

# Glaubwürdigkeit und Selektion von Suchergebnissen

*Der Einfluss von Platzierung, Reputation, Neutralität und sozialen Empfehlungen bei der Nutzung von Suchmaschinen\**

Alexander Haas / Julian Unkel

*Die Studie untersucht das Selektionsverhalten bei der Nutzung von Suchmaschinen. Diesen kommt in vielen Fällen die Rolle als Gatekeeper zu, wenn Rezipienten online nach Informationen suchen. Dabei scheinen sich die Nutzer relativ unreflektiert an den Vorgaben zu orientieren, da sie hauptsächlich die erstplatzierten Suchergebnisse aufrufen. Bislang ist nicht geklärt, ob sie diesen ersten Suchergebnissen eine höhere Glaubwürdigkeit zuschreiben oder ob Urteilen über die Glaubwürdigkeit kein nennenswerter Einfluss auf die Selektion zukommt. In einer Beobachtungsstudie (n = 247) werden neben der Reihenfolge der präsentierten Suchergebnisse drei Glaubwürdigkeitsindikatoren systematisch variiert: die Reputation der Quelle, die Neutralität des Suchergebnisses und die Anzahl an Empfehlungen durch andere Nutzer. Die Ergebnisse belegen die vorrangige Bedeutung der Platzierung, geben aber auch Hinweise darauf, in welchen Situationen die Beurteilung der Glaubwürdigkeit die Auswahlentscheidungen stärker beeinflussen könnte.*

**Schlagwörter:** Glaubwürdigkeit, Selektion, Suchmaschinen, Online, Beobachtung, soziale Empfehlungen

## 1. Selektion im Internet

Die Nutzung massenmedialer Inhalte setzt seit jeher ein gewisses Maß an Selektion voraus. Dies galt und gilt für traditionelle Massenmedien, trifft aber in noch stärkerem Maße auf die Nutzung von Online-Medien zu. Wirth und Brecht konstatierten in diesem Zusammenhang bereits vor 16 Jahren einen erhöhten „Selektionsdruck“ (1999: 150). Dieser Zwang zur Selektion durch die Nutzer<sup>1</sup> hat sich im Zuge der Ausweitung des Online-Angebots weiter verstärkt. Dabei unterscheidet sich die Online-Umgebung aus Nutzerperspektive nicht nur hinsichtlich der Menge der zur Verfügung stehenden Angebote und Informationen von dem klassischen Medienangebot. Die Nutzer sind zudem mit einer größeren Unsicherheit hinsichtlich Herkunft und Qualität der aufzufindenden Informationen konfrontiert.

Suchmaschinen sind für den Großteil der Nutzer ein Mittel, um sich innerhalb des unübersichtlichen Angebots zu orientieren und Informationen und Websites aufzufinden (Kink & Hess 2008: 24–25). Bei der Evaluation derer Ergebnislisten sind Nutzer mit folgender Situation konfrontiert: Im Vergleich zu klassischen Massenmedien fällt im Online-Bereich der Anteil der den Rezipienten bekannten „Informationsanbieter“ deutlich geringer aus. Rezipienten verbinden mit Fernsehsendern bzw. -sendungen, mit Tageszeitungen usw. ein gewisses Image, das sich über Jahre herausgebildet hat. Solche

\* Die Autoren bedanken sich bei den Reviewerinnen/Reviewern für die vielen konstruktiven und hilfreichen Anmerkungen, bei der LMU München für die Förderung des Projekts im Rahmen von LMUexcellent sowie bei Nora Denner, Lena Storck und vor allem Ramona Berger, die uns als Versuchsleiterinnen und Codiererinnen unterstützt haben.

1 Aus Gründen der Lesbarkeit verwenden wir überwiegend die männliche Form. Natürlich sind dabei immer beide Geschlechter, also beispielsweise Nutzerinnen und Nutzer gemeint.

Images sind dabei weniger als rational gebildete Summe aller Nutzungserfahrungen zu verstehen. Vielmehr handelt es sich um häufig emotional geprägte Urteile (Schweiger 2007: 253). Dennoch liefern sie Rezipienten eine gewisse Orientierung. Und die Bedeutung solcher Images wächst mit der Größe des insgesamt zur Verfügung stehenden Medienangebots (Bentele 1988: 421). Je weniger stark ausgeprägt solche Images sind, in desto stärkerem Maße basieren die Selektionsentscheidungen von Rezipienten auf anderen Kriterien. Gleichzeitig sind die finanziellen und organisatorischen Voraussetzungen, um selbst ein Angebot zu erstellen, im Online-Bereich niedrig. Prinzipiell ist jeder dazu in der Lage und kann somit Informationen verbreiten. Die Beurteilung der Güte von Informationen stellt somit eine Kernaufgabe für Rezipienten dar (Metzger, Flanagin, Eyal, Lemus & McCann 2003: 294), deren Vernachlässigung mit der Gefahr einhergeht, inkorrekte Informationen zu rezipieren.

Vorliegende Ergebnisse (z. B. Tremel 2010: 249) lassen allerdings den Schluss zu, dass sich Rezipienten dieser Aufgabe, zumindest bei der Nutzung von Suchmaschinen, nicht bewusst sind. Vielmehr orientieren sie sich bei der Selektion hauptsächlich an der vorgegebenen Platzierung. Andere Indikatoren für die Zuverlässigkeit oder Vertrauenswürdigkeit der zu erwartenden Informationen finden *scheinbar* kaum Berücksichtigung. Ob dies tatsächlich so ist, lässt sich anhand von Logfile-Analysen allerdings nicht beantworten. Es besteht durchaus die Möglichkeit, dass Nutzer solche Indikatoren sehr wohl wahrnehmen und in ihre Auswahlentscheidungen mit einfließen lassen, die erstplatzierten Suchergebnisse aber von den Nutzern eben auch anhand solcher Indikatoren positiv beurteilt werden. Wir untersuchen daher, ob und wie unterschiedliche Glaubwürdigkeitsindikatoren die Selektionsentscheidung beeinflussen. Im Zentrum stehen dabei die Fragen, ob diesen Glaubwürdigkeitsindikatoren überhaupt ein nennenswerter Einfluss zukommt, wenn gleichzeitig die Platzierung der Suchergebnisse berücksichtigt wird, und ob es interindividuelle Unterschiede zwischen den Studienteilnehmern gibt.

## 2. Nutzung von Suchmaschinen

Suchmaschinen sind für den Großteil der Internetnutzer inzwischen zum festen Bestandteil ihres Nutzungsverhaltens geworden. In der letzten Welle der ARD/ZDF-Onlinestudie gaben 82 Prozent der Onliner an, Suchmaschinen mindestens wöchentlich zu nutzen. Dieser Wert liegt sogar leicht über dem Wert hinsichtlich der Nutzung von E-Mails (van Eimeren & Frees 2014: 386). Bei der Online-Suche nach Informationen haben Nutzer grundsätzlich drei Möglichkeiten. Erstens ist ein entsprechendes Angebot, das mit einer gewissen Wahrscheinlichkeit die gewünschten Informationen bereitstellt (z. B. *Wikipedia*), bereits bekannt und wird durch Eingabe der entsprechenden URL (*www.wikipedia.de*) direkt ausgewählt. Zweitens werden Suchmaschinen aber trotz Kenntnis eines solchen Angebotes genutzt, um dieses (wieder) zu finden. Man spricht von *navigational tasks* (Broder 2002: 5) bzw. navigationsorientierten Anfragen (Lewandowski 2015: 70), wenn etwa „spon“ als Suchanfrage formuliert wird, um zum Angebot von *SpiegelOnline* zu gelangen. In beiden Fällen weiß der Nutzer aber schon vorab, welches Angebot er aufrufen will. Drittens, und aus unserer Perspektive interessanter, formulieren Nutzer informationsorientierte Anfragen (Lewandowski 2015: 70), wenn sie noch kein konkretes Angebot im Sinn haben. In diesen Fällen fungieren Suchmaschinen quasi als „Gatekeeper“ (Machill, Beiler & Zenker 2008: 591). Dabei ist ihre Rolle aber nicht mit der von Journalisten als klassischen Gatekeepern identisch. Suchmaschinen haben keinen Einfluss darauf, *ob* ein Inhalt erstellt und online zur Verfügung gestellt wird. Sie präsentieren den Nutzern vor allem keine „Agenda“, anhand derer diese sich über das aktuelle Geschehen informieren könnten (vgl. dazu ausführlich Röhle 2010:

29–34). In Anbetracht des Verhaltens des überwiegenden Teils der Nutzer (ausführlicher weiter unten) kommt ihnen jedoch eine vergleichbare Rolle zu.

Prinzipiell nehmen Suchmaschinen nur eine Reihung vor und blenden nur in Ausnahmefällen (z. B. Rechtswidrigkeiten, Unterlassungsansprüche) Inhalte aus. Allerdings werden Suchergebnisse, die nicht auf den ersten Seiten platziert sind, nicht in nennenswertem Umfang genutzt, sodass sie auffindbar sind, aber nicht wahrgenommen werden. Selbst bei bekannten Nachrichtenseiten gelangt ein relevanter Anteil der Besucher über eine Suchmaschine zu den Angeboten und Artikeln.<sup>2</sup>

### 2.1 Platzierung als Selektionskriterium

Ist die Entscheidung für eine Suchmaschine gefallen, wenden Nutzer durchaus unterschiedliche Strategien bei der Selektion innerhalb der präsentierten Ergebnislisten an. Trotz dieser Unterschiede ist der Hauptbefund hinsichtlich der Auswahl aber, dass ein Großteil der Klicks auf eines der ersten Suchergebnisse, in vielen Fällen sogar auf das erste Suchergebnis entfällt (Agichtein, Brill, Dumais & Ragno 2006: 4; Machill u. a. 2008: 603; Machill, Neuberger, Schweiger & Wirth 2004: 330). Eine vergleichende Analyse der Ergebnisse mehrerer Studien zur Suchmaschinennutzung kam darüber hinaus zu dem Schluss, dass der Anteil der Nutzer, der lediglich die erste Ergebnisseite betrachtet, seit der Etablierung von Suchmaschinen sogar gestiegen ist (Jansen & Spink 2006: 257). Dies könnte darauf zurückzuführen sein, dass Suchmaschinen inzwischen von der überwiegenden Mehrheit der Nutzer großes Vertrauen entgegengebracht wird. Diese gehen möglicherweise davon aus, dass die relevantesten und glaubwürdigsten Suchergebnisse auch auf der ersten Seite zu finden sind.

