

# Nachrichtenqualität aus Nutzersicht

*Ein Vergleich zwischen Leserurteilen und wissenschaftlich-normativen Qualitätsansprüchen*

Katrin Jungnickel

*Kosteneinsparungen, Konkurrenzdruck und der Zwang, möglichst aktuell und exklusiv über die neuesten Ereignisse und Entwicklungen zu berichten, machen es Redaktionen nicht immer leicht, journalistische Qualitätsstandards einzuhalten. Dieser Entwicklung entgegenzuwirken, obliegt auch dem Publikum, das durch eine Bevorzugung von Nachrichten hoher Qualität deren Nachfrage steigern kann. Allerdings ist unklar, ob Leser die gleichen Qualitätsansprüche haben, die auch aus wissenschaftlich-normativer Sicht an Nachrichtenbeiträge gestellt werden, und ob sie zu ähnlichen Qualitätsurteilen kommen. Dieser Frage geht die Studie mit Hilfe eines Mehrmethodendesigns nach. Zunächst wurden bei einer Inhaltsanalyse Artikel guter und schlechter Qualität identifiziert. In einem zweiten Schritt bewerteten Nutzer diese Beiträge im Rahmen eines Online-Experiments. Untersucht wurde, inwieweit die Bewertungen in Abhängigkeit verschiedener Beitrags-eigenschaften (Qualität und Themenbereich) sowie verschiedener Rezipientenmerkmale (Bildung, Mediennutzung, Themeninteresse) voneinander abwichen. Die Ergebnisse zeigen, dass Mediennutzer Qualitätsunterschiede durchaus erkennen können. Außerdem erwiesen sich der Themenbereich und das themenspezifische Interesse als wichtige Einflussfaktoren auf die Bewertung der präsentierten Nachrichtenartikel.*

**Schlüsselwörter:** Nachrichtenqualität, Nutzerqualität, Journalismus, Online-Experiment, Themeninteresse

## 1. Einleitung

Der technische Fortschritt und die damit verbundene rasante Entwicklung neuer Medien und potenzieller Informationsquellen versetzen den traditionellen Nachrichtenjournalismus derzeit in eine Phase des Wandels. Vor allem das Internet schwingt sich mit Onlinezeitungen, Blogs und seit Neuestem auch mit Twitter zu einem der stärksten Konkurrenten der gedruckten Zeitungen auf. Gesteigerter Aktualitäts- und Exklusivitätszwang und wirtschaftliche Schwierigkeiten aufgrund sinkender Verkaufszahlen bestimmen vielerorts den Redaktionsalltag. Personalmangel und die zum Teil auftretende Doppelbelastung, sowohl die Print- als auch die Onlineausgaben bedienen zu müssen, erschweren die Einhaltung wichtiger journalistischer Qualitätsstandards (vgl. Beck et al. 2010). Bei dem Versuch, Kosten zu reduzieren, werden immer weniger Redakteure fest angestellt oder die Dienste von Nachrichtenagenturen abbestellt (vgl. Siegert 2009: 14), in der Hoffnung, die Leser würden entsprechende Einsparungen nicht bemerken.

Aus demokratietheoretischer Perspektive erwartet man jedoch, dass das Publikum langfristig Nachrichtenmedien hoher Qualität bevorzugt und damit einem möglichen Qualitätsverlust entgegenwirkt. Diese Hoffnung stützt sich auf die Annahme, dass Rezipienten mangelnde Qualität in der Berichterstattung auch erkennen. Ob das tatsächlich so ist, ist allerdings in der kommunikationswissenschaftlichen Qualitätsforschung bisher noch nicht ausreichend überprüft worden. Mögliche Gründe dafür nennt Heinrich (1999: 39ff.), der darauf hinweist, dass Mediennutzer durch einen Mangel an Informationen vor dem Kauf eines Medienprodukts dessen Qualität nicht beurteilen können.

Medien sind jedoch auch „Erfahrungsgüter“ (Beck et al. 2010: 32): Wer schon oft Nachrichtenbeiträge einer bestimmten Zeitung gelesen hat, kann abschätzen, was ihn in der nächsten Ausgabe erwartet. Sinkt die Qualität einer Zeitung kontinuierlich, ist eine Abwanderung der Leserschaft, vor allem der Abonnenten, denkbar. Langfristig gesehen kann somit das Qualitätsbewusstsein der Rezipienten, das sich aus den Rezeptionserfahrungen mit einzelnen Nachrichtenbeiträgen eines Mediums speist, den Erfolg dieses Mediums beeinflussen. In diesem Zusammenhang gewinnt die Frage an Bedeutung, wie Rezipienten einzelne Nachrichtenbeiträge beurteilen und ob sie dabei mit den Ansprüchen übereinstimmen, die wissenschaftlich-normativ an die Berichterstattung der Medien gestellt werden. Antworten darauf soll die im vorliegenden Beitrag vorgestellte empirische Studie liefern.

## 2. Forschungsstand

### 2.1 Kriterien der Nachrichtenqualität

Angestoßen durch den Qualitätsdiskurs in den USA (u. a. McQuail 1992) hat sich die kommunikationswissenschaftliche Forschung in Deutschland vor allem zu Beginn der 1990er-Jahre vermehrt mit der Definition und Erforschung journalistischer Qualität beschäftigt. Dabei stellte man schnell fest, dass Nachrichtenqualität als Konstrukt von vielen verschiedenen Faktoren abhängt. Welche Faktoren dazu aus wissenschaftlicher Perspektive gehören, wurde in zahlreichen Kriterienkatalogen festgehalten (u. a. Ruß-Mohl 1992, Schatz / Schulz 1992, Hagen 1995; später Pöttker 2000, Rager 2000, Wyss 2002, Arnold 2009). Parallel dazu lieferten Regelungen des Grundgesetzes und der Landespressegesetze sowie die Statuten des deutschen Pressekodex Anhaltspunkte für die Anforderungen an eine qualitätsvolle Berichterstattung.

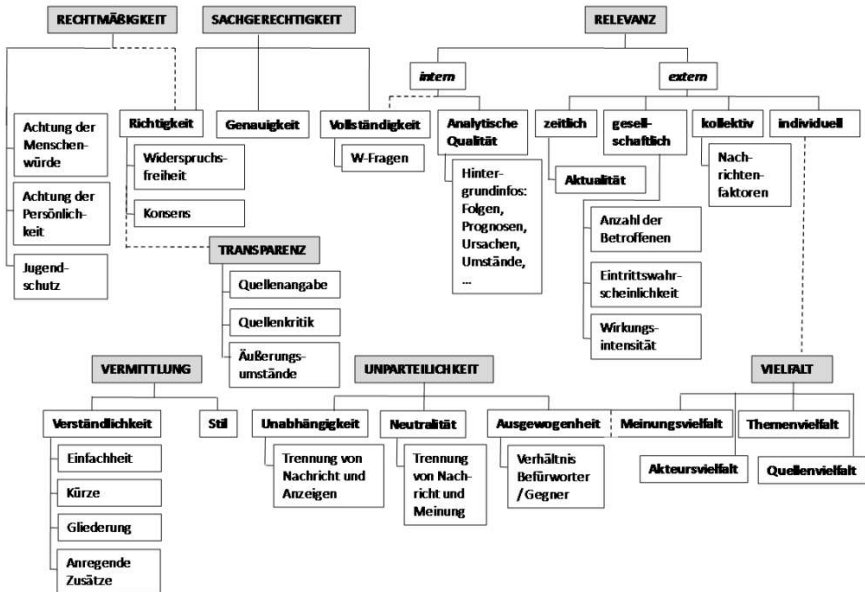
Insgesamt lassen sich daraus sieben Qualitätsdimensionen ableiten, die so oder ähnlich benannt in den meisten Kriterienkatalogen vorkommen: *Rechtmäßigkeit*, *Sachgerechtigkeit*, *Relevanz*, *Vermittlung*, *Transparenz*, *Unparteilichkeit* und *Vielfalt*. Abbildung 1 gibt einen Überblick über diese sieben Dimensionen und deren Unterkriterien sowie über empirisch messbare Indikatoren, durch die sie sich beschreiben lassen.

Die Systematik stellt dabei in erster Linie Aspekte der Qualität von Nachrichten dar. Strukturelle Faktoren, die sich eher auf das Mediensystem als Ganzes beziehen (wie z. B. Außenpluralismus als Teil der Vielfalt), werden hier ebenso wenig berücksichtigt wie Kriterien für andere journalistische Stilformen. Die sieben Dimensionen und deren Unterkriterien dienen für viele Studien als Grundlage, um journalistische Produkte verschiedener Mediengattungen im Hinblick auf ihre Qualität inhaltsanalytisch zu untersuchen: Im Bereich der Printmedien überprüfte z. B. Hagen (1995) die Qualität von Nachrichtenagenturmeldungen, während bei Fahr (2001) der Fokus auf der Berichterstattung über ein einzelnes Ereignis lag.

