 126–135.

Hörger, M. (2010). Participant Dropout as a Function of Survey Length in Internet-Mediated University Studies: Implications for Study Design and Voluntary Participation in Psychological Research. *Cyberpsychology, Behavior, and Social Networking*, 13(6), 697–700. <https://doi.org/10.1089/cyber.2009.0445>

Hölig, S., & Hasebrink, U. (2020). Reuters Institute Digital News Report 2020: Ergebnisse für Deutschland.

https://leibniz-hbi.de/uploads/media/default/cms/media/66q2yde_AP50_RIDNR20_Deutschland.pdf

Hölig, S., Hasebrink, U., & Hans-Bredow-Institut. (2019). Reuters Institute Digital News Report 2019: Ergebnisse für Deutschland.

Jungnickel, K. (2011). Nachrichtenqualität aus Nutzersicht. Ein Vergleich zwischen Leserurteilen und wissenschaftlich-normativen Qualitätsansprüchen. *Medien & Kommunikationswissenschaft*, 59(3), 360–378.
<https://doi.org/10.5771/1615-634x-2011-3-360>

Kreutner, G. (2019, September 23). Apotheken Umschau: Bestnoten für Glaubwürdigkeit und Verständlichkeit.
<https://www.presseportal.de/pm/52139/4382211>

Lacy, S. (2000). Commitment of financial resources as a measure of quality. In R. G. Picard (Hrsg.), *Measuring media content, quality, and diversity: Approaches and issues in content research* (S. 25–50). Media Economics, Content and Diversity Project and Media Group.

Lacy, S., & Rosenstiel, T. (2015). Defining and Measuring Quality Journalism. 1–66.

Lee, A. M., & Chyi, H. I. (2015). The Rise of Online News Aggregators: Consumption and Competition. *International Journal on Media Management*, 17(1), 3–24. <https://doi.org/10.1080/14241277.2014.997383>

Leidenberger, J. (2015). Boulevardisierung von Fernsehnachrichten: Eine Inhaltsanalyse deutscher und französischer Hauptnachrichtensendungen. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-658-08094-5>

Levene, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, 278–292.

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), 50–60.
<https://doi.org/10.1214/aoms/1177730491>

McNair, B. (2012). Trust, truth and objectivity: Sustaining quality journalism in the era of the content generating user. Bd. *Rethinking Journalism: Trust and Participation in a Transformed News Landscape* (C. Peters & M. J. Broersma, Hrsg.). Taylor & Francis Group. <http://ebookcentral.proquest.com/lib/ub-lmu/detail.action?docID=1092745>

McQuail, D. (1992). *Media performance: Mass communication and the public interest*. Sage Publications.

Meyen, M., & Jandura, O. (2011). Das Publikum der Qualitätsmedien. Eine repräsentative Studie zu Reichweite, sozialer Verortung und Nutzungsmotiven. In R. Blum, H. Bonfadelli, K. Imhof, & O. Jarren (Hrsg.), *Krise der Leuchttürme öffentlicher Kommunikation* (S. 207–220). VS Verlag für Sozialwissenschaften.
https://doi.org/10.1007/978-3-531-93084-8_13

Neuberger, C., Nuernbergk, C., & Rischke, M. (Hrsg.). (2009). *Journalismus im Internet: Profession – Partizipation – Technisierung*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-91562-3>

Newman, N. (2019). Reuters Institute Digital News Report 2019. 156.