*H1: Je höher die Platzierung eines Suchergebnisses ist, desto höher ist seine Selektionswahrscheinlichkeit.*

Gleichzeitig liegen Ergebnisse vor, die darauf schließen lassen, dass die Nutzer, trotz der grundlegenden Präferenz für eines der ersten Suchergebnisse, durchaus unterschiedliche Strategien bei der Selektion anwenden (vgl. Beiler 2005: 175–177). Die Studie offenbart verschiedene Dimensionen des Auswahlverhaltens, die neben dem Rangfolgen-Einfluss auch intuitiv-spontane, zufällige und evaluative Auswahl umfassen, ergo sich auf Kriterien abseits der Position des Resultats beziehen.

Im Folgenden gehen wir daher auf die Kriterien ein, die innerhalb von Suchmaschinen zur Glaubwürdigkeitsbeurteilung herangezogen werden können. Diese lassen sich idealtypisch danach unterscheiden, ob sie mit der wahrgenommenen Kompetenz oder Vertrauenswürdigkeit in Zusammenhang stehen sollten.

### 2.2 Glaubwürdigkeitsindikatoren als Selektionskriterien

Die Bedeutung von Urteilen über die Glaubwürdigkeit von Quellen oder Aussagen wurde im Rahmen der Medienwirkungsforschung bereits relativ früh erkannt (Self 2009: 435). Der Großteil der empirischen Studien konzentriert sich auf die postkommunikative Phase und analysiert den Einfluss der zugeschriebenen Glaubwürdigkeit auf das Auftreten und die Stärke von Medienwirkungen. Dies trifft auch auf den Online-Bereich zu. Hier finden sich vor allem Studien aus dem Bereich der Gesundheitskommunikation,

2 Verlässliche Zahlen zum exakten Anteil sind kaum verfügbar. Das Web-Analytics-Unternehmen *SimilarWeb* ([www.similarweb.com](http://www.similarweb.com)) gibt für deutsche Newssites wie *Zeit.de* und *sueddeutsche.de* Werte zwischen 20 und 25 Prozent an.

die etwa den Einfluss der zugeschriebenen Glaubwürdigkeit auf die Veränderung von Einstellungen und Meinungen untersuchen (z. B. Hu & Sundar 2009). Der Einfluss der zugeschriebenen Glaubwürdigkeit auf die *Auswahl* einer Mediengattung, eines bestimmten Medienangebots (z. B. einer bestimmten Tageszeitung) oder eines konkreten Inhalts (z. B. eines bestimmten Artikels) ist dagegen deutlich seltener untersucht worden (Sundar, Knobloch-Westerwick & Hastall 2007; Winter & Krämer 2012). Dabei hat die Beurteilung der Glaubwürdigkeit für Rezipienten, unabhängig von der Phase im Selektionsprozess, einen heuristischen Wert. Das Urteil hilft bei der Einordnung neuer Informationen, ohne dass die Botschaft selbst einer aufwändigen inhaltlichen Überprüfung unterzogen werden muss (O’Keefe 2008: 1475). Dies ist vor allem in Situationen hilfreich, in denen eine solche Überprüfung außerhalb der Möglichkeiten des Rezipienten oder eine große Menge an Informationen zur Verfügung steht. Beide Kriterien dürften auf einen Großteil der Online-Suchanfragen von Rezipienten zutreffen. Im Hinblick auf Nachrichten sollte also durchaus eine Evaluation der Glaubwürdigkeit stattfinden, da es unglaubwürdige Nachrichten für Rezipienten in den meisten Situationen nicht wert sind, gelesen zu werden oder sich mit ihnen auseinanderzusetzen (Sundar u. a. 2007: 369).

Wie bei anderen häufig analysierten Konstrukten fällt die konkrete Operationalisierung des Konstruktes *Glaubwürdigkeit* in empirischen Studien sehr unterschiedlich aus (vgl. Nawratil 2006). Konsens ist die Sichtweise, dass es sich bei Glaubwürdigkeit nicht um eine inhärente Eigenschaft eines Objektes, etwa einer Quelle, sondern um eine subjektive Zuschreibung handelt (Bentele 1988: 408). Diese Zuschreibung kann sich auf unterschiedliche Objekte beziehen. Im Hinblick auf Medieninhalte sind drei Bezugsobjekte relevant: die Quelle einer Aussage (source credibility), die Aussage selbst (message credibility) sowie der mediale Kanal im Allgemeinen (media credibility) (Metzger u. a. 2003: 297, 302, 306). Eine erste wissenschaftliche Auseinandersetzung mit der Quellenglaubwürdigkeit findet sich in den Yale-Studien von Hovland und Kollegen (1953), die die Kompetenz und die Vertrauenswürdigkeit der Quelle als entscheidende Dimensionen herausarbeiten. Glaubwürdigkeit verfügt demnach über eine faktische und eine moralische Komponente, anders gesagt: *Kann* (Kompetenz) und *will* (Vertrauenswürdigkeit) die Quelle die Wahrheit sagen.

Welches Objekt Gegenstand der Beurteilung wird, hängt von den Möglichkeiten in der konkreten Selektions- bzw. Rezeptionssituation, aber auch von der Motivation des Rezipienten ab. So wird der Inhalt der Aussage mit höherer Wahrscheinlichkeit evaluiert, wenn ein Rezipient ein starkes Interesse an dem entsprechenden Themenbereich hat oder sogar persönlich betroffen ist. Die Beurteilung der Quellenglaubwürdigkeit dient dagegen eher weniger involvierten Rezipienten als Indikator, da diese Evaluation mit geringerem Aufwand verbunden ist. Gleichzeitig existieren allerdings Selektionssituationen, in denen es Rezipienten unabhängig von ihrer Motivation überhaupt nicht möglich ist, die Aussagenglaubwürdigkeit zu beurteilen. Dies trifft auch auf die Evaluation und Auswahl von Suchergebnissen innerhalb der Ergebnisliste einer Suchmaschine zu. Tremel (2010: 249) geht in dieser Situation von einem geringen Einfluss der Glaubwürdigkeit auf die Selektionswahrscheinlichkeit aus, der deutlich hinter dem Rangfolgeeffekt zurückbleibt; die Probanden wichen vor allem dann von dem vorgegebenen Ranking ab, wenn sich besonders unglaubwürdige Suchergebnisse oben befanden. Allerdings wurde die Glaubwürdigkeit der Suchergebnisse nicht anhand bestimmter Attribute variiert, sondern per Vorstudie mit einer anderen Stichprobe erfasst. Welche spezifischen Glaubwürdigkeitsindikatoren die Auswahl beeinflussen, konnte daher nicht gezeigt werden.

### 2.2.1 Suchmaschineninhalte und wahrgenommene Kompetenz

Auch wenn die Wahrnehmung und Wirkung spezifischer Glaubwürdigkeitsindikatoren von Suchergebnissen bislang kaum untersucht wurden, finden sich in den Ergebnislisten von Suchmaschinen durchaus Indikatoren, die bereits in Studien zur Rolle der Glaubwürdigkeit im Online-Bereich aufgegriffen wurden. Von 51 untersuchten Website-Elementen werden von Fogg und Kollegen sieben der Dimension Kompetenz zugeordnet (Fogg u. a. 2001: 64). Dazu gehören vorrangig die Zugehörigkeit zu einer auch außerhalb des Internets respektierten Nachrichtenorganisation, das Hervorheben der Befähigung des Autors sowie Quellenangaben. Allein das Vorhandensein von weiteren Angaben über Autoren hat einen Effekt. So können sowohl Informationen über deren Biographie bzw. Tätigkeit als auch Autorenfotos für eine höhere wahrgenommene Glaubwürdigkeit sorgen (Johnson & Wiedenbeck 2009: 340–341). Zwar tauchen solche Informationen in Ergebnislisten von Suchmaschinen selten auf. *Google* zeigt allerdings sowohl Autorennamen und -fotos über bzw. neben dem Suchergebnis an, sofern dieser bei der eigenen sozialen Netzwerk-Seite *Google Plus* angemeldet ist. Dutta-Bergmann (2004: 264) weist zudem auf den Einfluss der *Information Completeness* – dies umfasst bei wissenschaftlichen Ergebnissen etwa Erläuterungen, methodische Details und Statistiken – hin: Je stärker diese erfüllt sei, desto glaubwürdiger werde die Quelle eingeschätzt. Zwar lässt sich die *Information Completeness* nur an den vollständigen Artikeln evaluieren; aber auch die Evaluation der Suchergebnisse (bestehend aus Überschrift und Snippet) in den Ergebnislisten kann einen ersten Eindruck vermitteln, wie sehr dieses Kriterium erfüllt sein wird.

Den stärksten Einfluss auf die wahrgenommene Kompetenz hat nach den Ergebnissen von Fogg und Kollegen (2001: 64) allerdings die Zugehörigkeit zu einer bekannten Marke bzw. Institution. Dies weist bereits darauf hin, dass die *Reputation* der Seite einen entscheidenden Einfluss auf die perzipierte Glaubwürdigkeit hat. Weitere Untersuchungen bestätigen dieses Ergebnis und zeigen, dass aufgrund der unterschiedlichen Reputation Nachrichtenseiten generell als glaubwürdiger als andere Seiten, wie etwa private Blogs, wahrgenommen werden (Flanagin & Metzger 2007: 329; Greer 2003: 22–23; Meyer, Marchionni & Thorson 2010: 111–112).