### 2.2 Nachrichtenqualität aus Nutzersicht

Allerdings genügt es nicht, die Qualität von Nachrichten oder generell von Medieninhalten nur daran zu messen, was aus wissenschaftlicher oder demokratietheoretischer Sicht wünschenswert wäre. Damit sie von ihrer Zielgruppe, den Nutzern, akzeptiert wird, muss die Berichterstattung auch deren Ansprüchen genügen. Grundsätzlich kann das Publikum bei der Rezeption und Bewertung von Medieninhalten verschiedene Rollen einnehmen (vgl. Hasebrink 2008: 515ff.): Erstens befriedigen die Nutzer als mündige *Konsumenten* individuelle kognitive, affektive oder soziale Bedürfnisse. Zweitens be-

Abbildung 1: Systematisierung der Qualitätsdimensionen



Eigene Darstellung in Anlehnung an Schatz & Schulz (1992), Wyss (2002), Maurer & Reinemann (2006) und Daschmann (2009).

nötigen sie als *schutzbedürftige Individuen* gesetzliche Regelungen u. a. zum Jugend- und Persönlichkeitsschutz, die sie vor möglichen negativen Medienwirkungen bewahren sollen. Drittens nehmen sie als *Bürger* auch eine gesellschaftsbezogene Perspektive ein, aus der sie Medienangebote nach ihrer Relevanz für die Allgemeinheit und nach ihrer Funktionalität für die öffentliche Meinungsbildung beurteilen können. Klassische Qualitätsmodelle (u. a. Schatz / Schulz 1992) berücksichtigen die Nutzer oftmals lediglich als Konsumenten und fragen nach der Akzeptanz von Medieninhalten beim Publikum. Dass das Publikum jedoch auch die Bürgerrolle einnimmt und dabei ähnliche Qualitätsansprüche hat, denen aus normativer Perspektive Bedeutung beigemessen wird, zeigen u. a. die Ergebnisse einer Studie von Arnold (2009). Er ermittelte faktorenanalytisch fünf Hauptdimensionen der Nutzerqualität: *Präsentation* (= Vermittlung, Gestaltung), *Nähe* (= Relevanz, Themeninteresse), *Orientierung* (= Themenvielfalt, Aktualität), *Engagement* (= Recherche, Quellentransparenz) und *Fairness* (= Neutralität, Ausgewogenheit, Achtung der Persönlichkeit). Einige dieser Dimensionen beziehen sich vor allem auf die Ansprüche des Publikums als Konsumenten (Präsentation, Nähe), während andere sich an wissenschaftlich-normativen Kriterien orientieren, die ein Nutzer in der Rolle des Bürgers erwartet (Orientierung, Engagement, Fairness). Diese zwei verschiedenen Perspektiven können dabei auch im Widerspruch zueinander stehen, nämlich dann, wenn ein Nutzer Medieninhalte zwar als Bürger qualitativ hochwertig einschätzt, sie ihm aber als Konsument nicht gefallen, da sie seinen individuellen Bedürfnissen nicht entsprechen. Zusätzlich bedeuten gleiche Erwartungen in der Rolle des Bürgers noch

lange nicht, dass das Publikum bei der Qualitätseinschätzung auch zu den gleichen Ergebnissen wie Experten und Wissenschaftler kommt, da den Rezipienten dafür möglicherweise entsprechende Kompetenzen fehlen (vgl. Daschmann 2009: 259).

Bisher ist das Verhältnis zwischen Gefallen, Qualitätsbewertungen der Nutzer und normativen Qualitätsurteilen wenig untersucht worden. Studien, die sich der Nachrichtenqualität aus Nutzersicht gewidmet haben, erfassten diese vor allem mit Hilfe von Befragungen. Grob lassen sich die Untersuchungen in drei Gruppen einteilen:

1. *Bewertung von konkreten Medienprodukten oder Mediengattungen* (z. B. Langzeitstudie Massenkommunikation, vgl. Reitze / Ridder 2006, oder Studien zur Qualitätsbeurteilung von Nachrichten auf öffentlich-rechtlichen bzw. privaten Fernsehsendern, u. a. Zubayr / Geese 2009);
2. *Wichtigkeitsrankings verschiedener Kriterien aus Nutzersicht* (u. a. erhoben bei Trepte et al. 2008, Arnold 2009) und
3. *Untersuchungen zu allgemeinen Qualitätsansprüchen und Kritikpunkten an Journalisten* (z. B. Donsbach et al. 2009, Arnold 2009, Beck et al. 2010<sup>1</sup>).

Die bisherigen Studien ließen jedoch einige Fragen noch offen: Bei der Bewertung von Medienprodukten konnten Glaubwürdigkeit und langfristig aufgebaute Images die Qualitätsbeurteilung verzerren, und Wichtigkeitsrankings sagen noch nichts darüber aus, wie Leser die Qualität im Einzelfall wirklich einschätzen<sup>2</sup>. Außerdem beschäftigen sie sich ausschließlich mit der Nutzerperspektive. Neuere Untersuchungen gingen noch einen Schritt weiter. Sie haben mit einer Kombination aus Befragungen und Inhaltsanalysen versucht, Gemeinsamkeiten und Unterschiede zwischen normativen Qualitätsansprüchen und Nutzerqualitätsbewertungen festzustellen. So verglich Wolling (2006) die Erwartungen von Radiohörern an die Qualität von Rundfunkprogrammen mit dem tatsächlichen Inhalt und konnte zeigen, dass ein Programm dann häufiger genutzt wird, wenn es den Erwartungen der Hörer entspricht. Außerdem untersuchten Dahinden et al. (2004) und Rössler (2004) Onlinezeitungen inhaltsanalytisch und ließen ihre Versuchspersonen diese nach einiger Zeit der freien Exploration in Bezug auf verschiedene klassische und webspezifische Qualitätskriterien bewerten. Allerdings wurde in der Untersuchung nicht genau erfasst, welche Artikel die Nutzer eigentlich gelesen hatten, daher konnten die Forscher auch keine genauen Aussagen machen, wie die Qualitätsbewertungen zustande kamen. Grund dafür ist die unterschiedliche Analyseebene: Die Inhaltsanalyse erfolgte auf Beitrags-, die Befragung hingegen auf Medienproduktebene. Diese gleiche Ebene versuchten Berdi et al. (1992) herzustellen, die einen qualitativ schlechten, einen mittelmäßigen und einen guten Artikel von etwa 100 Lesern und 30 Redakteuren bewerten ließen. Insgesamt war die Übereinstimmung der Befragten-Urteile mit der normativ bestimmten Qualität der drei Artikel eher gering, bei den Journalisten jedoch erwartungsgemäß höher als bei den Lesern (vgl. ebd.: 187). Ihre Notengebung begründeten die Leser vor allem mit den Dimensionen Relevanz, Präsentation und Objektivität; weniger kümmerten sie sich beispielsweise darum, ob der Artikel gut recherchiert war. Diese Studie war unter den Untersuchungen lange Zeit die einzige, die sich mit Nutzerqualitätsbewertungen auf der Ebene einzelner Artikel beschäftigte. Al-

1 Die Autoren befragten dabei „professionelle Leser“, also Wirtschaftsakteure und Vertreter von politischen, kulturellen oder gesellschaftlichen Institutionen.

2 Dies führte auch zu sehr unterschiedlichen Rangplätzen für die einzelnen Kriterien: So rangiert die Aktualität als Teil der Relevanz einer Nachricht in der Befragung von Trepte und Kollegen (2008) beispielsweise hinter Kriterien wie Richtigkeit, Objektivität, Verständlichkeit und Vielfalt, während sie bei Arnold (2009) nach der Kürze das wichtigste Qualitätskriterium aus Nutzersicht darstellt.

lerdings traten hier auch methodische Probleme auf, die vor allem den nur dichotom skalierten Variablen, der ungleichen Gewichtung der Qualitätskriterien und den wenig differenzierten Bewertungsmöglichkeiten vonseiten der Versuchspersonen geschuldet waren. Neben Berdi und Kollegen haben bisher im Printbereich nur Trepte et al. (2008), deren Studie sich jedoch ausschließlich auf ethische Qualitätsaspekte beschränkt, ebenfalls Nutzerbewertungen auf Beitragsebene untersucht.

Anders ist die Situation in der Rundfunkforschung. Im Rahmen des Programmcontrollings beim WDR bewerten bereits seit 1998 regelmäßig ausgewählte Zuschauer Sendungen im Hinblick auf ihre Qualität. Deren Einschätzung wird durch Expertenmeinungen ergänzt, wobei diese im Verlauf der Zeit immer seltener genutzt wurden, da die Kritik der Zuschauer oft schon präzise genug war (vgl. Tebert 2010: 81f.). Ein direkter Vergleich mit wissenschaftlich normativen Ansprüchen fand hierbei jedoch nicht statt.

### 2.3 Einflussfaktoren auf Nutzerqualitätsbewertungen

Zentrale Fragen zum Verhältnis von normativer Qualität und Nutzerqualität konnten bisherige Vergleichsstudien nicht oder nur teilweise beantworten. Gründe dafür sind vor allem die unterschiedlichen Analyseeinheiten (Beitragsebene in Inhaltsanalysen, Medienproduktebene bei Befragungen) und die Vernachlässigung intervenierender Angebotsseigenschaften wie Glaubwürdigkeit und Medienimage.

Außerdem wurden *Rezipienteneigenschaften* bisher seltener beachtet (Ausnahmen sind u. a. Arnold 2009 sowie Reitze / Ridder 2006), obwohl verschiedene Ansätze insbesondere aus dem Bereich der Mediensozialisationsforschung, der Sozialpsychologie und der Wissensklufforschung Anhaltspunkte dafür bieten. Darunter ist auch das Konzept der *Medienkompetenz* zu zählen, definiert als „die Fähigkeit, Medien kritisch, selbstbestimmt und verantwortlich nutzen, verstehen, bewerten und gestalten zu können“ (Trepte 2008: 102). Groeben (2004) unterscheidet insgesamt sieben Kompetenzen, darunter auch Medienwissen und medienbezogene Kritikfähigkeit. In Verbindung mit Nutzerqualitätsbewertungen erfüllt das Konzept der Medienkompetenz zwei Funktionen. Zum einen lässt sich vermuten, dass Aspekte wie Medienwissen die Qualitätseinschätzungen beeinflussen; zum anderen sind die Nutzerurteile selbst möglicherweise ein Indikator für medienbezogene Kritikfähigkeit. Bisher ist dies kaum empirisch untersucht worden, vor allem weil es an geeigneten Kriterien zur Operationalisierung fehlte (vgl. Trepte et al. 2008).