- Newman, N., Fletcher, R., Levy, D. A., & Nielsen, R. K. (2016). Digital News Report 2016. <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/research/files/Digital%2520News%2520Report%25202016.pdf>
- Nielsen, H. (2020). Bereinigter Werbetrend 2019. <https://www.vau.net/werbeumsaetze/content/brutto-werbeumsaetze-2019>
- Nohr, H. (2001). Management der Informationsqualität. In W.-F. Riekert & M. Michelson (Hrsg.), *Informationswirtschaft* (S. 57–77). Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-322-87395-8_4
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>
- Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, 32(1), 1–13. <https://doi.org/10.1007/BF02289400>
- Oberauer, J. (2020). Datenbank der Journalistenpreise in Deutschland. <https://journalistenpreise.de/>
- Peters, C., & Broersma, M. J. (Hrsg.). (2013). *Rethinking journalism: Trust and participation in a transformed news landscape*. Routledge.
- Porst, R. (2014). Fragebogen. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-02118-4>
- Pöttker, H. (2000). Kompensation von Komplexität—Journalismustheorie als Begründung journalistischer Qualitätsmaßstäbe. In M. Löffelholz (Hrsg.), *Theorien des Journalismus* (S. 375–390). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-97091-6_19
- Prinzing, M. (2016). Qualität im Journalismus. In J. Heesen (Hrsg.), *Handbuch Medien- und Informationsethik* (S. 106–113). J.B. Metzler. https://doi.org/10.1007/978-3-476-05394-7_14
- Purcell, K., Mitchel, L., Rosenstil, T., & Olmstead, K. (2010). *Understanding the participatory news consumer*. Pew Research Center.
- R. Pennington, Nancy A. Pachana, Sa, H. (2001). USE OF THE FACTS ON AGING QUIZ IN NEW ZEALAND: VALIDATION OF QUESTIONS, PERFORMANCE OF A STUDENT SAMPLE, AND EFFECTS OF A DON'T KNOW OPTION. *Educational Gerontology*, 27(5), 409–416. <https://doi.org/10.1080/03601270152053438>
- Rager, G. (1994). Dimensionen der Qualität Weg aus den allseitig offenen Richter-Skalen? Bd. *Publizistik in der Gesellschaft* (G. Bentele & K. R. Hesse, Hrsg.; S. 198–209).
- Reinemann, C., Fawzi, N., & Obermaier, M. (2017). Die „Vertrauenskrise“ der Medien – Fakt oder Fiktion? Zu Entwicklung, Stand und Ursachen des Medienvertrauens in Deutschland. (S. 77–94).
- Ruggiero, T. E. (2000). Uses and Gratifications Theory in the 21st Century. *Mass Communication and Society*, 3(1), 3–37. https://doi.org/10.1207/S15327825MCS0301_02
- Ruß-Mohl, S. (1992). Am eigenen Schopfe... Qualitätssicherung im Journalismus—Grundfragen, Ansätze, Näherungsversuche. *Publizistik*, 37(1), 83–96.

- Schatz, H., & Schulz, W. (1992). Qualität von Fernsehprogrammen: Kriterien und Methoden zur Beurteilung von Programmqualität im dualen Fernsehsystem. *Media Perspektiven*, 690–712.
- Schwaderer, H. (2020). D21 Digitalindex 2019/2020.
https://initiated21.de/app/uploads/2020/02/d21_index2019_2020.pdf
- Shoemaker, P. J., & Reese, S. D. (1996). *Mediating the message: Theories of influences on mass media content* (2nd ed). Longman.
- Spohr, D. (2017). Fake news and ideological polarization: Filter bubbles and selective exposure on social media. *Business Information Review*, 34(3), 150–160. <https://doi.org/10.1177/0266382117722446>
- Spurk, C., & Lublinski, J. (o. J.). *Content Analysis: Measuring the Success of Journalism Capacity Building*. *Content Analysis*, 8.
- Struckmann, S., & Karnowski, V. (2016). News consumption in a changing media ecology: An MESM-study on mobile news. *Telematics and Informatics*, 33(2), 309–319. <https://doi.org/10.1016/j.tele.2015.08.012>
- Student. (1908). The Probable Error of a Mean. *Biometrika*, 6(1), 1. <https://doi.org/10.2307/2331554>
- Urban, J., & Schweiger, W. (2013). News Quality from the Recipients' Perspective: Investigating recipients' ability to judge the normative quality of news. *Journalism Studies*, 15(6), 821–840.
<https://doi.org/10.1080/1461670X.2013.856670>
- Uwe Hasebrink. (2011). Giving the Audience a Voice: The Role of Research in Making Media Regulation More Responsive to the Needs of the Audience. *Journal of Information Policy*, 1, 321.
<https://doi.org/10.5325/jinfopoli.1.2011.0321>
- Vlasic, A. (2004). Über Geschmack lässt sich nicht streiten – über Qualität schon? Zum Problem der Definition von Maßstäben für publizistische Qualität. In K. Beck, W. Schweiger, & W. Wirth (Hrsg.), *Gute Seiten, schlechte Seiten: Qualität in der Onlinekommunikation* (S. 15–31). Verlag Reinhard Fischer.
- Willnata, L., & Weaver, D. H. (1998). Public Opinion on Investigative Reporting in the 1990s: Has Anything Changed since the 1980s? *Journalism & Mass Communication Quarterly*, 75(3), 449–463.
<https://doi.org/10.1177/107769909807500302>
- Wolling, J. (2002). Aufmerksamkeit durch Qualität? Empirische Befunde zum Verhältnis von Nachrichtenqualität und Nachrichtennutzung. https://www.db-thueringen.de/receive/dbt_mods_00027019