*H2: Suchergebnisse mit einer hohen Reputation werden mit einer höheren Wahrscheinlichkeit ausgewählt als solche mit einer niedrigen Reputation.*

### 2.2.2 Suchmaschineninhalte und wahrgenommene Vertrauenswürdigkeit

Als Indikatoren für die moralische Komponente der Glaubwürdigkeitsbeurteilung, die zugeschriebene Vertrauenswürdigkeit der Website, werden in der Literatur besonders Kriterien genannt, die entweder auf der Zielseite direkt (z. B. Abwesenheit von Werbung, Verlinkung von Artikelquellen) oder gar auf anderen Unterseiten dieser (z. B. Verlinkung von Redaktionskodizes und -richtlinien) zu finden sind (Fogg u. a. 2001: 64; Metzger u. a. 2003: 300–301). Ebenso spielt eine Rolle, von wo aus auf die Seite verlinkt wurde – in diesem Fall also die Suchmaschine selbst. Wird diese als vertrauenswürdig wahrgenommen, so wird auch die Zielseite, also die Seite, die sich hinter dem Suchergebnis verbirgt, als vertrauenswürdiger eingestuft – zumal zumindest ein Teil der Nutzer davon ausgeht, dass Suchmaschinen ihre Ergebnisse nach deren Glaubwürdigkeit sortieren (Nakamura u. a. 2007: 41). Eine höhere Platzierung würde demnach auch eine höhere Vertrauenswürdigkeit bedeuten (vgl. Westerwick 2013: 205). Als weitere Quellenindikatoren für eine hohe Vertrauenswürdigkeit, die auch in den Ergebnislisten von Suchmaschinen sichtbar sind, können der Name und die URL der Seite dienen. So können

Top-Level-Domains wie *.edu*, *.gov* oder *.org* auf einen nicht-kommerziellen Hintergrund der Seite verweisen (Wathen & Burkell 2002: 138), während mit dem Namen der Seite/des Seitenherausgebers entweder bereits eine bestimmte Vertrauenswürdigkeit verknüpft ist oder sich aus diesem eine parteiische oder kommerzielle Absicht herauslesen lassen kann.

Weiterhin kann ein Suchergebnis auch auf Aussagenebene anhand von Indikatoren zur Vertrauenswürdigkeit evaluiert werden. Vertrauenswürdigkeit im Sinne einer fairen, unparteiischen Darstellung kann über die vermutete *Neutralität* des Artikels (bzw. dessen Vorschau durch Überschrift und Snippet) abgebildet werden. So werden bei kontroversen Themen Aussagen als ehrlicher bzw. vertrauenswürdiger wahrgenommen, wenn diese zweiseitig sind, also Argumente für beide Seiten liefern (Winter & Krämer 2012: 82).

*H3: Suchergebnisse mit einer hohen Neutralität werden mit einer höheren Wahrscheinlichkeit ausgewählt als solche mit einer niedrigen Neutralität.*

Das Aufkommen der unter dem Schlagwort „soziale Medien“ zusammengefassten Webtechnologien ermöglicht es dem Nutzer, auf einfache Weise die Erfahrungen Dritter in die eigene Glaubwürdigkeitsbeurteilung einzubeziehen (vgl. Metzger, Flanagin & Medders 2010: 420–424). Bereits 1999 stellten Tseng und Fogg den Einfluss Dritter auch in der Online-Kommunikation unter dem Titel „reputed credibility“ heraus (Tseng & Fogg 1999: 42). Inwiefern sich soziale Empfehlungen auf Selektionsentscheidungen auswirken können, wurde bislang vor allem im E-Commerce-Kontext untersucht: Eine Reihe von Studien liefert Hinweise für das vorhandene Einflusspotenzial im Hinblick auf Produktbewertungen oder Kaufintentionen (Park, Lee & Han 2007; Ziegele & Weber 2015). Soziale Empfehlungen können auch in Suchmaschinen eine Rolle spielen. So ermöglicht es beispielsweise *Google* Seitenbetreibern, in den Snippets Informationen wie Produktbewertungen oder Teile des *Google-Plus*-Profils einzubinden. Die Forschung zu sozialen Empfehlungen und nicht zuletzt zu deren Rolle bei der Nachrichtenselektion in sozialen Medien (Leino, Riihã & Finnberg 2011; Xu 2013) geht von einer vorhandenen Wirkung aus. Auch in einem Eyetracking-Experiment zu *Google* konnte der Einfluss von sozialen Empfehlungen auf die Selektion gezeigt werden (Terbeck 2012: 51).

*H4: Suchergebnisse mit einem hohen Rating werden mit einer höheren Wahrscheinlichkeit ausgewählt als solche mit einem niedrigen Rating.*

### 2.2.3 Additivität und Wahrnehmung von Glaubwürdigkeitsindikatoren

Natürlich kann ein Suchergebnis auch alle drei der eben diskutierten Merkmale (Reputation, Neutralität, soziale Empfehlungen) aufweisen, die dem Nutzer als Glaubwürdigkeitsindikatoren dienen könnten. Sundar und Kollegen (2007: 370) formulieren in diesem Zusammenhang eine Additivitätshypothese und gehen von einem kumulativen Effekt aus. Je mehr Glaubwürdigkeitsindikatoren auf ein Objekt (in unserem Fall ein Suchergebnis) zutreffen, desto höher sollte die Beurteilung der Glaubwürdigkeit ausfallen. Gleichzeitig weisen die Autoren allerdings darauf hin, dass möglicherweise auch das Vorhandensein eines bestimmten Indikators bereits für eine hohe Glaubwürdigkeitsbeurteilung ausreichend sein könnte. Im Hinblick auf Suchergebnisse könnte etwa die Zugehörigkeit zu einem bekannten Offline-Medium (Reputation) bereits ausreichen, damit Nutzer das Suchergebnis als glaubwürdig beurteilen. Wir untersuchen daher den Effekt des Vorhandenseins mehrerer Glaubwürdigkeitsindikatoren.



*FF1: Kommt es beim Vorhandensein von zwei oder drei Glaubwürdigkeitsindikatoren zu einer steigenden Selektionswahrscheinlichkeit (additiver Effekt)?*

Zumindest zwei Gründe sprechen allerdings gegen einen *einheitlichen* Einfluss der Glaubwürdigkeitsindikatoren auf das individuelle Selektionsverhalten. Zum einen ist es plausibel, dass die Indikatoren nicht von allen Nutzern überhaupt wahrgenommen und in der gleichen Weise verarbeitet werden. So schlägt etwa Metzger (2007: 2087–2089) ein Modell zur Beurteilung der Glaubwürdigkeit von Websites vor, das (ähnlich dem *Elaboration Likelihood Model* – ELM) zwei unterschiedliche Modi der Informationsverarbeitung postuliert. In Abhängigkeit von Motivation und Fähigkeit evaluiert der Nutzer eine Website systematisch (zentral), heuristisch (peripher) oder überhaupt nicht. Übertragen auf die Evaluation der Glaubwürdigkeit von Suchergebnissen bedeutet dies, dass Wahrnehmung und Verarbeitung der Indikatoren von Motivation und Vorwissen abhängen sollten. Sind Nutzer in ein Thema involviert, sollten die Glaubwürdigkeitsindikatoren mit größerer Wahrscheinlichkeit ihre Selektion beeinflussen, während sie sich bei einer heuristischen Evaluation stärker an der Platzierung der Suchergebnisse orientieren sollten. Weiterhin zeigten die Ergebnisse von Beiler (2005: 175–177), dass Suchmaschinennutzer durchaus unterschiedliche Strategien anwenden. Nutzer, die Suchergebnisse einer gründlichen Evaluation unterziehen, sollten sich bei ihrer Selektion eher von Glaubwürdigkeitsindikatoren leiten lassen als solche, die Suchergebnisse intuitiv-spontan auswählen.

Zum anderen handelt es sich bei der Beurteilung der Glaubwürdigkeit um eine individuelle Zuschreibung. Selbst für den Fall, dass die Mehrheit der Nutzer einen Indikator wahrnimmt, ist damit nicht sichergestellt, dass sie diesen auch mehrheitlich in gleicher Weise beurteilt. So werden Nachrichtenseiten *generell* als glaubwürdiger wahrgenommen als Blogs. Gleichzeitig gibt es allerdings Hinweise darauf, dass zumindest unter Bloglesern solche Angebote allgemein als glaubwürdig, bisweilen sogar als glaubwürdiger als Nachrichtenseiten beurteilt werden (Johnson & Kaye 2004: 630–631; Schmierbach & Oeldorf-Hirsch 2012: 320). Auch die allgemeine Mediennutzung kann somit einen Einfluss auf die Beurteilung der Glaubwürdigkeit ausüben. Unsere zweite Forschungsfrage lautet daher:

*FF2: Beeinflussen Suchmaschinennutzung, allgemeine Mediennutzung und das Vorwissen über die Suchaufgaben die Präferenz für einzelne Glaubwürdigkeitsindikatoren?*

### 3. Methode

Zur Beantwortung der Fragen wurde im April 2014 in einem universitären Forschungslabor eine Beobachtungsstudie durchgeführt. Den Teilnehmern wurde mitgeteilt, dass es sich um eine Studie zur Nutzung einer neuen (tatsächlich existierenden) Suchmaschine handle. Nach einigen Einstiegsfragen zu Internet- und Suchmaschinennutzung wurden die Teilnehmer instruiert, sich anhand einer vorgegebenen Suchanfrage zu einem Thema zu informieren und im Anschluss an die Recherche einige Fragen zu diesem Thema zu beantworten. Hierzu wurde per *iFrame* eine selbst erstellte Ergebnisseite der Suchmaschine *DuckDuckGo*<sup>3</sup> in den Fragebogen eingebunden und mit einer rudimentären Browsersteuerung versehen (siehe Abb. 1). Die Teilnehmer konnten die Ergebnisliste

3 Die eher unbekanntere Suchmaschine *DuckDuckGo* wurde ausgewählt, da große Suchmaschinen wie *Google* von nahezu der gesamten Stichprobe alltäglich mehrfach genutzt werden und somit Manipulationen sofort ins Auge fallen würden.

navigieren und einzelne Ergebnisse aufrufen, es gab lediglich die Anweisung, die vorab eingestellte Suchanfrage nicht zu verändern. Für die Rechercheaufgabe gab es jeweils ein Zeitlimit von fünf Minuten, nach denen automatisch zu den Fragen zum Thema überleitet wurde. Die Teilnehmer konnten die Recherche aber auch selbst beenden und vor Ablauf der Zeit zu diesen Fragen wechseln. Die Rechercheaufgabe wurde dann mit einem zweiten Thema wiederholt. Im Anschluss wurden weitere Fragen, etwa zum allgemeinen Umgang mit Suchmaschinen gestellt. Nach Beendigung des Fragebogens erhielt jeder Teilnehmer eine Aufwandsentschädigung von zehn Euro.