Es ist nicht unproblematisch, festlegen zu wollen, wer dazu fähig ist, Medien angemessen zu kritisieren, und wer nicht. Allerdings könnte die Übereinstimmung mit Bewertungen, die inhaltsanalytisch anhand wissenschaftlicher Kriterien vorgenommen wurden, darauf hinweisen, in welchem Maße ein Leser über medienbezogene Kritikfähigkeit verfügt. Wie stark diese bei Nutzern ausgeprägt ist, hängt potenziell von mehreren Faktoren ab. Zunächst liegt die Vermutung nahe, dass die Häufigkeit der *Mediennutzung*, die Anzahl und Art der genutzten Medien sowie die rezipierten Inhalte dabei eine Rolle spielen. So berichtet beispielsweise Bonfadelli (2004: 250) in seinem Forschungsüberblick über eine positive Korrelation zwischen Printmediennutzung und Informationsstand, während die Zusammenhänge zwischen Rezeption und Wissen bei den übrigen Medien generell eher schwach ausgeprägt sind.

Ebenso kann die *Bildung* eine Rolle bei der Qualitätsbeurteilung spielen. Hinweise darauf liefert z. B. die Wissenskluffhypothese (vgl. Tichenor et al. 1970), die postuliert, dass Statushöhere und Gebildetere in der Informationsgesellschaft ihr Wissen stärker ausbauen können als Statusniedrigere. Empirische Studien konnten die Wissenskluffhypothese nicht immer bestätigen, da intervenierende Variablen wie Motivation, Alter,

Geschlecht und habitualisierte Mediennutzung den Effekt zum Teil deutlich abschwächen (vgl. Wirth 1997: 80). Dennoch zeigt die bisherige Forschung, dass die Schulbildung vor allem bei der Entwicklung von integrativem Wissen und Interpretationsfähigkeit eine Rolle spielt. Daraus lässt sich schlussfolgern, dass medienbezogene Kritikfähigkeit bei höher Gebildeten stärker ausgeprägt sein sollte.

Als weitere Disposition mit möglichem Einfluss auf die Qualitätseinschätzung ist das *Interesse* für das Beitragsthema zu werten und damit verbunden die Motivation, sich mit der jeweiligen Nachricht genauer zu beschäftigen. Petty und Cacioppos (1986) Elaboration Likelihood Model (ELM) besagt, dass Informationen entweder komplex und rational-kognitiv auf der zentralen Route verarbeitet werden oder eher affektiv bzw. mit Hilfe von Heuristiken auf der peripheren Route (ebd.: 128f.). Ausschlaggebend für die Routenwahl sind dabei insbesondere die Fähigkeiten (z. B. Vorwissen) und das Involvement (persönliche Relevanz, Interesse) der betreffenden Person. Ähnliche Annahmen machen das heuristisch-systematische Modell von Chaiken (1980) und Brosius' (1995) Modell der Alltagsrationalität. Die unterschiedlichen Verarbeitungsmodi könnten dazu führen, dass interessierte Nutzer, die eine Nachricht zentral verarbeiten, möglicherweise auf andere Qualitätsaspekte achten als unmotivierte Rezipienten. Gerade Nachrichten, die komplizierte Themen beinhalten, könnten mit einer alltagsrationalen Verarbeitung schnell als unverständlich eingestuft werden, während sie für hoch involvierte Rezipienten logisch erscheinen.

Neben Rezipienteneigenschaften können auch *Eigenschaften des Beitragsthemas* eine Rolle bei der Qualitätsbeurteilung spielen. So liegt die Vermutung nahe, dass bestimmte Themenbereiche die einzelnen Publikumsrollen verschieden stark ansprechen. Während politische Themen vom Mediennutzer möglicherweise eher in seiner Rolle als Bürger beurteilt werden, steht bei anderen Themen, z. B. Sport, die Konsumentenperspektive klar im Vordergrund. Schließlich ist die Sportberichterstattung nicht nur reine Information, sondern beinhaltet immer auch Unterhaltungselemente (vgl. Maurer / Reine mann 2006: 216). Demnach ist zu erwarten, dass das Publikum an Artikel über Politik andere Ansprüche stellt als an Artikel über Sport.

### 3. Studiendesign und Forschungsfragen

Um die Korrespondenz zwischen Nutzerbewertungen und wissenschaftlich-normativen Qualitätsansprüchen bei Textnachrichten in Abhängigkeit von Rezipienteneigenschaften zu messen, war die Entwicklung eines Mehrmethodendesigns nötig, bei dem beide Perspektiven der Qualitätsbeurteilung auf der gleichen Analyseebene und am gleichen Ausgangsmaterial untersucht werden. Dafür wurden zunächst die wissenschaftlich-normativen Kriterien operationalisiert, um mit Hilfe einer Inhaltsanalyse die Qualität einzelner Nachrichtenbeiträge zu messen. Anschließend lasen Versuchspersonen in einem Online-Experiment einige der vorher inhaltsanalytisch untersuchten Artikel und bewerteten deren Qualität. Die Daten der Inhaltsanalyse und des Experiments konnten dann auf Gemeinsamkeiten und Unterschiede in der Qualitätseinschätzung überprüft werden.

Von Interesse war zunächst, wie Mediennutzer zu einem Gesamturteil über einen Beitrag kommen. Untersucht werden sollte dabei das Verhältnis zwischen Konsumenten- und Bürgerrolle, also inwieweit die Qualitätseinschätzung einen Einfluss darauf hat, ob Lesern der Beitrag gefällt.

*Forschungsfrage F1: Gefällt den Lesern ein Beitrag besser, wenn sie dessen Qualität höher einschätzen?*

Im Anschluss wurde geprüft, ob die Nutzerbewertungen insgesamt und bei den einzelnen Kriterien Ähnlichkeiten zur inhaltsanalytisch ermittelten Qualität der Beiträge aufwiesen:

*Forschungsfrage F2: Bewerten Nutzer einen inhaltsanalytisch als qualitativ voll identifizierten Beitrag positiver als einen weniger qualitativ vollen Beitrag?*

Ein weiterer Aspekt, der in diesem Zusammenhang Berücksichtigung fand, ist der Themenbereich<sup>3</sup> des Artikels als Beitragseigenschaft neben seiner inhaltsanalytisch ermittelten Qualität. Im Experiment wurden zwei gegensätzliche Themenbereiche gewählt: Politik als demokratietheoretisch besonders relevanter Bereich und Sport als Bereich, der eher der Unterhaltung zugeordnet werden kann.

*Forschungsfrage F3: Hat der Themenbereich des Nachrichtenbeitrags einen Einfluss auf die Nutzerbewertungen?*

Neben der Beantwortung dieser Fragen war es außerdem Ziel der Studie, rezipientenspezifische Variablen zu identifizieren, die möglicherweise die Art beeinflussen, wie Leser Nachrichtenbeiträge beurteilen. Dazu sollte vor allem der Einfluss von drei Gruppen unabhängiger Rezipientenvariablen auf die Qualitätseinschätzung, die Gesamtbewertung des Beitrags und das Nutzungsinteresse untersucht werden: *Bildung* (höchster Bildungsabschluss), *Mediennutzung* (Häufigkeit der Nachrichtennutzung, Nutzung von Printmedien) sowie *Interesse und Vorwissen* (themenspezifisches Vorwissen, Interesse und persönliche Relevanz).

*Forschungsfrage F4: Haben Bildung, Mediennutzung sowie Interesse und Vorwissen einen Einfluss auf die Nutzerbewertungen?*

## 4. Methode

### 4.1 Inhaltsanalyse

Um Nachrichten unterschiedlicher Qualität finden zu können, wurde zunächst eine Inhaltsanalyse von Nachrichten in Tageszeitungen durchgeführt, bei denen im Allgemeinen Qualitätsunterschiede vermutet werden können: die überregionalen Tageszeitungen „Süddeutsche Zeitung“ und „Frankfurter Allgemeine Zeitung“ (FAZ), die lokalen Tageszeitungen „Sächsische Zeitung“ und „Dresdner Neueste Nachrichten“ (DNN) sowie die Straßenverkaufszeitungen „BILD Dresden“ und „Dresdner Morgenpost“. Alle Medien haben in ihrem jeweiligen Verbreitungsgebiet eine vergleichsweise hohe Auflage und werden dementsprechend auch häufig von Lesern genutzt (vgl. AWA 2009, IVW 2009).

Die Untersuchung fand am 24.07.2009 statt, an dem pro Zeitung 20 Artikel, also insgesamt 120 Nachrichtenbeiträge anhand einschlägiger Kriterien im Hinblick auf ihre journalistische Qualität analysiert wurden (siehe Tabelle 1)<sup>4</sup>.

3 Der Begriff „Themenbereich“ wird im Folgenden für die übergeordneten Themenkategorien „Sport“ und „Politik“ verwendet, während mit „Thema“ immer eine konkrete im Beitrag angesprochene Thematik (z. B. Stürmersuche des VfB Stuttgart) gemeint ist.

4 Für den Pretest haben zwei Codierer unabhängig voneinander eine Stichprobe von zwölf Beiträgen (zwei Artikel pro Medium) analysiert. Unter Verwendung der Formel von Früh (2001: 179) zur Berechnung der Intercoderreliabilität konnte eine Übereinstimmung von 92 Prozent über alle Variablen hinweg festgestellt werden. Für einzelne Variablen lag die Übereinstimmung jeweils bei mindestens 75 Prozent.

Tabelle 1: Operationalisierung der Qualitätsdimensionen

Dimension	Variable	Variable angelehnt an
Vielfalt	Thema/Anlass	–
	Meinungsvielfalt	–
	Quellenvielfalt	Hagen 1995
Transparenz	Quellentransparenz	Berdi et al. 1992
	Äußerungsanlässe	Hagen 1995
	Quellenangabe	Hagen 1995
	Urhebertransparenz	–
Sachgerechtigkeit	<i>Vollständigkeit des Hauptereignisses: Wer? Was? Wann? Wo?</i>	Schmiedgen 2006
	<i>Analytische Qualität: Ursachen / Hintergründe / Geschichtliches, Ablauf / Umstände, Folgen / Prognosen</i>	Schmiedgen 2006
	<i>Plausibilität</i>	Vehlow 2006
Relevanz	<i>Gesellschaftliche Relevanz: Anzahl der Betroffenen, Wirkungsintensität, Eintrittswahrscheinlichkeit</i>	Schulz 1976
	<i>Kollektive Relevanz: Nachrichtenfaktoren</i>	Schulz 1976
Unparteilichkeit	Neutralität	Schmiedgen 2006
Rechtmäßigkeit	Rechtmäßigkeit	–
Vermittlung	<i>Verständlichkeit: Semantische Einfachheit, Syntaktische Einfachheit, Gliederung / Ordnung, Kürze / Prägnanz, Anregender Zusatz: Direkte Rede</i>	Langner et al. 2006
	<i>Sprache</i>	–

Im Anschluss an die Inhaltsanalyse wurde für alle untersuchten Beiträge ein sogenannter „Qualitätspunktwert“ berechnet. Dieses Vorgehen war nötig, um einen standardisierten Wert zu erhalten, anhand dessen geeignete Artikel für das im Anschluss durchzuführende Experiment ausgewählt werden konnten. Die Berechnung dieses Qualitätspunkt-werts (QW), der zwischen 0 (sehr schlechte Qualität) und 100 (sehr gute Qualität) liegen kann, erfolgte in drei Schritten: Zunächst wurden die Codebuchskalen auf Werte zwischen 0 und 100 angepasst<sup>5</sup>, dann für jede einzelne der sieben Qualitätsdimensionen aus den Werten der ihnen zugehörigen Kriterien ein Mittelwert berechnet. Aus den sieben Punktwerten der einzelnen Dimensionen konnte im letzten Schritt wiederum der mittlere Gesamtpunktwert berechnet werden, der die Qualität eines untersuchten Artikels beschreibt. Im Mittel erreichten die Beiträge der Stichprobe 72 Punkte; das Minimum lag bei 44 und das Maximum bei 89 Punkten.

Insgesamt unterscheidet sich die Art der Berechnung des Qualitätsgesamtpunkt-werts grundsätzlich von der Methode von Berdi und Kollegen (1992). Im Gegensatz zu ihrer Untersuchung sind es bei der vorliegenden Studie nicht die einzelnen Qualitätskriterien,

5 Bsp.: Vollständigkeit (Wer, was, wann, wo): 0 = keine W-Fragen beantwortet; 25 = eine W-Frage beantwortet; 50 = zwei W-Fragen beantwortet; 75 = drei W-Fragen beantwortet; 100 = alle vier W-Fragen beantwortet.

die gleich gewichtet werden, sondern die sieben Dimensionen. Damit soll eine unausgewogene Betonung von Transparenz- und recherspezifischen Kategorien vermieden werden, die entstehen kann, wenn es für einzelne Dimensionen wesentlich mehr Indikatoren gibt als für andere. Außerdem schien die Gleichgewichtung hier auch mangels konkreter Anhaltspunkte in der bisherigen Forschung über die optimale Gewichtung der einzelnen Qualitätsdimensionen (vgl. Hassemer / Rager 2006: 20, 23) als die geeignetste Methode.

#### 4.2 Experiment

Im Anschluss an die Inhaltsanalyse wurde die Übereinstimmung der wissenschaftlich-normativen Qualitätseinstufungen mit den Nutzerbewertungen untersucht. Zur Beantwortung der Forschungsfragen fiel die Wahl auf ein experimentelles 2x2-Design mit einer Online-Befragung als Untersuchungsmethode.

Die Faktoren waren Themenbereich (Sport vs. Politik) und inhaltsanalytisch gemessene Qualität (gut vs. schlecht). Zusätzlich wurden verschiedene Rezipienteneigenschaften (Bildung, Mediennutzung, Interesse am Thema) als potenzielle Einflussfaktoren berücksichtigt. Von Interesse war deren Wirkung auf folgende abhängige Variablen: die gemittelte Nutzerbewertung der sieben klassischen Qualitätsdimensionen (Vielfalt, Transparenz, Sachgerechtigkeit, Relevanz, Unparteilichkeit, Vermittlung, Rechtmäßigkeit) und die allgemeine Gesamtbewertung (= Gefallen des Beitrags).

Für die Auswahl des Stimulusmaterials wurden mit Hilfe des nach der Inhaltsanalyse berechneten Qualitätspunktwerts die zehn schlechtesten (QW zwischen 44 und 60) und die zehn besten Artikel (QW zwischen 83 und 89) der Untersuchungsstichprobe identifiziert. Neben den Qualitätspunktwerten mussten bei der Auswahl auch die Themenbereiche der Artikel berücksichtigt werden, die entweder zu Sport oder Politik zugeordnet werden konnten und innerhalb dieser Kategorien möglichst ähnliche Themen behandeln sollten. Die Wahl der experimentellen Stimuli fiel somit auf die folgenden vier Nachrichten:

##### *Politik:*

- „Carstensen löst Landtag in Kiel auf“, FAZ vom 24.07.2009, Seite 1, 137 Wörter, 89 Qualitätspunkte (Gruppe 1)
- „Grüner Lichdi sagt der CDU den Kampf an“, DNN vom 24.07.2009, Seite 13, 159 Wörter, 60 Qualitätspunkte (Gruppe 2)

##### *Sport:*

- „Stürmersuche geht weiter“, Süddeutsche Zeitung vom 24.07.2009, Seite 28, 608 (nach Kürzung<sup>6</sup> 394) Wörter, 85 Qualitätspunkte (Gruppe 3)
- „VfB jetzt scharf auf den Sex-Bomber“, BILD Dresden vom 24.07.2009, Seite 15, 344 Wörter, 57 Qualitätspunkte (Gruppe 4).

Die Meinung der Studienteilnehmer zu den präsentierten Artikeln wurde mittels eines Online-Fragebogens erfasst, der neben den allgemeinen soziodemografischen Angaben aus zwei Teilen bestand: einem allgemeinen Teil, der Nachrichtennutzungsgewohnheiten abfragte, und einem speziellen Teil, in dem der präsentierte Nachrichtenartikel (je einer pro Experimentalgruppe, ohne Angabe der Quelle zum Ausschluss von Markenimageeffekten) anhand der sieben Qualitätsdimensionen Rechtmäßigkeit, Relevanz, Vielfalt, Sachgerechtigkeit, Unparteilichkeit, Vermittlung und Transparenz sowie an-

6 Um die Länge der Beiträge aus dem gleichen Themenbereich relativ konstant zu halten, wurde der Artikel ohne Beeinträchtigung der Qualität leicht gekürzt.

hand von vier nutzerspezifischen Kriterien (Themeninteresse<sup>7</sup>, persönliche Relevanz, Unterhaltsamkeit und Stil) bewertet werden sollte<sup>8</sup>. Zusätzlich wurde hier auch erfasst, wie viel Vorwissen bereits zum angesprochenen Thema vorhanden war.

Insgesamt nahmen 190 Versuchspersonen am Experiment teil<sup>9</sup>. Die Teilnehmer wurden im Schneeballverfahren per E-Mail rekrutiert, also zur Teilnahme und zur Weiterleitung der Einladung aufgefordert. Damit ist die Befragung zwar nicht repräsentativ, da das Forschungsziel jedoch nicht rein deskriptiv ist, sondern experimentellen Charakter hat, wurde aus forschungsökonomischen Gründen keine Zufallsstichprobe gezogen, sondern nur für die zufällige Verteilung der Versuchspersonen auf die Experimentalgruppen gesorgt (vgl. Diekmann 2008: 195)<sup>10</sup>. Frauen waren in der Stichprobe leicht überrepräsentiert (60 % der Befragten), ebenso Jüngere (Durchschnittsalter ist 33, die Altersspanne reicht von 13 bis 79 Jahren) und höher Gebildete (drei Viertel der Befragten besitzen einen Hochschulabschluss oder streben diesen gerade an). Um mögliche intervenierende Effekte auszuschließen, war die heterogene Zusammensetzung der vier Experimentalgruppen von Bedeutung. In Bezug auf Alter, Geschlecht und Bildung gab es zwischen den Experimentalgruppen jedoch keine signifikanten Unterschiede. Somit kann davon ausgegangen werden, dass die Gruppen soziodemografisch ähnlich zusammengesetzt waren und keine Konfundierungen mit den stimulusabhängigen Variablen auftraten<sup>11</sup>. Auch hinsichtlich der Häufigkeit ihres Nachrichtenkonsums, der Anzahl der genutzten Informationsquellen oder der Nutzung von Printmedien unterschieden sich die Experimentalgruppen nicht signifikant voneinander.