### 3.1 Stimulus

Um eine möglichst natürliche Nutzungssituation zu erzeugen, wurden sowohl die Suchmaschinen-Ergebnisseiten als auch alle Zielseiten voll funktional in HTML nachgebaut. Auf jeder Ergebnisseite wurden unter einem als Werbung gekennzeichneten Eintrag acht manipulierte Suchergebnisse verlinkt (siehe Abb. 1).<sup>4</sup> Als Reaktion auf Aussagen im Pretest wurden außerdem zwei irrelevante Ergebnisse (z. B. Link zu einem Übersetzungsservice), jeweils an vierter und achter Stelle, eingebaut, um ein realistischeres Bild abzugeben.

Bei beiden Rechercheaufgaben wurden die Ergebnisseiten nach demselben Schema systematisch auf zwei Ebenen variiert. Zum einen wurden jeweils acht Stimulusversionen (d. h. Ergebnisseiten) erstellt, in denen sich die *Platzierung* der acht Suchergebnisse unterschied, wobei jedes Suchergebnis in jeder Gruppe eine jeweils andere Platzierung einnahm, also jeweils einmal an erster Stelle, einmal an zweiter Stelle etc. stand. Die Teilnehmer wurden per Urnenziehung gleichmäßig auf diese acht Versionen verteilt. Zum anderen wurden innerhalb der acht Suchergebnisse die Faktoren *Reputation*, *Neutralität* und *soziale Empfehlung* auf je zwei Stufen variiert, sodass es jede Faktorkombination genau einmal gab. Jeder Teilnehmer sah also dieselben acht Suchergebnisse unabhängig davon, welcher Stimulusversion er oder sie zugeordnet wurde, nur jeweils in einer anderen Reihenfolge.

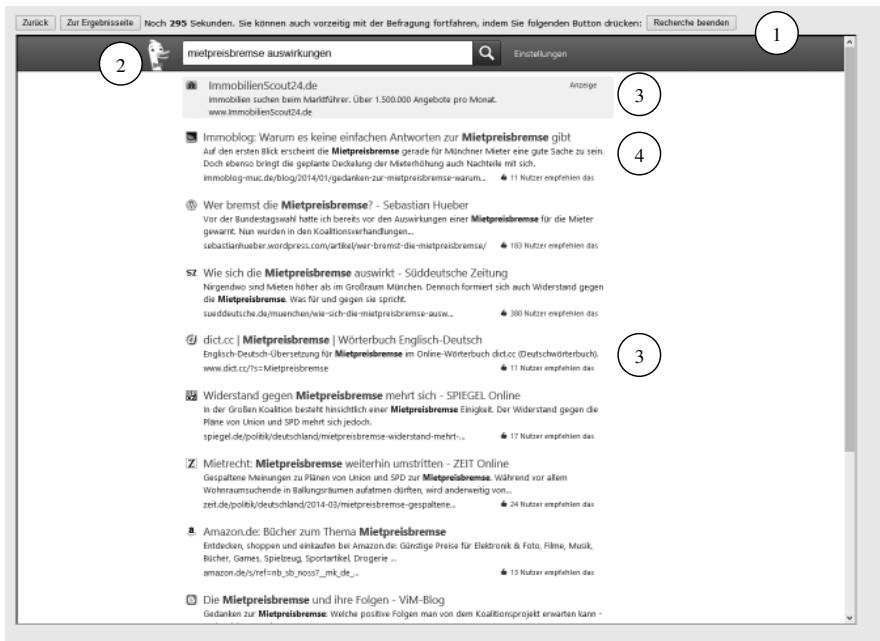
- (1) Für die *Reputation* wurden je vier Suchergebnisse einer bekannten Nachrichtenseite (hohe Reputation) und je vier einem unbekanntem Blog oder einer privaten Website (geringe Reputation) zugeordnet. Ersichtlich war dies über URL, Nennung des Seitennamens in der Überschrift sowie eine Logo-Einblendung neben dem Suchergebnis.
- (2) Hinsichtlich der *Neutralität* konnten die Suchergebnisse das jeweilige Thema entweder zweiseitig (hohe Neutralität) oder einseitig (geringe Neutralität) darstellen. Ersichtlich war dies über eindeutige Formulierungen („Was für und was gegen die Mietpreisbremse spricht“, „Wer bremst die Mietpreisbremse?“) in Überschrift und Snippet.
- (3) Schließlich wurde neben jedem Suchergebnis ein *soziales Rating* („xxx Nutzer empfehlen das“) angezeigt, das entweder im dreistelligen (hohes Rating) oder im niedrigen zweistelligen Bereich (geringes Rating) angesiedelt war.

---

4 *DuckDuckGo* bietet eigentlich eine endlos scrollende Ergebnisliste, für diese Studie wurde daher eine Paginierung ergänzt. Auch die zweite Ergebnisseite ließ sich aufrufen, bot aber nur noch irrelevante Ergebnisse. Wollte der Nutzer zu einer anderen Ergebnisseite navigieren oder die Suchanfrage verändern, wurde eine Fehlerseite eingeblendet, die die Rückkehr zur Ergebnisliste ermöglichte.



Abbildung 1: Eingebundene Ergebnisseite mit (1) rudimentärer Browser-Steuerung und Countdown, (2) voreingestellter Suchanfrage, (3) Werbung und irrelevanten Suchergebnissen und (4) manipulierten Suchergebnissen.



Bei der Erstellung der Zielseiten dienten tatsächlich existierende Nachrichtenseiten und Blogs als Vorlage. Diese wurden mit eigenen Artikeltexten versehen. Dabei wurde darauf geachtet, dass die acht unterschiedlichen Seiten bzw. Texte jeweils den gleichen Informationsgehalt aufwiesen, um mögliche Ausstrahlungseffekte zu vermeiden.

Die Themen für die Rechercheaufgaben waren die *Mietpreisbremse* (Suchanfrage im Wortlaut: „auswirkungen mietpreisbremse“; im Folgenden kurz: *Miete*) und *Rechtslage von Streaming und Download von TV-Serien* („serien streaming download rechtslage“; *Streaming*). Beide Themen wurden aufgrund von Überlegungen über die externe Validität der Ergebnisse ausgewählt. Illegale Streamingangebote werden von einer Mehrheit der Studierenden selbst genutzt (in unserer Stichprobe gaben 63,2 Prozent an, solche Portale zu nutzen). Das Thema Mietpreise bzw. deren Begrenzung spielt in der Stadt, in der die Studie durchgeführt wurde, ebenfalls eine wichtige Rolle.

### 3.2 Erhobene Merkmale

#### *Such- und Informationsverhalten*

Alle Aktionen der Teilnehmer wurden per Bildschirmaufnahmesoftware aufgezeichnet und von drei studentischen Mitarbeiterinnen codiert. Hierbei wurden die ausgewählten Suchergebnisse sowie die Auswahlreihenfolge erfasst. Außerdem wurde die Zeitspanne bis zum ersten Klick codiert.

### Vorwissen

Das Vorwissen wurde pro Thema anhand von vier Items erfasst (fünf-stufige Likert-Skala von 1 = *trifft überhaupt nicht zu* bis 5 = *trifft voll und ganz zu*): „Ich habe mich bereits umfassend mit diesem Thema auseinandergesetzt“, „Ich habe mir bereits eine recht feste Meinung zu diesem Thema gebildet“, „Ich kenne die wichtigsten Fakten zu diesem Thema“ und „Ich könnte andere in einer Diskussion von meinem Standpunkt zu diesem Thema überzeugen“ (Miete:  $\alpha = ,86$ ;  $M = 2,16$ ;  $SD = 0,88$ ; Streaming:  $\alpha = ,90$ ;  $M = 2,70$ ;  $SD = 1,03$ ).

### Auswahlstrategien

Der allgemeine Nutzungsstil von Suchmaschinen wurde anhand einer gekürzten Skala nach Beiler (2005) abgefragt, die per Hauptkomponentenanalyse<sup>5</sup> auf die drei Faktoren *Rangfolgen-Auswahl* (Beispielitem: „Ich entscheide mich für Suchergebnisse, die sehr weit oben stehen“), *spontan-intuitive Auswahl* (z. B. „Ich klicke spontan auf ein Suchergebnis“) und *evaluative Auswahl* (z. B. „Ich überlege mir gründlich, welche Suchergebnisse ich anklicke“) reduziert wurde.

### Mediennutzung

Die Nutzung aller sechs in den Ergebnislisten verlinkten Nachrichtenseiten ( $\alpha = ,75$ ;  $M = 2,59$ ;  $SD = 0,81$ ) und von „Blogs“ ( $M = 1,93$ ;  $SD = 1,18$ ) wurden mit der Frage „Wie häufig nutzen Sie die folgenden Internetangebote, um sich über aktuelle Ereignisse zu informieren?“ auf einer fünf-stufigen Skala von 1 (*gar nicht*) bis 5 (*täglich*) erfasst.

### Internet- und Suchmaschinennutzung

Die tägliche Internetnutzung wurde anhand einer sechs-stufigen Skala von 1 (*weniger als 30 Minuten*) bis 6 (*mehr als sechs Stunden*) erfasst. 74 Prozent der Teilnehmer nutzen das Internet mehr als zwei Stunden am Tag. Die Nutzung von Suchmaschinen wurde anhand einer fünf-stufigen Skala von 1 (*weniger als einmal wöchentlich*) bis 5 (*mehrmals am Tag*) erfasst. 92 Prozent der Teilnehmer nutzen Suchmaschinen wenigstens täglich.

### 3.3 Stichprobe

Als Teilnehmer wurden Studierende rekrutiert. Insgesamt nahmen 250 Personen an der Studie teil. Von diesen wurden drei aufgrund zu kurzer Verweildauer ausgeschlossen. In der finalen Stichprobe sind daher 247 Studierende, davon 171 (69,2 %) weiblich. Das Durchschnittsalter beträgt 22,00 Jahre ( $Mdn = 21$ ;  $SD = 3,50$ ). Die Teilnehmer wurden per Urnenziehung gleichmäßig auf eine der acht Gruppen verteilt. Die Gruppen unterscheiden sich weder nach Geschlecht,  $\chi^2(7, n = 247) = 5,49$ ;  $p = ,60$ ; noch nach Alter,  $F(7, 239) = 1,26$ ;  $p = ,27$ . Bis auf einen Teilnehmer nutzen alle Probanden Suchmaschinen mindestens mehrfach wöchentlich, die Suchmaschine *DuckDuckGo* war nur etwa jedem zehnten Teilnehmer bekannt.