## 5. Ergebnisse

### 5.1 Beurteilung der einzelnen Qualitätsdimensionen und Gesamteindruck

Bei der Untersuchung medienbezogener Kritikfähigkeit im Zusammenhang mit den abhängigen Variablen war zunächst die Frage von Interesse, ob die Versuchspersonen überhaupt bereit waren, Angaben zu den einzelnen Qualitätsdimensionen zu machen, oder sich eher für die Ausweichkategorie „kann ich nicht beurteilen“ entschieden. Es zeigte sich, dass die Befragten in der Regel durchaus eine Meinung zur Qualität des gelesenen Beitrags hatten. Das Kriterium, bei dem die meisten Studienteilnehmer (18 %) angaben, sie könnten es nicht beurteilen, war die Vielfalt („Der Journalist hat verschiedene Meinungen und Quellen berücksichtigt“). 15 Prozent konnten nicht angeben, ob der Artikel plausibel, vollständig und gut recherchiert ist, und 10 Prozent hatten Probleme, die Transparenz („Man erkennt, woher die im Artikel genannten Informationen stammen“) einzuschätzen. Bei allen anderen Kriterien lag der Anteil derer, die „kann ich nicht beurteilen“ angaben, bei weniger als zehn Prozent.

7 Hiermit ist das Interesse am tatsächlichen Thema gemeint, nicht am Themenbereich.

8 Operationalisiert jeweils durch ein Statement pro Dimension / nutzerspezifisches Kriterium, bei dem die Teilnehmer den Grad der Zustimmung angeben mussten, z. B. für Vielfalt: „Der Journalist hat verschiedene Meinungen und Quellen berücksichtigt.“

9 Verteilung der Versuchspersonen auf die Experimentalgruppen: Gruppe 1: n = 55; Gruppe 2: n = 39; Gruppe 3: n = 49; Gruppe 4: n = 47.

10 Dies wurde mit Hilfe einer Einstiegsfrage nach dem Geburtsmonat erreicht. Bsp.: Alle im Januar, Mai und September Geborenen kamen in Gruppe 1, alle im Februar, Juni und Oktober Geborenen in Gruppe 2 etc.

11 Eine Gleichheit konnte allerdings mit Hilfe des Beta-Tests auch nicht zweifelsfrei nachgewiesen werden, da die Teststärke nicht hoch genug war.

Dementsprechend lohnt sich ein Blick auf die Zusammenhänge zwischen den Bewertungen der einzelnen Qualitätsdimensionen und dem Gesamteindruck, den die Leser vom Beitrag gewonnen haben. Die Antworten der Versuchspersonen auf die Frage „Wie hat Ihnen dieser Artikel ganz allgemein gefallen?“ (=Gesamtbewertung) korrelierten dabei erwartungsgemäß signifikant positiv ( $p < 0.01$ ) mit den Bewertungen der Teilnehmer in allen abgefragten Qualitätsdimensionen mit Ausnahme der Vielfalt. Besonders stark war der Zusammenhang mit den nutzerspezifischen Kriterien „Ich finde das Thema des Artikels interessant“ ( $r = .60$ ) und „Der Artikel ist in einem guten journalistischen Stil verfasst worden“ ( $r = .66$ ). Eine lineare Regressionsanalyse bestätigte dieses Ergebnis: Mit den beiden Variablen konnten zusammen 60 Prozent der Varianz in der Gesamtbewertung erklärt werden ( $r^2$ ). Integriert man alle elf Qualitätskriterien (die sieben Dimensionen und die vier nutzerspezifischen Kriterien) in die Analyse, verbessert sich dieser Wert nur um etwa vier Prozent. Das liegt auch daran, dass die Bewertungen der einzelnen Kriterien untereinander korrelieren und daher selbst wenig zur Erhöhung der erklärten Varianz beitragen können. Tabelle 2 zeigt die beta-Koeffizienten für alle die Qualitätskriterien, deren Einfluss sich im Regressionsmodell als signifikant erwies. Auffällig ist, dass von diesen nur eines aus der Gruppe der „klassischen“ Qualitätsdimensionen stammt.

*Tabelle 2: Einfluss der Qualitätskriterien auf das Gefallen*

	Gefallen (n = 171-188) beta
Der Artikel ist in einem guten journalistischen Stil verfasst worden (= Nutzerkriterium Stil).	+ .41
Ich finde das Thema des Artikels interessant (= Nutzerkriterium Interesse).	+ .24
Man erkennt, woher die im Artikel genannten Informationen stammen (= Transparenz).	+ .16
<b>Erklärte Varianz durch alle elf Qualitätskriterien</b>	<b><math>r^2 = .64</math></b>

Multiple Regression; nur signifikante beta-Koeffizienten ausgewiesen ( $p < .05$ ); weitere acht in die Analyse einbezogene Kriterien lieferten keine signifikanten beta-Koeffizienten.

Insgesamt können aber auch die nutzerspezifischen Kriterien nur zum Teil erklären, ob ein Beitrag den Lesern gefällt. Etwa 35 Prozent der Varianz werden also offenbar durch andere Faktoren bestimmt. Auskunft über diese noch nicht berücksichtigten Kriterien geben die Angaben der Studienteilnehmer auf die offen gestellte Frage nach Aspekten, die ihnen am Nachrichtenbeitrag gefallen bzw. nicht gefallen haben. Die Aussagen wurden nach der Untersuchung zu Kategorien zusammengefasst (Tabelle 3) und im Datensatz entsprechend verschlüsselt.

Die Ergebnisse bestätigen einerseits die mit Hilfe der Regression festgestellte Tendenz, dass das Interesse für das Thema und die Bewertung des Stils für die Nutzer von besonderer Bedeutung sind. Häufig kamen jedoch auch Anmerkungen, die die Beiträge als zu weitschweifig und detailliert herausstellten: „zu viele unwichtige Informationen“, „zu viele Zahlen und Fakten“, „zu lang“. Dieses für Leser offenbar sehr wichtige Kriterium der Kürze und Prägnanz wird theoretisch oft nur als Unterkriterium der Verständlichkeit verortet (siehe Langner et al. 2006). Für Mediennutzer scheint es jedoch von besonderer, von der Verständlichkeit teilweise unabhängiger Bedeutung zu sein,

Tabelle 3: Zusammenfassung der Nutzerkommentare

Gefallen	Anzahl der Nennungen	nicht gefallen	Anzahl der Nennungen
Kürze & Prägnanz	26	Uninteressantes Thema	42
Hintergründe / Recherche	18	Schlechter Stil	38
Neutralität	16	Unverständlichkeit	36
Guter Stil	13	Weitschweifigkeit, unwichtige Informationen	35
		Schlechter Aufbau	21
		Oberflächlichkeit	17

Auswahl der häufigsten Nennungen (n > 10)

dass ein Artikel sich auf die wesentlichen Informationen beschränkt und nicht zu viele, für sie uninteressante Fakten enthält.

### 5.2 Vergleich von normativer Qualitätseinschätzung und Nutzerbewertung

Ob Lesern ein Artikel gefällt, hängt also stark davon ab, inwieweit sie das Thema interessiert, ihnen der Beitrag stilistisch zusagt und ob er kurz, knapp und verständlich die wichtigsten Fakten zusammenfasst. Damit nehmen Mediennutzer eine andere Gewichtung vor, als dies in der Inhaltsanalyse stattfand, bei der alle Qualitätsdimensionen gleichwertig in die Berechnung des Qualitätspunktwerts einbezogen wurden. Vor diesem Hintergrund muss Forschungsfrage F2 beantwortet werden: Bewerten Nutzer trotz dieser Unterschiede einen inhaltsanalytisch als qualitativ voll identifizierten Beitrag positiver als einen weniger qualitativ vollen Beitrag?

Aus der qualitativen Auswertung der Kommentare ließen sich bereits Erkenntnisse zum Verhältnis zwischen wissenschaftlich-normativen und Nutzerqualitätsbewertungen ablesen:

- Mediennutzer ziehen noch andere als die „klassischen“ Qualitätsdimensionen zur Beurteilung zu Rate (vor allem Themeninteresse und Stil).
- Die Befunde der Inhaltsanalyse stimmten mit den Bewertungen der Versuchspersonen weitgehend überein. Allerdings hatten die Befragten andere Schwerpunkte und machten selten Aussagen zu Transparenz oder Vielfalt. Dadurch fielen ihre Einschätzungen insgesamt negativer aus.

Zusätzlich sollte im Rahmen der quantitativen Auswertung geklärt werden, wie sich die Bewertung der vier verschiedenen Beiträge jeweils in den einzelnen Qualitätsdimensionen voneinander unterscheiden (siehe Tabelle 4). Dazu wurden einfache t-Tests gerechnet. Zum Vergleich zeigt Tabelle 4 auch den inhaltsanalytisch ermittelten Qualitätspunktwert für die jeweiligen Dimensionen.

Artikel 1 und Artikel 2 zum Thema „Politik“ unterschieden sich in ihren Qualitätspunktwerten aus der Inhaltsanalyse hauptsächlich in den Dimensionen Vielfalt, Transparenz, Vermittlung und Unparteilichkeit. In allen vier schnitt Artikel 1 deutlich besser ab. Auch die Versuchspersonen schätzten Artikel 1 signifikant vielfältiger und unparteilicher ein als Artikel 2. Zusätzlich hielten sie ihn für gesellschaftlich relevanter und waren eher der Meinung, er achte das Privatleben der Betroffenen. Artikel 2 wurde hingegen von den Versuchspersonen als unterhaltsamer eingestuft.

Bei den Beiträgen zum Thema Sport wurde Artikel 3 in der inhaltsanalytischen Untersuchung vor allem in den Dimensionen Vielfalt, Unparteilichkeit und Vermittlung als

Tabelle 4: Nutzerbewertungen der einzelnen Qualitätskriterien

	Politik		Sport	
	Artikel 1 Qualität + (n=48-55)	Artikel 2 Qualität - (n=36-39)	Artikel 3 Qualität + (n=37-49)	Artikel 4 Qualität - (n=38-45)
Der Journalist hat verschiedene Meinungen und Quellen berücksichtigt. <sup>a</sup>	3,0	2,4	3,2	3,0
Vielfalt (QW Inhaltsanalyse)	75	0	75	50
Man erkennt, woher die im Artikel genannten Informationen stammen.	2,6	2,8	2,6	2,5
Transparenz (QW Inhaltsanalyse)	100	75	67	67
Der Artikel ist plausibel, vollständig und gut recherchiert. <sup>b</sup>	3,3	2,9	3,4	2,3
Sachgerechtigkeit (QW Inhaltsanalyse)	92	73	100	100
Das Thema des Artikels ist für die Gesellschaft und die Leser von Bedeutung. <sup>a, b</sup>	4,1	2,7	2,0	1,7
Relevanz (QW Inhaltsanalyse)	68	60	52	43
Der Journalist hat neutral und wertfrei berichtet. <sup>a, b</sup>	4,1	3,0	3,8	2,0
Unparteilichkeit (QW Inhaltsanalyse)	100	50	100	0
Der Journalist hat die Informationen sprachlich korrekt und verständlich dargestellt. <sup>b</sup>	3,6	3,5	3,3	2,7
Vermittlung (QW Inhaltsanalyse)	90	60	100	40
Der Journalist hat das Privatleben und die Ehre der Personen geachtet, über die er berichtet hat. <sup>a, b</sup>	4,6	4,2	3,9	1,7
Rechtmäßigkeit (QW Inhaltsanalyse)	100	100	100	100
<i>Nutzerqualitätskriterien</i>				
Das Thema des Artikels ist für mich selbst von Bedeutung.	2,0	2,2	1,2	1,3
Ich finde das Thema des Artikels interessant.	3,0	2,6	1,6	1,4
Der Artikel ist unterhaltsam. <sup>a, b</sup>	2,0	2,4	2,0	2,5
Der Artikel ist in einem guten journalistischen Stil verfasst worden. <sup>b</sup>	3,0	2,7	2,6	1,8

MW auf einer Skala von 1 = „stimme gar nicht zu“ bis 5 = „stimme voll und ganz zu“; <sup>a</sup> Artikel 1 und 2 unterscheiden sich signifikant voneinander (t-Test); <sup>b</sup> Artikel 3 und 4 unterscheiden sich signifikant voneinander; QW = Qualitätspunktwert (Wertebereich liegt zwischen 0 und 100).

wesentlich qualitätsvoller als Artikel 4 eingeschätzt. Die Studienteilnehmer bewerteten ihn zusätzlich als signifikant besser recherchiert, gesellschaftlich relevanter und stilistisch ansprechender. Außerdem stimmten sie eher der Aussage zu, dass der Beitrag das Privatleben der Betroffenen achte. Schließlich stuften sie wie schon bei den politischen Nachrichten auch hier den weniger qualitätsvollen Artikel als unterhaltsamer ein.

Insgesamt ähnelten die Bewertungen der Befragten also tendenziell den Ergebnissen der Inhaltsanalyse. Allerdings gab es auch deutliche Unterschiede, insbesondere bei der

Dimension Rechtmäßigkeit. Wahrscheinlich liegt dies in der Tatsache begründet, dass sich die Definition dieser Dimension in der Inhaltsanalyse und deren Auslegung durch die Nutzer im Experiment möglicherweise gravierend voneinander unterscheiden: In der Inhaltsanalyse sollte ein Artikel nur dann als „unrechtmäßig“ eingestuft werden, wenn er im juristischen Sinn eindeutig die persönliche Ehre oder Privatsphäre der Personen verletzt, über die berichtet wurde – und zwar so, dass das Recht auf Meinungsfreiheit demgegenüber zurücktreten müsste. Das war bei keinem der vier präsentierten Artikel der Fall. Im Experiment sollten die Versuchspersonen jedoch nur einschätzen, ob der Journalist die persönliche Ehre und Privatsphäre der Personen achtet, über die er berichtet; eine Abwägung mit der Meinungsfreiheit wurde nicht gefordert. Dennoch ist dieses Ergebnis ein Indiz dafür, dass Nutzer der Meinungsfreiheit möglicherweise engere Grenzen setzen würden als der Gesetzgeber. Da sie aber nicht wirklich nach einer Abwägung gefragt wurden, müsste die Frage zur Bestätigung dieser Vermutung in einer späteren Untersuchung anders formuliert werden.

### *5.3 Einfluss von Qualität und Themenbereich auf die Nutzerbewertungen*

Dimensionsübergreifend war für die Studie entsprechend der Forschungsfragen F2 und F3 hauptsächlich von Interesse, inwieweit es von der Qualität und vom Themenbereich abhängt, ob Lesern ein Beitrag gefällt, und wie sie seine Qualität einstufen.

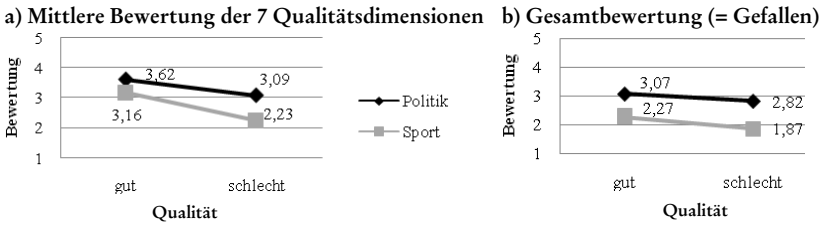
Abbildung 2 zeigt die Mittelwerte der vier Experimentalgruppen für die jeweiligen abhängigen Variablen. Für die mittlere Einschätzung der sieben „klassischen“ Qualitätsdimensionen ließ sich mit Hilfe einer mehrfaktoriellen Varianzanalyse sowohl ein signifikanter Haupteffekt des Themenbereichs als auch der inhaltsanalytisch ermittelten Qualität nachweisen. Die Politik-Beiträge wurden als deutlich qualitätsvoller eingeschätzt als die Sport-Beiträge, und die inhaltsanalytisch als qualitativ höherwertig identifizierten Beiträge erhielten auch von den Versuchspersonen eine positivere Bewertung als die schlechter eingestuften Artikel. Dabei ergab sich zwischen Themenbereich und Qualität sogar ein signifikanter Interaktionseffekt: Die Bewertungen der beiden Sportbeiträge verschiedener Qualität unterschieden sich deutlicher voneinander als die Bewertungen der beiden Politikbeiträge. Ähnliche Ergebnisse zeigten sich in Bezug auf die Gesamtbewertung. Eine mehrfaktorielle Varianzanalyse ergab auch hier einen signifikanten Haupteffekt beider unabhängiger Variablen (Themenbereich und Qualität), allerdings keinen Interaktionseffekt.

Insgesamt wurden die Beiträge zum Thema Politik in beiden abhängigen Variablen (Gefallen und mittlere Einschätzung der Qualitätsdimensionen) positiver bewertet. Dieser Effekt ist zum Teil darauf zurückzuführen, dass der Spielertransfer beim VfB Stuttgart als Thema insgesamt auf besonders wenig Interesse stieß. Außerdem legten die Befragten viel Wert auf Kürze und Prägnanz und präferierten vielleicht auch deshalb die kürzeren Nachrichtenbeiträge zum Thema Politik gegenüber den deutlich längeren Sportartikeln.

### *5.4 Einfluss von Rezipienteneigenschaften auf Nutzerbewertungen*

Die Forschungsfragen F2 und F3 können also mit Einschränkungen bejaht werden: Sowohl Themenbereich als auch Qualität beeinflussen die Nutzerbewertungen zum Teil deutlich. Diese Effekte bleiben auch dann bestehen, wenn mehrere unabhängige Rezipientenvariablen in das Wirkungsmodell einbezogen werden. Ihr Einfluss auf die Nutzerbewertungen kann damit als relativ stabil angesehen werden.