5 Promax-Rotation; kumulativ erklärte Varianz = 52,5 %, KMO = ,66; Bartlett-Test auf Sphärität:  $\chi^2(36) = 235,54$ ,  $p < ,001$ .

## 4. Ergebnisse

### 4.1 Bearbeitung der Rechercheaufgaben

Bei der Auswertung und Ergebnisdarstellung wurden jeweils nur die für uns relevanten Suchergebnisse berücksichtigt. Klicks auf die Werbung oder den Übersetzungsdienst gingen nicht in die Analysen mit ein. Die Teilnehmer nahmen sich durchaus einen kurzen Moment Zeit, um die Ergebnisliste zu betrachten, bevor sie ihr erstes Suchergebnis auswählten und somit die erste Zielseite öffneten (erster Klick nach Sekunden: Miete:  $M = 14,66$ ;  $SD = 11,94$ ; Streaming:  $M = 9,35$ ;  $SD = 7,56$ ). Ein Großteil der Studierenden nutzte die maximale Zeit von fünf Minuten pro Rechercheaufgabe komplett aus (Bearbeitungsdauer in Sekunden: Miete:  $M = 280,24$ ;  $SD = 52,71$ ; Streaming:  $M = 273,95$ ;  $SD = 55,52$ ). Dabei wählten sie pro Aufgabe im Durchschnitt drei Suchergebnisse aus (Miete:  $M = 3,18$ ;  $SD = 1,16$ ; Streaming:  $M = 2,89$ ;  $SD = 1,05$ ). Die Ergebnisse sind hinsichtlich der Anzahl der aufgerufenen Suchergebnisse mit denen von Studien vergleichbar, die auf Basis von Logfile-Analysen das Selektionsverhalten bei der Nutzung von Suchmaschinen untersuchen (Machill u. a. 2008: 603); die Zeitspanne bis zum ersten Klick liegt im Bereich der Ergebnisse vergleichbarer Untersuchungen (vgl. Tremel 2010: 233; siehe hierzu auch Kap. 4.4). Die Unterschiede zwischen beiden Rechercheaufgaben dürften in erster Linie auf die Reihenfolge zurückzuführen sein. Da allen Teilnehmern zunächst das Thema Miete präsentiert wurde, benötigten sie bei dieser Rechercheaufgabe einige Sekunden länger, um sich für das erste Suchergebnis zu entscheiden. Tabelle 1 liefert einen Überblick darüber, um welche Suchergebnisse es sich dabei handelte. Dabei wird zwischen dem ersten Klick pro Rechercheaufgabe und den folgenden Selektionsentscheidungen unterschieden.

Tabelle 1: Klickhäufigkeit nach Faktoren und Faktorkombinationen

Faktor- kombination Rep / Neu / SoE	Anteil an Klicks in Prozent				
	Rechercheaufgabe Miete		Rechercheaufgabe Streaming		Gesamt (n = 1.500)
	Erster Klick (n = 247)	Folgende Klicks (n = 538)	Erster Klick (n = 247)	Folgende Klicks (n = 468)	
+ / + / +	34,0	18,0	19,4	16,7	20,5
+ / + / -	18,2	15,8	27,1	16,7	18,4
+ / - / +	11,7	14,1	22,3	15,2	12,9
+ / - / -	13,4	17,3	6,9	15,6	16,9
- / + / +	7,3	10,4	7,7	12,2	10,0
- / + / -	4,9	9,7	5,7	8,1	7,8
- / - / -	5,3	7,8	7,7	9,2	7,8
- / - / +	5,3	6,9	3,2	6,4	5,8
Faktoren					
Hohe Reputation	77,3	65,2	75,7	64,1	68,6
Hohe Neutralität	64,4	53,9	59,9	53,6	56,5
Hohe Anzahl SoE	58,3	49,4	52,6	50,9	51,9

Anmerkung: Rep = Reputation; Neu = Neutralität; SoE = Soziale Empfehlung. Prozentwerte über 100 sind rundungsbedingt.

Quellen mit hoher Reputation werden häufiger ausgewählt als solche mit niedriger Reputation. Vor allem beim ersten Klick entscheiden sich mehr als drei Viertel der Teilnehmer für ein Suchergebnis mit hoher Reputation (Miete: 77,3 %; Streaming: 75,7 %), von den folgenden Klicks entfallen rund zwei Drittel auf solche Quellen. Auch neutrale Suchergebnisse werden häufiger ausgewählt als solche, deren Snippets auf eine einseitige Beschäftigung mit dem Thema verweisen. Wiederum tritt dieser Effekt besonders bei der ersten Entscheidung auf (64,4 bzw. 59,9 % der Klicks auf Suchergebnisse mit hoher Neutralität).

Suchergebnisse mit einer hohen Anzahl an sozialen Empfehlungen werden dagegen lediglich beim Thema Miete und auch dort nur bei der ersten Selektion bedeutend häufiger ausgewählt (58,3 %) als Suchergebnisse mit einer geringen Anzahl; bei den folgenden Selektionsentscheidungen der Miete-Rechercheaufgabe sowie bei der Streaming-Aufgabe generell entfallen jeweils rund die Hälfte der Klicks auf Suchergebnisse mit hoher bzw. niedriger Anzahl.

Auf den ersten Blick entscheiden sich die Teilnehmer also häufiger für solche Suchergebnisse, bei denen zwei oder alle drei Faktoren „positive“ Ausprägungen annehmen. Allerdings könnte man die Zahlen auch dahingehend interpretieren, dass alleine eine hohe Reputation bereits eine ausreichende Bedingung für eine höhere Selektionswahrscheinlichkeit darstellt.

#### 4.2 Einflussfaktoren auf die Selektionsentscheidung

Um den Einfluss der unterschiedlichen Glaubwürdigkeitsindikatoren insgesamt analysieren zu können, war eine Umstrukturierung der Daten notwendig. Jeder Teilnehmer wurde mit einer Ergebnisseite konfrontiert, innerhalb derer jede der acht möglichen Faktorkombinationen durch jeweils *ein* Suchergebnis repräsentiert war. Die Ergebnislisten unterschieden sich lediglich in der Reihenfolge der Suchergebnisse. Dies bedeutet, dass jede Entscheidung für ein Suchergebnis gleichzeitig eine Entscheidung *gegen* die übrigen sieben Suchergebnisse darstellt. Somit müssen bei der Analyse des Einflusses der unterschiedlichen Merkmale auf die Selektionsentscheidung auch die *nicht* ausgewählten Suchergebnisse (und deren Merkmale) berücksichtigt werden.

Die Umstrukturierung der Daten auf Selektionsentscheidungsebene führt zu einer dichotomen abhängigen Variablen, die die Komplementärereignisse „Suchergebnis selektiert“ (1) und „Suchergebnis nicht selektiert“ (0) unterscheidet. Es wurde daher für die erste Selektionsentscheidung und für die folgenden Selektionsentscheidungen jeweils ein hierarchisch-logistisches Regressionsmodell berechnet, in das schrittweise erstens die Platzierung des Suchergebnisses, zweitens die drei Glaubwürdigkeitsindikatoren und schließlich drittens Interaktionseffekte zwischen den drei Indikatoren als Prädiktoren aufgenommen wurden (siehe Tab. 2). Als Kennwerte werden die Regressionskoeffizienten  $b$  und die Wald-Statistik  $W$  betrachtet.<sup>6</sup> Zur Beurteilung der Modellgüte wird Nagelkerkes  $R^2$  herangezogen.

6 Die logistische Regression nimmt eine nichtlineare Modellierung der Wahrscheinlichkeiten vor, weshalb sich an den Regressionskoeffizienten lediglich die Richtung des Effekts ablesen lässt. Um den alleinigen Beitrag der betrachteten Prädiktoren zum Gesamtmodell angeben zu können, wird daher auf die Wald-Statistik zurückgegriffen, die – vergleichbar mit dem t-Test in linearen Regressionsmodellen – testet, ob sich der Regressionskoeffizient signifikant von Null unterscheidet ( $w = \frac{b}{\frac{b}{\sqrt{b}}}$ ) und gegen die  $\chi^2$ -Verteilung ( $df = 1$ ) geprüft werden kann (Backhaus, Erichson, Plinke & Weiber 2011: 280; Best & Wolf 2010: 831).