Abbildung 2: Nutzerbewertungen in Abhängigkeit von Themenbereich und Qualität

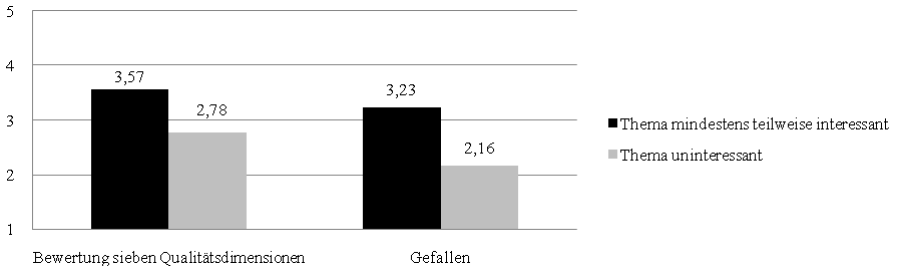


a) MW der Bewertung der sieben Qualitätsdimensionen auf einer Skala von 1 = „stimme gar nicht zu“ (Kriterien nicht erfüllt) bis 5 = „stimme voll und ganz zu“ (Kriterien erfüllt)  
 Haupteffekte: Thema:  $F = 68,0$ ;  $p < .001$ ; Qualität:  $F = 55,4$ ;  $p < .001$   
 Interaktionseffekt Thema x Qualität:  $F = 5,3$ ;  $p < .05$

b) MW auf einer Skala von 1 = „gar nicht gefallen“ bis 5 = „sehr gut gefallen“  
 Haupteffekte: Thema:  $F = 49,2$ ;  $p < .001$ ; Qualität:  $F = 6,6$ ;  $p < .05$

Ein separater Einfluss der Rezipientenvariablen selbst konnte hingegen ähnlich wie bei Berdi et al. (1992) entgegen der eingangs aufgestellten Vermutungen bei den meisten Eigenschaften nicht gemessen werden. Das Interesse am Thema war damit die einzige Variable, die unabhängig vom Themenbereich und der Qualität der Beiträge dafür sorgte, dass die Nutzer die Artikel sowohl insgesamt als auch innerhalb der Qualitätsdimensionen besser bewerteten (siehe Abb. 3). Vorwissen zum jeweiligen Thema hatte zumindest auf Letzteres ebenfalls einen signifikanten Haupteffekt.

Abbildung 3: Nutzerbewertungen in Abhängigkeit vom Themeninteresse



Gesamtbewertung: MW auf einer Skala von 1 = „gar nicht gefallen“ bis 5 = „sehr gut gefallen“,  
 Bewertungen der einzelnen Qualitätsdimensionen: MW auf einer Skala von 1 = „stimme gar nicht zu“ (Kriterien nicht erfüllt) bis 5 = „stimme voll und ganz zu“ (Kriterien erfüllt)  
 t-Tests signifikant,  $p < .001$

Die eingangs vermuteten Interaktionseffekte mit der Qualität, insbesondere, dass gebildete und interessierte Leser gute Beiträge besser und schlechte Beiträge schlechter bewerten als die restlichen Versuchspersonen, konnten nicht bestätigt werden. Allerdings weisen zumindest die offenen Antworten darauf hin, dass am Thema interessierte Leser, ähnlich wie es das ELM voraussagt, eher auf Argumente und Fakten achten (siehe Tabelle 5).

Tabelle 5: Nutzerkommentare nach Interesse am Thema

Aussagen zu	nicht am Thema des Beitrags interessiert in % (n = 122)	am Thema des Beitrags interessiert in % (n = 66)
Länge des Beitrags (n = 61)	31	35
Stil (n = 51)	31	20
Thema* (n = 44)	31	9
Verständlichkeit (n = 42)	20	27
Hintergründe/Recherche* (n = 35)	14	26
Neutralität (n = 26)	12	17
Aufbau (n = 23)	12	12

\* Unterschiede zwischen dem Anteil der Interessierten, die Aussagen zum jeweiligen Qualitätskriterium machen, und dem Anteil der Uninteressierten sind signifikant.

26 Prozent der am Thema Interessierten und nur 14 Prozent der Uninteressierten machten Angaben zu Hintergründen und Recherche. Im Gegenzug wurde das Thema als Grund dafür, dass der Beitrag (nicht) gefällt, häufiger von den nicht Interessierten genannt.

## 6. Zusammenfassung und Diskussion

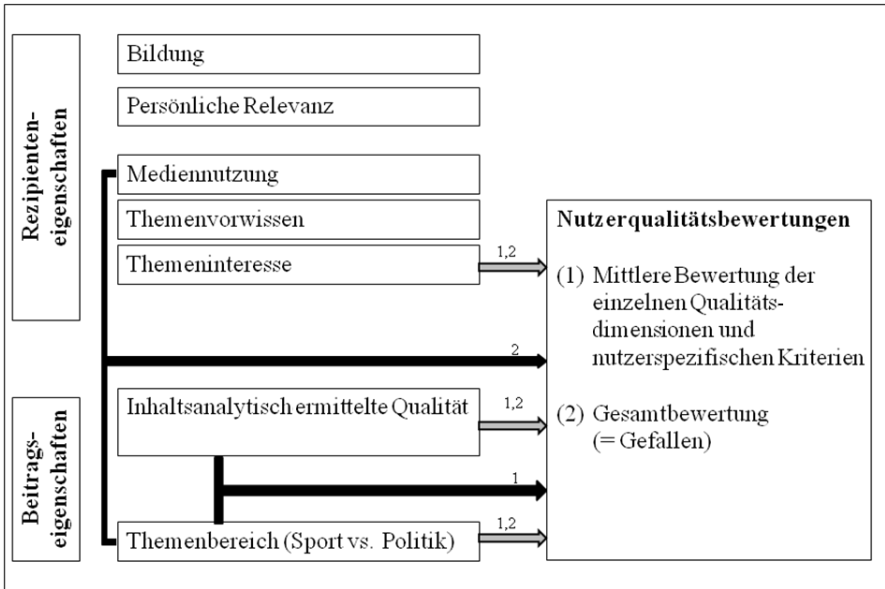
Die vorliegende Studie kann erste Erkenntnisse darüber liefern, wie Leser die Qualität von Nachrichten beurteilen. Als Untersuchung, die Nutzerbewertungen auch auf Beitragsebene gemessen hat, ermöglichte sie es den Versuchspersonen, ihre Meinung zur Nachrichtenqualität anhand eines konkreten Beispiels abzugeben. Im Gegensatz zu bisherigen Studien wurden Einschätzungen gemessen, die frei von Vorprägungen durch bestimmte Medienimages und ohne Zwang zu Verallgemeinerungen getroffen worden sind. Dabei hat sich gezeigt, dass Mediennutzer durchaus einen Unterschied zwischen Beiträgen schlechter Qualität und Beiträgen guter Qualität erkennen können – auch wenn ihre Qualitätsbewertung und vor allem die Gewichtung verschiedener Dimensionen nicht immer mit den der Inhaltsanalyse zugrundeliegenden normativen Einschätzungen übereinstimmten.

Zusammenfassend zeigt Abbildung 4, welche eingangs vermuteten Einflüsse auf die Beitragsqualität und die Nutzerqualitätsbewertungen bestätigt werden konnten.

Hervorzuheben ist hierbei insbesondere das Themeninteresse, das sich positiv auf die Bewertung auswirkte, allerdings – entgegen der eingangs in Anlehnung an das ELM aufgestellten Vermutung – unabhängig von der Qualität der Beiträge. Weitere Interaktionseffekte von Bildung und Mediennutzung mit der inhaltsanalytisch gemessenen Qualität, die gemäß dem Konzept der Medienkompetenz oder der Wissenskluthypothese vermutet wurden, ließen sich ebenfalls nicht bestätigen. In der Stichprobe waren Akademiker allerdings überrepräsentiert, sodass es möglich ist, dass sich ein Einfluss der Bildung auch wegen der geringen Anzahl an Nicht-Abiturienten nicht als signifikant erwies.

Dennoch liefert die Studie einen ersten Ansatz zur Beantwortung der eingangs formulierten Fragestellung: Nutzer können grundsätzlich hohe von niedriger Qualität unterscheiden. Dabei argumentieren sie jedoch hauptsächlich mit Aspekten der Verständlichkeit, des Stils, des Themas und der Beitragslänge – also eher konsumentenorientierten Kriterien –, während sie weniger auf Transparenz oder Vielfalt achten. Dies ist ein möglicher Hinweis darauf, dass es dem Publikum in seiner Bürgerrolle noch stellenweise an

Abbildung 4: Einflussfaktoren auf die Nutzerbewertungen



Die Zahlen geben an, auf welche der zwei verschiedenen Nutzerbewertungsvariablen Beitrags- bzw. Rezipienteneigenschaften einen signifikanten Einfluss hatten (graue Pfeile = Haupteffekte, schwarze Pfeile = Interaktionseffekte).

medienbezogener Kritikfähigkeit mangelt. Außerdem scheinen Leser letztendlich ebenso stark von den behandelten Themen beeinflusst zu werden und bewerten Beiträge nur dann gut, wenn ein gewisses Mindestinteresse vorhanden ist. Daraus lässt sich schlussfolgern, dass Zeitungen ihre Leser hauptsächlich dadurch binden können, dass sie für sie interessante und relevante Themen auswählen und diese verständlich und stilistisch ansprechend aufbereiten. Allerdings wurden im Experiment auch oft Angaben zur Recherchequalität gemacht, sodass die Hoffnung besteht, dass Einsparungen auf diesem Sektor ebenfalls von den Lesern bemerkt werden. Aufgrund der geringen Stichprobengröße und der fehlenden Repräsentativität der Untersuchung müssen diese allgemeinen Schlussfolgerungen über die einzelnen Experimentalgruppen hinweg jedoch vorerst noch als Vermutungen betrachtet werden. Folgestudien sollten daher eine diversifizierte und größere Stichprobe in die Untersuchung einbeziehen. Zusätzlich könnten sie, anstatt ausschließlich authentische Beiträge als Stimulusmaterial zu verwenden, identische Nachrichten jeweils in Bezug auf eine bestimmte Qualitätsdimension manipulieren. Schließlich erhöhte die Verwendung „echter“ Beiträge zwar die externe Validität der Studie, dies ging jedoch mit einem Verlust an interner Validität einher, da das authentische Material viele unkontrollierbare Variablen enthalten kann, die miteinander konfundieren. Eine Möglichkeit wäre, online ein Studienpanel einzurichten, bei dem die Teilnehmer regelmäßig manipulierte aktuelle Nachrichten lesen und bewerten. In diesem Zusammenhang könnte man noch zusätzlich zu den einzelnen Qualitätsdimensionen und dem Gefallen ein Gesamtqualitätsurteil der Nutzer als abhängige Variable erheben, um einen Eindruck davon zu bekommen, inwieweit sich dieses Urteil vom Mittelwert

der einzelnen Qualitätsdimensionen und vom Gefallen als eher konsumentenorientierter Form einer Gesamtbewertung unterscheidet.