Das Gesamtmodell zur ersten Selektionsentscheidung weist einen guten Fit auf ( $R^2 = ,40$ ; vgl. Backhaus, Erichson, Plinke & Weiber 2011: 276). Der Platzierung eines Suchergebnisses kommt der größte Einfluss zu ( $b = -0,71$ ;  $W = 413,08$ ;  $p < ,001$ ): Je weiter unten ein Suchergebnis in der Ergebnisliste steht, desto geringer ist die Selektionswahrscheinlichkeit. Unter den drei Glaubwürdigkeitsindikatoren hat lediglich die Reputation einen erwähnenswerten Einfluss auf die Selektionsentscheidung, eine hohe Reputation erhöht ebenfalls die Auswahlwahrscheinlichkeit ( $b = 0,98$ ;  $W = 19,23$ ;  $p < ,001$ ). Weiterhin zeigt sich im Gesamtmodell ein Interaktionseffekt zwischen der Reputation eines Suchergebnisses und dessen Neutralität ( $b = 0,77$ ;  $W = 9,61$ ;  $p < ,01$ ), der den in Schritt 2 noch vorhandenen alleinigen Einfluss der Neutralität überlagert. Der Interaktionseffekt lässt sich daher dahingehend interpretieren, dass Nutzer nur von Newssites (hohe Reputation) eine ausgewogene Darstellung erwarten, während sie beim Lesen eines Blogs sogar eher eine eindeutige Positionierung bevorzugen (Johnson & Kaye 2004: 633). Insgesamt können durch das Modell 89,3 Prozent der Suchergebnisse (32,6 % der geklickten und 97,5 % der nicht geklickten) während der ersten Selektionsentscheidung richtig klassifiziert werden (geklickt/nicht geklickt).<sup>7</sup>

Die beiden Modelle in Tabelle 2 unterscheiden sich in zweierlei Hinsicht. Zunächst steht jeder Teilnehmer im ersten Modell für zwei Selektionsentscheidungen: der jeweils erste Klick zum Thema *Miete* und zum Thema *Streaming*. Im Modell zu den folgenden Entscheidungen „liefern“ die Teilnehmer dagegen eine unterschiedliche Anzahl von Fällen, da sie unterschiedlich viele Suchergebnisse aufgerufen haben. Weiterhin sollten bei der Analyse der zweiten und folgenden Selektionsentscheidung(en) die vorangegangenen Klicks berücksichtigt werden. Die Logik, dass eine Entscheidung für ein Suchergebnis zugleich die Entscheidung gegen *sieben* andere darstellt, trifft nur im ersten Modell zu. Wir haben daher für die Analyse der folgenden Selektionsentscheidungen alle bereits aufgerufenen Suchergebnisse nicht mehr berücksichtigt. Ein Beispiel soll dies verdeutlichen: Hat ein Teilnehmer beim ersten und zweiten Klick jeweils bereits ein Suchergebnis mit hoher Reputation ausgewählt, stehen ihm nur noch zwei Suchergebnisse mit hoher und vier mit niedriger Reputation zur Verfügung (außer er ruft ein Suchergebnis wiederholt auf). Einer Entscheidung für ein bestimmtes Suchergebnis im dritten Klick werden daher nur noch die fünf *bislang nicht ausgewählten* Suchergebnisse zugeordnet.

Wie schon beim ersten Klick präferieren die Teilnehmer auch bei einer der folgenden Entscheidungen eines der oberen (bislang nicht ausgewählten) Suchergebnisse. Da in den meisten Fällen bei der ersten Selektionsentscheidung ein oben platziertes Ergebnis ausgewählt wird, sinkt der Einfluss der Platzierung jedoch entsprechend, bleibt aber der stärkste Prädiktor ( $b = -0,20$ ;  $W = 134,84$ ;  $p < ,001$ ). Der Einfluss der Reputation steigt hingegen: Auch bei den folgenden Entscheidungen wird ein Suchergebnis mit hoher Reputation eher ausgewählt als eines mit geringer Reputation ( $b = -0,96$ ;  $W = 53,86$ ;  $p < ,001$ ).

Neutralität und soziale Empfehlungen erweisen sich einzeln betrachtet erneut nicht als relevante Prädiktoren. Auch der bei der ersten Selektionsentscheidung beobachtete Interaktionseffekt zwischen Neutralität und Reputation verschwindet. Allerdings zeigt sich ein Interaktionseffekt zwischen der Neutralität und der Anzahl sozialer Empfehlungen ( $b = -0,44$ ;  $W = 9,24$ ;  $p < ,001$ ). Die Güte des Gesamtmodells sinkt allerdings im Vergleich zum ersten Modell beträchtlich ( $R^2 = ,10$ ). Dies zeigt sich nicht zuletzt daran,

7 32,6 Prozent richtige Klassifikationen mögen auf den ersten Blick gering erscheinen; der Wert liegt aber deutlich über dem Erwartungswert von 12,5 Prozent, der sich bei der Selektion einer von insgesamt acht Optionen ergibt.

dass im zweiten Modell in keinem Fall die anhand der Prädiktoren vorhergesagte Selektionswahrscheinlichkeit 50 Prozent überschreitet, also alle Suchergebnisse als nicht geklickt klassifiziert werden. Dies lässt darauf schließen, dass nach einer bewussten ersten Selektion weitere Suchergebnisse eher zufällig ausgewählt werden.

*Tabelle 2: Hierarchisch-logistische Regressionen: Einflüsse auf die Selektionsentscheidungen*

Modell	Erste Selektionsentscheidung			Folgende Selektionsentscheidungen		
	<i>b</i>	<i>W</i>	R <sup>2</sup>	<i>b</i>	<i>W</i>	R <sup>2</sup>
<b>Schritt 1</b>						
Platzierung <sup>a</sup>	-0,71	409,81***		-0,16	102,56***	
Konstante	0,40	15,49***		-0,95	151,36***	
Nagelkerke-R <sup>2</sup>			,30			,03
<b>Schritt 2</b>						
Platzierung <sup>a</sup>	-0,77	414,71***		-0,19	134,32***	
Reputation <sup>b</sup>	1,66	175,98***		1,09	212,96***	
Neutralität <sup>b</sup>	0,75	41,83***		0,37	25,53*	
Soziale Empfehlung <sup>b</sup>	0,34	9,10**		0,12	2,73	
Konstante	-1,02	42,60***		-1,63	260,96***	
Nagelkerke-R <sup>2</sup>			,40			,09
<b>Schritt 3</b>						
Platzierung <sup>a</sup>	-0,78	413,08***		-0,20	134,84***	
Reputation <sup>b</sup>	0,98	19,23***		0,95	53,86***	
Neutralität <sup>b</sup>	0,24	1,10		0,41	0,09	
Soziale Empfehlung <sup>b</sup>	0,05	0,04		-0,16	1,27	
Rep × Neu	0,77	9,61**		0,18	1,40	
Rep × SoE	0,47	3,65		0,07	0,22	
Neu × SoE	-0,05	0,04		0,44	9,24*	
Konstante	-0,54	6,83**		-1,41	113,75***	
Nagelkerke-R <sup>2</sup>			,40			,10

*Anmerkung:* *b* = Regressionskoeffizient; *W* = Wald-Statistik (df = 1); Rep = Reputation; Neu = Neutralität; SoE = Soziale Empfehlung.

*n* = 3.984 (erste Selektionsentscheidung), 6.256 (folgende Selektionsentscheidungen).

<sup>a</sup>Platzierung in der Ergebnisliste, 1-8, von oben nach unten. <sup>b</sup>Kategoriale Variable, 0/1-codiert (1 = hohe Ausprägung).

\*  $p < ,05$ ; \*\*  $p < ,01$ ; \*\*\*  $p < ,001$ .

Insgesamt bestätigen unsere Analysen den Einfluss der Platzierung (H1) und der Reputation (H2) auf die Selektionsentscheidung. Ein Effekt der Neutralität (H3) oder der sozialen Empfehlungen (H4) konnte dagegen nicht festgestellt werden. Im Hinblick auf einen möglichen additiven Effekt beim Vorliegen mehrerer Glaubwürdigkeitsindikatoren (FF1) zeigen sich nur schwache Interaktionseffekte zwischen der Neutralität und der Reputation (erste Selektionsentscheidung) bzw. sozialen Empfehlungen (folgende Selektionsentscheidungen).



### 4.3 Interindividuelle Unterschiede in der Präferenz für Glaubwürdigkeitsindikatoren

Bislang standen die Selektionsentscheidungen und der Einfluss der manipulierten Glaubwürdigkeitsindikatoren der Suchergebnisse im Vordergrund. Unsere zweite Forschungsfrage zielt dagegen auf den Einfluss unterschiedlicher *Personenmerkmale* auf die Präferenz für Suchergebnisse etwa mit hoher Reputation ab. Die Frage, der wir abschließend nachgehen wollen, ist daher: Lassen sich die Präferenzen für Ergebnisse mit hoher Reputation, hoher Neutralität und einer hohen Anzahl an sozialen Empfehlungen durch Merkmale auf Personenebene erklären?

Zur Beantwortung untersuchen wir nun nicht mehr den Vergleich von ausgewählten und nicht ausgewählten Suchergebnissen, sondern konzentrieren uns nur auf die tatsächlich ausgewählten Suchergebnisse. Da die vorherigen Ergebnisse für eine eher zufallsgeleitete Auswahl bei den späteren Selektionsentscheidungen sprechen, betrachten wir hier außerdem nur die erste Selektionsentscheidung. Für jeden Glaubwürdigkeitsindikator werden dabei die Komplementärereignisse „Suchergebnis mit hoher Ausprägung selektiert“ (1) und „Suchergebnis mit geringer Ausprägung selektiert“ (0) betrachtet. Es wurden daher drei logistische Regressionsmodelle gerechnet, die sich jeweils auf einen der drei Glaubwürdigkeitsindikatoren beziehen (siehe Tabelle 3).

Tabelle 3: Logistische Regressionen: Einflüsse auf die Präferenz für die Glaubwürdigkeitsindikatoren

Modell	Reputation		Neutralität		Soziale Empfehlung	
	<i>b</i>	<i>W</i>	<i>b</i>	<i>W</i>	<i>b</i>	<i>W</i>
Internetnutzung <sup>a</sup>	-0,05	0,24	0,04	0,15	-0,02	0,03
Suchmaschinennutzung <sup>b</sup>	0,05	0,09	0,04	0,06	-0,09	0,34
FA Rangfolgen-Selektion <sup>c</sup>	-0,13	1,35	0,09	0,74	-0,11	1,34
FA Spontane Selektion <sup>c</sup>	0,06	0,26	0,10	1,02	-0,05	0,25
FA Evaluative Selektion <sup>c</sup>	0,17	1,83	0,05	0,21	0,02	0,03
Nutzung Blogs <sup>b</sup>	0,00	0,00	0,08	0,78	-0,05	0,44
Nutzung trad. Nachrichtenseiten <sup>b</sup>	0,42	7,86*	0,17	1,71	0,20	2,40
Vorwissen <sup>b</sup>	-0,16	2,02	-0,03	0,09	0,00	0,00
Konstante	0,52		0,74	0,18	0,32	0,20
Nagelkerke-R <sup>2</sup>	,05		,02		,02	

Anmerkung: *b* = Regressionskoeffizient; *W* = Wald-Statistik (df = 1).

*n* = 494; <sup>a</sup>6-stufige Skala, höhere Werte entsprechen höherer Nutzung; <sup>b</sup>5-stufige Skala, höhere Werte entsprechen höherer/m Nutzung/Wissen; <sup>c</sup>Normalverteilte Faktoren, höhere Werte entsprechen höherer Tendenz zur jeweiligen Auswahlstrategie.

\* *p* < ,05.

Die Modelle weisen allesamt nur eine sehr geringe Güte auf (Reputation:  $R^2 = ,05$ ; Neutralität:  $R^2 = ,02$ ; soziale Empfehlung:  $R^2 = ,02$ ), die Präferenz der Studienteilnehmer für ein Suchergebnis mit hoher Reputation, hoher Neutralität oder hoher Anzahl sozialer Empfehlungen kann also durch die Prädiktoren nicht erklärt werden. Nur der am ehesten zu erwartende Zusammenhang tritt schwach zutage: Wer angab, generell häufig News-sites wie *SpiegelOnline* oder *Süddeutsche.de* zu nutzen, entscheidet sich mit höherer Wahrscheinlichkeit für ein Suchergebnis mit hoher Reputation ( $b = 0,42$ ;  $W = 7,76$ ;  $p < ,05$ ). Die Nutzungshäufigkeit von Blogs hat dagegen keinen Einfluss auf die Auswahl.

Auch die übrigen Merkmale entfalten in keinem der Modelle eine Wirkung. So hätten wir etwa einen Zusammenhang zwischen dem Nutzungsstil *spontane Selektion* und einer Präferenz für Suchergebnisse mit hoher Reputation aufgrund der Orientierung an optischen Reizen wie den bekannten Logos der Nachrichtenseiten, und zwischen dem Nutzungsstil *evaluative Selektion* und der Präferenz für neutrale Ergebnisse erwartet (vgl. Kapitel 3.3). Die erhobenen Daten können diese Zusammenhänge jedoch nicht bestätigen.

## 5. Diskussion und Ausblick

Unsere Ergebnisse lassen sich zunächst als Beleg für die Bedeutung von Suchmaschinen für die Sichtbarkeit von Informationen im Online-Bereich interpretieren. Die Studienteilnehmer folgten zu großen Teilen der vorgeschlagenen Reihenfolge. Diese Übereinstimmung mit vorliegenden Ergebnissen (z. B. Agichtein u. a. 2006: 4; Pan u. a. 2007: 814) kann als Indiz für die externe Validität der Ergebnisse gewertet werden. Unter den Glaubwürdigkeitsindikatoren beeinflusste lediglich die Reputation der Quelle die Selektionswahrscheinlichkeit, wenn auch in deutlich geringerem Maße als die Platzierung der Suchergebnisse. Diese Befunde lassen sich auf unterschiedliche Weise interpretieren.

Erstens können sie als Beleg für die untergeordnete Bedeutung der Glaubwürdigkeit bei der Selektion von Suchergebnissen interpretiert werden. Dieser Lesart folgend, wären sich die Rezipienten nach wie vor nicht bewusst, dass die Beurteilung der Qualität von Informationen im Online-Bereich in ihrer eigenen Verantwortung (Metzger u. a. 2003: 294) liegt. Sie folgen vielmehr unreflektiert den Vorgaben der Suchmaschine und lassen dabei Glaubwürdigkeitsindikatoren, wie etwa die Neutralität oder soziale Empfehlungen, größtenteils außer Acht.

Zweitens lassen sich die Ergebnisse dahingehend interpretieren, dass Nutzer(innen) von Suchmaschinen zwar sehr wohl Überlegungen über die Glaubwürdigkeit von Informationen und Quellen anstellen, dabei aber andere Indikatoren (als die in dieser Studie variierten) berücksichtigen bzw. diese unterschiedlich gewichten. Bei den Studienteilnehmern handelte es sich größtenteils um erfahrene Suchmaschinennutzer, die solche Dienste bereits seit Jahren regelmäßig nutzen. Es besteht die Möglichkeit, dass sie im Laufe der Zeit ein gewisses Vertrauen in die vorgegebene Platzierung entwickelt haben (vgl. auch Jansen & Spink 2006). Dies würde bedeuten, dass eine hohe Platzierung als Indikator für hohe Glaubwürdigkeit wahrgenommen werden könnte (Westerwick 2013: 205). Nicht zuletzt versuchen Suchmaschinenanbieter zunehmend, auch automatisierte Glaubwürdigkeitsattributionen in die Rankings einfließen zu lassen (Lewandowski 2012: 134; siehe auch Dong u. a. 2015).

Auch der Einfluss der Neutralität eines Suchergebnisses ist differenziert zu betrachten. Die Neutralität hat zumindest dann einen Einfluss auf die Selektionsentscheidung, wenn es sich bei dem „unparteiischen“ Suchergebnis gleichzeitig um eine Quelle mit hoher Reputation handelt. Unter den Quellen mit niedriger Reputation (Blogs) werden Suchergebnisse, die eine neutrale Auseinandersetzung mit einem Thema suggerieren, nicht häufiger ausgewählt als solche, bei denen dies nicht der Fall ist. Dies deutet darauf hin, dass Nutzer bei unterschiedlichen Typen von Quellen auch unterschiedliche Kriterien anlegen. Von Blogs erwarten sie evtl. keine Neutralität, von klassischen Nachrichtensites dagegen schon (vgl. Johnson & Kaye 2004: 630–631). Dass der Neutralität insgesamt aber eine geringe Bedeutung zukommt, dürfte nicht zuletzt auf die Operationalisierung dieses Glaubwürdigkeitsindikatoren über das Snippet zurückzuführen sein. Diese werden von den Nutzern eher nach Stichwörtern abgesucht als vollständig gelesen, die Haltung zum Thema kann so kaum erfasst werden (Lorigo u. a. 2008: 1044).

Schließlich hatte eine hohe Anzahl sozialer Empfehlungen, entgegen der Erwartung auf Basis vorliegender Befunde, keinen messbaren Einfluss auf die Selektionsentscheidung. Allerdings wurde der positive Einfluss der Empfehlungen zuvorderst in Hinblick auf Produktauswahl, z. B. bei Unterhaltungselektronik (Park u. a. 2007), analysiert. In diesem Kontext sind die Bewertungen anderer Nutzer bereits seit längerer Zeit gebräuchlich und somit auch unseren Studienteilnehmern bekannt. Die Anzeige von Bewertungen einzelner Suchresultate bei Suchmaschinen ist dagegen bislang unüblich. Der in unserer Studie fehlende Einfluss sozialer Empfehlungen ist somit nicht zwangsläufig ein Hinweis auf deren Wirkungslosigkeit.

Drittens wurden in unserer Studie die Merkmale der Situation und des Kontextes konstant gehalten. Unter Umständen waren die Instruktion der Teilnehmer und die Themen der Rechercheaufgaben nicht dazu geeignet, Bedingungen zu schaffen, unter denen der Beurteilung der Glaubwürdigkeit eine der Realität entsprechende Bedeutung zukommt. Dies könnte allerdings sowohl zu einer Unter- wie zu einer Überschätzung der Bedeutung der Glaubwürdigkeitsindikatoren führen. Auf der einen Seite stellt die Selektion von Suchergebnissen im monetären Sinn immer eine Niedrigkostensituation dar. Im Normalfall haben die Nutzer aber ein Interesse daran, relevante Informationen in möglichst kurzer Zeit aufzufinden. Kaum jemand wird daher, außer aus Neugier, bei der Suche nach Informationen ein Suchergebnis öffnen, von dem keine glaubwürdigen Informationen zu erwarten sind, wenn gleichzeitig als kompetenter oder vertrauenswürdiger eingeschätzte Alternativen vorliegen. Auch sind die Themen (Mietpreisbremse, Legalität von Streaming) zwar für Studierende direkt erfahrbar (obtrusive). Dennoch dürften nur wenige Teilnehmer tatsächlich hoch involviert gewesen sein. Mit dem eher niedrigen Involvement würde auch eine geringere Wahrscheinlichkeit einer elaborierten Informationsverarbeitung einhergehen. Niedriges Involvement und Habitualisierung begünstigen eine Auswahl, die vorhandene Informationen nicht oder nicht vollständig berücksichtigt und sich stattdessen einfacher Heuristiken, wie eben der Selektion des ersten Suchergebnisses, bedient (Metzger 2007: 2088; Wirth & Schweiger 1999: 58). Auf der anderen Seite könnte unsere Studie die Bedeutung der Glaubwürdigkeitsindikatoren allerdings auch überschätzen lassen. Zwar war den Teilnehmern nicht bewusst, dass ihre Selektionsentscheidungen aufgezeichnet werden. Sicherlich wird ein Teil von ihnen aber über den Sinn der Untersuchung nachgedacht haben. Andere waren unter Umständen bestrebt, die gestellte Aufgabe möglichst gut zu bearbeiten. Die relativ lange Zeitspanne bis zum ersten Klick legt somit die Vermutung nahe, dass die Teilnehmer die Snippets sogar aufmerksamer wahrgenommen haben als dies in einer realen Suchsituation der Fall wäre. Allerdings trifft diese Einschränkung der Übertragbarkeit unserer Studienergebnisse vor allem auf navigationsorientierte Anfragen zu, bei denen Nutzer lediglich schnell ein bekanntes Angebot auffinden wollen. Im Hinblick auf informationsorientierte Anfragen sollten die Ergebnisse, eine hinreichende Motivation der Nutzer vorausgesetzt (dazu gleich mehr), durchaus ein gewisses Maß an externer Validität besitzen.

Interessant ist viertens, dass die Selbstauskunft über den Nutzungsstil von Suchmaschinen scheinbar kaum geeignet ist, die tatsächliche Nutzung abzubilden. Diese Selbstauskünfte der Studienteilnehmer ließen unterschiedliche Nutzungsstrategien erwarten. Manche gaben an, relativ spontan Suchergebnisse auszuwählen, andere orientierten sich stärker an der vorgegebenen Reihenfolge. Es gab aber auch Teilnehmer, die von sich behaupteten, sich die Suchergebnisse genau anzuschauen und sich ihre Selektionsentscheidung gut zu überlegen. Unsere Ergebnisse zeigen allerdings, dass diese Selbstauskunft wenig mit dem tatsächlichen Selektionsverhalten (zumindest in der Erhebungssituation) zu tun hatte.