## Literatur

- Arnold, K. (2009): Qualitätsjournalismus. Die Zeitung und ihr Publikum. Konstanz: UVK.
- Allensbacher Markt- und Werbeträgeranalyse (AWA) (2009), [www.awa-online.de](http://www.awa-online.de) [07.07.2011].
- Beck, K.; Reineck, D.; Schubert, S. (2010): Journalistische Qualität in der Wirtschaftskrise. Eine Studie im Auftrag des Deutschen Fachjournalisten-Verbandes (DFJV), [www.dfv.de/fileadmin/user\\_upload/pdf/Studie\\_Journalistische\\_Qualitaet\\_03\\_2010.pdf](http://www.dfv.de/fileadmin/user_upload/pdf/Studie_Journalistische_Qualitaet_03_2010.pdf) [07.07.2011].
- Berdi, C.; Ebert, A.; Faust, G.; Graber, T. (1992): Hauptsache gut aufgesetzt?! Qualität und Qualitätsurteile im Lokalen. Unveröffentlichte Diplomarbeit Universität Dortmund.
- Bonfadelli, H. (2004): Medienwirkungsforschung 1 (3. üb. Auflage). Konstanz: UVK Medien.
- Brosius, H.-B. (1995): Alltagsrationalität in der Nachrichtenrezeption: Ein Modell zur Wahrnehmung und Verarbeitung von Nachrichteninhalten. Opladen: VS.
- Chaiken, S. (1980): Heuristic versus Systematic Information processing and the Use of Source versus Message Cues in Persuasion. *Journal of Personality and Social Psychology*, 39 (5): 752-766.
- Dahinden, U.; Kaminski, P.; Niederreuther, R. (2004): ‚Content is king‘ – Gemeinsamkeiten und Unterschiede bei der Qualitätsbeurteilung aus Angebots- und Rezipientenperspektive. In: K. Beck, W. Schweiger, W. Wirth (Hrsg.), *Gute Seiten – schlechte Seiten. Qualität in der Onlinekommunikation* (S. 103-126). München: Reinhard Fischer.
- Daschmann, G. (2009): Qualität von Fernsehnachrichten: Dimensionen und Befunde. *Media Perspektiven*, Heft 5: 257-266.
- Diekmann, A. (2008): *Empirische Sozialforschung. Grundlagen, Methoden, Anwendung* (19. Aufl.). Hamburg: Rowohlt.
- Donsbach, W.; Rentsch, M.; Schielicke, A.-M.; Degen, S. (2009): Entzauberung eines Berufs. Was die Deutschen vom Journalismus erwarten und wie sie enttäuscht werden. Konstanz: UVK.
- Fahr, A. (2001): Katastrophale Nachrichten? Eine Analyse der Qualität von Fernsehnachrichten. München: Reinhard Fischer.
- Früh, W. (2001): *Inhaltsanalyse. Theorie und Praxis* (5. üb. Auflage). Konstanz: UVK.
- Groeben, N. (2004): Medienkompetenz. In: R. Mangold, P. Vorderer, G. Bente (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 27-49). Göttingen: Hogrefe.
- Hagen, L. M. (1995): Informationsqualität von Nachrichten. Meßmethoden und ihre Anwendung auf die Dienste von Nachrichtenagenturen. Opladen: Westdeutscher Verlag.
- Hasebrink, U. (2008): Das multiple Publikum. Paradoxien im Verhältnis von Journalismus und Mediennutzung. In: B. Pörksen, W. Loosen, A. Scholl (Hrsg.), *Paradoxien des Journalismus. Theorie – Empirie – Praxis* (S. 513-530). Wiesbaden: VS.
- Hassemer, G.; Rager, G. (2006): Das Bessere als Feind des Guten. Qualität in der Tageszeitung. In: G. Rager, K. Graf-Szczuka, G. Hassemer; S. Süper (Hrsg.), *Zeitungsjournalismus. Empirische Leserschaftsforschung* (S. 19-26). Konstanz: UVK.
- Heinrich, J. (1999): *Medienökonomie. Band 2: Hörfunk und Fernsehen* (1. Auflage). Wiesbaden: VS.
- Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e. V. (IVW) (2009): Auflagenliste 2/2009, [http://daten.ivw.eu/download/20092\\_Auflagenliste.zip](http://daten.ivw.eu/download/20092_Auflagenliste.zip) [07.07.2011].
- Langner, I.; Schulz von Thun, F.; Tausch, R. (2006): *Sich verständlich ausdrücken* (8. Aufl.). München: Ernst Reinhardt.
- Maurer, M.; Reinemann, C. (2006): *Medieninhalte. Eine Einführung*. Wiesbaden: VS.
- McQuail, D. (1992): *Media Performance. Mass Communication and the Public Interest*. London, Newbury Park, New Delhi: Sage.
- Petty, R. E.; Cacioppo, J. T. (1986): The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology*, 19: 123 -205.
- Pöttker, H. (2000): Kompensation von Komplexität. Journalismustheorie als Begründung journalistischer Qualitätsmaßstäbe. In: M. Löffelholz (Hrsg.), *Theorien des Journalismus. Ein diskursives Handbuch* (S. 375-390). Wiesbaden: Westdeutscher Verlag.

- Rager, G. (2000): Ethik – eine Dimension von Qualität? In C. Schicha, C. Brosda (Hrsg.), *Medienethik zwischen Theorie und Praxis. Normen für die Kommunikationsgesellschaft* (S. 76-89). Münster: Lit. Verlag.
- Reitze, H.; Ridder, C.-M. (Hrsg.) (2006): *Massenkommunikation VII. Eine Langzeitstudie zur Mediennutzung und Medienbewertung 1964-2005*. Baden-Baden: Nomos.
- Rössler, P. (2004): Qualität aus transaktionaler Perspektive. Zur gemeinsamen Modellierung von „User Quality“ und „Sender Quality“: Kriterien für Onlinezeitungen. In: K. Beck, W. Schweiger, W. Wirth (Hrsg.), *Gute Seiten - schlechte Seiten. Qualität in der Onlinekommunikation* (S. 127-145). München: Reinhard Fischer.
- Ruß-Mohl, S. (1992): „Am eigenen Schopfe ...“ Qualitätssicherung im Journalismus – Grundfragen, Ansätze, Näherungsversuche. *Publizistik*, 37: 83-96.
- Schatz, H.; Schulz, W. (1992): Qualität von Fernsehprogrammen. Kriterien und Methoden zur Beurteilung von Qualität im dualen Fernsehen. *Media Perspektiven*, Heft 11: 690-771.
- Schmiedgen, A. (2006): *Online-Zeitungen: Ersatz oder Ergänzung der Print-Ausgaben? Ein inhaltsanalytischer Vergleich der Qualität der Print- und Online-Versionen von Süddeutscher Zeitung, Bild und Spiegel*. Unveröffentlichte Magisterarbeit, TU Dresden.
- Schulz, W. (1976): *Die Konstruktion von Realität in den Nachrichtenmedien. Analyse der aktuellen Berichterstattung*. Freiburg & München: Alber.
- Siegert, S. (2009): Prinzip Baukasten. *Journalist*, Heft 10: 13-17.
- Tebert, M. (2010): Profil durch Qualitätsmanagement. *Media Perspektiven*, Heft 2: 78-89.
- Tichenor, P. J.; Donohue, G. A.; Olien, C. N. (1970): Mass Media Flow and Differential Growth in Knowledge. *Public Opinion Quarterly*, 34 (2): 159-170.
- Trepte, S. (2008): Medienkompetenz. In: N. C. Krämer, S. Schwan, D. Unz, M. Suckfüll (Hrsg.), *Medienpsychologie. Schlüsselbegriffe und Konzepte* (S. 102-107). Stuttgart: Kohlhammer.
- Trepte, S.; Reinecke, L.; Behr, K.-M. (2008): Qualitätserwartungen und ethischer Anspruch bei der Lektüre von Blogs und von Tageszeitungen. *Publizistik*, 53: 509-534.
- Vehlow, B. (2006): *Qualität von Spätnachrichten-Sendungen*. München: Reinhard Fischer.
- Wirth, W. (1997): *Von der Information zum Wissen. Die Rolle der Rezeption für die Entstehung von Wissensunterschieden*. Opladen: Westdeutscher Verlag.
- Wolling, J. (2006): Medienqualität aus Rezipientensicht. Test eines qualitätsbasierten Selektionsmodells im Rahmen eines Mehrmethodenprojekts. In: S. Weischenberg, W. Loosen, M. Beuthner (Hrsg.), *Medien-Qualitäten. Öffentliche Kommunikation zwischen ökonomischem Kalkül und Sozialverantwortung* (S. 457-473). Konstanz: UVK.
- Wyss, V. (2002): *Redaktionelles Qualitätsmanagement. Ziele, Normen, Ressourcen*. Konstanz: UVK.
- Zubayr, C.; Geese, S. (2009): Die Informationsqualität der Fernsehnachrichten aus Zuschauersicht. *Media Perspektiven*, Heft 4: 158-173.