Zusammenfassend können die Ergebnisse aus zwei Perspektiven interpretiert werden. Erstens lassen sie den Schluss zu, dass der Beurteilung der Glaubwürdigkeit keine Bedeutung zukommt, wenn Bürger mittels Suchmaschinen nach Informationen suchen, bzw. dass sich Nutzer dabei nahezu ausschließlich an der Platzierung der Suchergebnisse orientieren. Zweitens ist durchaus denkbar, dass es unserer Studie an externer Validität fehlte, da die Situation, in der die Teilnehmer die Rechercheaufgaben bearbeiteten, zu weit von der Realität entfernt war. Es bestand für die Teilnehmer unter Umständen kein hinreichend großer Anreiz, nach verlässlichen Informationen zu suchen und somit den Glaubwürdigkeitsindikatoren (v. a. der Neutralität und den sozialen Empfehlungen) Beachtung zu schenken. Natürlich wird dies auch auf einen Teil der Nutzungsepisoden von Suchmaschinen im täglichen Leben zutreffen. Es sind aber durchaus Situationen denkbar, in denen Nutzer stärker motiviert sind, die Glaubwürdigkeit von Informationen schon vor der Selektion zu beurteilen. Sei es, weil sie ein hohes Themeninvolvement aufweisen, weil sie sich verlässlich über ein zu kaufendes Produkt informieren oder weil sie die Informationen an andere weitergeben wollen. Einige Autoren integrieren in ihre Überlegungen zur Rolle der Glaubwürdigkeit in der Online-Kommunikation Modelle, die – ähnlich dem ELM – unterschiedliche Modi der Informationsverarbeitung vorsehen (z. B. Metzger 2007; Reinhard & Sporer 2005). In einer Folgestudie manipulieren wir daher die Motivation der Teilnehmer experimentell, um den Einfluss dieses Merkmals auf die Wahrnehmung und Wirkung der Glaubwürdigkeitsindikatoren überprüfen zu können.

## Literatur

- Agichtein, E., Brill, E., Dumais, S. & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 3–10). New York, NY: ACM.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (Hrsg.) (2011). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (13., überarb. Aufl.). Berlin: Springer.
- Beiler, M. (2005). Selektionsverhalten in den Ergebnislisten von Suchmaschinen. Modellentwicklung und empirische Überprüfung am Beispiel von Google. In M. Machill & N. Schneider (Hrsg.), *Suchmaschinen: Neue Herausforderungen für die Medienpolitik* (S. 165–189). Berlin: Vistas.
- Bentele, G. (1988). Der Faktor Glaubwürdigkeit. Forschungsergebnisse und Fragen für die Sozialisationsperspektive. *Publizistik*, 33, 406–426.
- Best, H. & Wolf, C. (2010). Logistische Regression. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 827–854). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Broder, A. (2002). A taxonomy of web search. In *ACM SIGIR Forum* (Bd. 36: 3–10).
- Dong, X. L., Gabrilovich, E., Murphy, K., Dang, V., Horn, W., Lugaresi, C., ... Zhang, W. (2015). Knowledge-based trust: Estimating the trustworthiness of web sources. *arXiv:1502.03519*, <http://arxiv.org/abs/1502.03519> [30.06.2015].
- Dutta-Bergman, M. J. (2004). The impact of completeness and web use motivation on the credibility of e-health information. *Journal of Communication*, 54, 253–269.
- Flanagin, A. J. & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9, 319–342.
- Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., ... Treinen, M. (2001). What makes web sites credible? A report on a large quantitative study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 61–68). New York, NY: ACM.

- Greer, J. D. (2003). Evaluating the credibility of online information: A test of source and advertising influence. *Mass Communication and Society*, 6, 11–28.
- Hovland, C. L., Janis, I. L. & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*. Westport, CT: Greenwood Press.
- Hu, Y. & Sundar, S. (2009). Effects of online health sources on credibility and behavioral intentions. *Communication Research*, 37, 105–132.
- Jansen, B. J. & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, 42, 248–263.
- Johnson, K. A. & Wiedenbeck, S. (2009). Enhancing perceived credibility of citizen journalism web sites. *Journalism & Mass Communication Quarterly*, 86, 332–348.
- Johnson, T. J. & Kaye, B. K. (2004). Wag the blog: How reliance on traditional media and the Internet influence credibility perceptions of weblogs among blog users. *Journalism & Mass Communication Quarterly*, 81, 622–642.
- Kink, N. & Hess, T. (2008). Search engines as substitutes for traditional information sources? An investigation of media choice. *The Information Society*, 24, 18–29.
- Leino, J., Rähkä, K.-J. & Finnberg: (2011). All the news that's fit to read: Finding and recommending news online. In P. Campos, N. Graham, J. Jorge, P. Palanque, N. Nunes & M. Winckler (Hrsg.), *Human-Computer Interaction - INTERACT 2011* (Bd. 3: 169–186). Heidelberg: Springer.
- Lewandowski, D. (2012). Credibility in web search engines. In M. Folk & S. Apostel (Hrsg.), *Online credibility and digital ethos: Evaluating computer-mediated communication* (S. 131–146). Hershey, PA: IGI Global.
- Lewandowski, D. (2015). *Suchmaschinen verstehen*. Heidelberg: Springer Vieweg.
- Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., ... Pan, B. (2008). Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, 59, 1041–1052.
- Machill, M., Beiler, M. & Zenker, M. (2008). Search-engine research: A European-American overview and systematization of an interdisciplinary and international research field. *Media, Culture & Society*, 30, 591–608.
- Machill, M., Neuberger, C., Schweiger, W. & Wirth, W. (2004). Navigating the Internet: A study of German-language search engines. *European Journal of Communication*, 19, 321–347.
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58, 2078–2091.
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R. & McCann, R. M. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. In P. J. Kalbfleisch (Hrsg.), *Communication Yearbook* (Bd. 27: 293–335). New York, NY: Routledge.
- Metzger, M. J., Flanagin, A. J. & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60, 413–439.
- Meyer, H. K., Marchionni, D. & Thorson, E. (2010). The journalist behind the news: Credibility of straight, collaborative, opinionated, and blogged „news“. *American Behavioral Scientist*, 54, 100–119.
- Nakamura, S., Konishi, S., Jatowt, A., Ohshima, H., Kondo, H., Tezuka, T., ... Tanaka, K. (2007). Trustworthiness analysis of web search results. In *Research and advanced technology for digital libraries* (Bd. 4675: 38–49). Berlin: Springer.
- Nawratil, U. (2006). *Glaubwürdigkeit in der sozialen Kommunikation* (2. Aufl.). München: Open Access LMU, <http://epub.ub.uni-muenchen.de/941/> [30.06.2015].
- O'Keefe, D. J. (2008). Elaboration likelihood model. In W. Donsbach (Hrsg.), *The international encyclopedia of communication* (S. 1475–1480). Malden, MA: Blackwell Publishing.
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G. & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12, 801–823.
- Park, D.-H., Lee, J. & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. *International Journal of Electronic Commerce*, 11, 125–148.

- Reinhard, M.-A. & Sporer, L. (2005). Mehr als Worte: Glaubwürdigkeitsattribution anhand non-verbaler und inhaltlicher Informationen als Funktion der Urteilsrelevanz. *Zeitschrift für Sozialpsychologie*, 36, 21–32.
- Röhle, T. (2010). *Der Google-Komplex: Über Macht im Zeitalter des Internets*. Bielefeld: Transcript.
- Schmierbach, M. & Oeldorf-Hirsch, A. (2012). A little bird told me, so I didn't believe it: Twitter, credibility, and issue perceptions. *Communication Quarterly*, 60, 317–337.
- Schweiger, W. (2007). *Theorien der Mediennutzung. Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Self, C. C. (2009). Credibility. In D. W. Stacks & M. B. Salwen (Hrsg.), *An integrated approach to communication theory and research* (2. Aufl.: 435–456). New York, NY: Routledge.
- Sundar, S., Knobloch-Westerwick, S. & Hastall, M. R. (2007). News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology*, 58, 366–378.
- Terbeck, V. H. (2012). Der Einfluss sozialer Empfehlungen auf das Selektionsverhalten in Suchmaschinen: eine Eyetracking-Studie der Google-Ergebnisseite. *Info* 7, 27, 45–52.
- Tremel, A. (2010). *Suchen, finden – glauben? Die Rolle der Glaubwürdigkeit von Suchergebnissen bei der Nutzung von Suchmaschinen* (Doktorarbeit). Ludwig-Maximilians-Universität, München, <http://d-nb.info/1009952234/> [30.06.2015].
- Tseng, S. & Fogg, B. J. (1999). Credibility and computing technology. *Communications of the ACM*, 42, 39–44.
- van Eimeren, B. & Frees, B. (2014). 79 Prozent der Deutschen online – Zuwachs bei mobiler Internetnutzung und Bewegtbild: Ergebnisse der ARD/ZDF-Onlinestudie 2014. *Media Perspektiven*, 378–396.
- Wathen, C. N. & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the web. *Journal of the American Society for Information Science and Technology*, 53, 134–144.
- Westerwick, A. (2013). Effects of sponsorship, web site design, and google ranking on the credibility of online information. *Journal of Computer-Mediated Communication*, 18, 80–97.
- Winter, S. & Krämer, N. C. (2012). Selecting science information in Web 2.0: How source cues, message sidedness, and need for cognition influence users' exposure to blog posts. *Journal of Computer-Mediated Communication*, 18, 80–96.
- Wirth, W. & Brecht, M. (1999). Selektion und Rezeption im WWW: Eine Typologie. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet: empirische Analysen zu einem Schlüsselkonzept* (S. 149–180). Opladen: Westdeutscher Verlag.
- Wirth, W. & Schweiger, W. (1999). Selektion neu betrachtet: Auswahlentscheidungen im Internet. In W. Wirth & W. Schweiger (Hrsg.), *Selektion im Internet: empirische Analysen zu einem Schlüsselkonzept* (S. 43–74). Opladen: Westdeutscher Verlag.
- Xu, Q. (2013). Social recommendation, source credibility, and recency: Effects of news cues in a social bookmarking website. *Journalism & Mass Communication Quarterly*, 90, 757–775.
- Ziegele, M. & Weber, M. (2015). Example, please! Comparing the effects of single customer reviews and aggregate review scores on online shoppers' product evaluations. *Journal of Consumer Behaviour*, 14, 103–114.