Confirmation Bias in Information Search with Social Tags

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Mag. Stefan Schweiger

aus Graz/Österreich

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 04.09.2019

Dekan: Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatterin: Prof. Dr. Ulrike Cress

2. Berichterstatter: Prof. Dr. Tobias Ley

To my parents

Acknowledgements

Table of Contents

Deutsche Zusammenfassung

Der Bestätigungsfehler ist die Tendenz Informationen so auszuwählen und zu bewerten, dass bestehende Einstellungen verstärkt werden. In dieser Dissertation untersuche ich, ob der Bestätigungsfehler auch in Online-Umgebungen auftritt, wo Nutzer auf geteilte Inhalte über Social Tag Clouds zugreifen. Personen sollten bei der gesundheitsbezogenen Suche nicht dazu tendieren, dass sie ihr Selbstkonzept aufrechterhalten (Verteidigungsmotiv), sondern sie sollten Inhalte gemäß ihrer Korrektheit auswählen (Genauigkeitsmotiv). Um die Korrektheit der Inhalte einzuschätzen berücksichtigen Nutzer internale, *individuelle Bewertungen (Vorwissen, Voreinstellungen, Einstellungsgewissheit)*, sowie externale, *kollektive Reize (Tag Popularität, Glaubwürdigkeit)*. Dies ist eine bedeutsame Erweiterung zur Literatur Sozialer Tagging Systeme, die bisher ausschließlich den Einfluss semantischer Assoziationen im menschlichen Gedächtnis und den Einfluss aggregierter, kollektiver Assoziationen untersucht.

In den Studien 2 und 3 zeigte sich ein linearer Einfluss von *Voreinstellungen* auf die Auswahl und Bewertung von Blogeinträgen. Hier testete ich auch, ob der Einfluss der Voreinstellungen durch *Einstellungsgewissheit* moderiert würde. Einstellungsgewissheit beeinflusste die Auswahl von Blogeinträgen und deren Bewertung.

In den Studien 1 und 3 zeigte sich ein Effekt der Tag-Popularität auf die Auswahl von Tags, Blogeinträgen und die Bewertung von Inhalten. Außerdem zeigte sich in der Studierendenstichprobe (Studie 2), dass hohe Glaubwürdigkeit den Einfluss von Einstellungen auf die Navigation reduzierte. In Studie 3, mit einer repräsentativen Stichprobe, wurde dies jedoch nicht gefunden. In Bezug auf die Bewertung von Inhalten hatte die Glaubwürdigkeit keinen Einfluss in Studie 2, aber in Studie 3 bewertete die repräsentative Stichprobe von Suchenden bei hoher Glaubwürdigkeit und gleichzeitig geringer Einstellungsgewissheit Inhalte günstiger, wenn sie einstellungs-kongruent waren.

Somit zeigte sich, dass Informationssuchende in der Tag-basierten Navigation dem Bestätigungsfehler unterliegen. Dies ist eine bedeutende Erweiterung zur Literatur, die gezeigt hat, dass Informationssuche und Bewertung in Social-Tagging-Umgebungen semantischen Assoziationen folgen. Die vorliegenden Ergebnisse zeigen interessante Implikationen für die Gestaltung von Social-Tagging Plattformen. So ist es beispielsweise wichtig, die Zielgruppe zu berücksichtigen, da sie bestimmt, ob Suchende im gesundheitsbezogenen Kontext durch Genauigkeits- oder Verteidigungsmotive geleitet sind.

English Summary

Confirmation bias is the tendency of information searchers to select and evaluate information that supports pre-existing attitudes favourably. The current dissertation investigates whether confirmation bias affects health-related search in online environments, where users share content and social tag clouds are the navigation interface for searchers. I assumed that when individuals search health-related issues, they are motivated to find accurate information (accuracy motivation), in contrast to defending their self-concept (defense motivation). To determine what information is accurate, I expect that searchers attend to internal, *individual evaluations* (*prior knowledge*, *prior attitudes*, and *attitude confidence)*, and external, *collective cues* (*tag popularity* and *source credibility)*.

Regarding the influence of *individual evaluations*, in studies 2 and 3, a linear influence of prior attitudes on the selection of blog posts (but not tags), and the evaluation of blog posts was found. In studies 2 and 3, I tested whether the influence of prior attitudes was moderated by confidence. I found that high confidence did affect the selection of blog posts but not tags in both studies, and confidence influenced the evaluation of tag-related blog posts.

Regarding the influence of the *collective cues*, tag popularity was manipulated in studies 1 and 3, where I found a main effect of tag popularity on the selection of tags, blog posts, and evaluation of content, showing that tag size influenced confirmation bias in a moderate to strong way. In the student sample (study 2), I found that high credibility reduced the influence of prior attitudes on the selection of tags and consequently blog posts. However, using a representative sample (study 3), no influence of source credibility was found. With respect to the searchers' evaluation of content, credibility had no influence in study 2, but in study 3, under high source credibility and low attitude confidence, searchers evaluated content more favourably when content was attitude consistent.

In conclusion, the present dissertation shows that confirmation bias and individual evaluations guide information searchers in tag-based navigation, extending the literature which showed behaviour in social tagging environments follows semantic associations. The results are interesting for the construction of content aggregation or social tagging platforms, and practitioners who provide health-related online content. Practitioners and platform providers pay attention to their target audience, as this will either elicit accuracy or defense motivation. So, different strategies can be implemented when the aim is to reduce the influence of confirmation bias on information search behaviour.

Published Journal Articles

Publication 1 (Study 1)

Schweiger, S., Oeberst, A., & Cress, U. (2014) Confirmation bias in web-based search: a randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of Medical Internet Research*, *16*(3).

Publication 2 (Study 2)

Schweiger, S., & Cress, U. (2019). Attitude confidence and source credibility in information foraging with social tags. *PLOS ONE*, *14*(1).

Publication 3 (Study 3)

Schweiger, S., & Cress, U. (2019). How Confidence in Prior Attitudes, Social Tag Popularity, and Source Credibility Shape Confirmation Bias Toward Antidepressants and Psychotherapy in a Representative German Sample: Randomized Controlled Web-Based Study. *Journal of Medical Internet Research*, *21*(4), e11081.

Author Contribution

In the following, I will highlight my contribution to the journal articles presented in the current dissertation. I contributed to the design, data acquisition and data analysis, interpretation of results and publication of all three studies. For each published paper, I was the main author. Co-authors provided valuable feedback for the drafts, and I revised the drafts accordingly. For all three studies I constructed the study materials, organized the online experiments and analyzed the results. For the first study, the ideas were conceived together with Prof. Dr. Aileen Oeberst and Prof. Dr. Ulrike Cress as part of a DFG proposal. The scientific ideas, study design, and interpretation of results were conceived collaboratively with both co-authors. The main draft of all three publications were written by me, the interpretation of results was conceived collaboratively by me and the co-authors. For the second study, I received feedback from both co-authors who were involved in study 1. In this study I presented ideas that I developed on my own which diverged from the original DFG proposal. After receiving feedback from anonymous peer reviewers, the concept of individual associations in semantic memory was replaced with individual attitudes since we measured propositional statements. Moreover, I introduced the idea that defense motivation and accuracy motivation would underly confirmation bias, and I interpreted my results also in terms of navigation between and within patches according to information foraging theory. In the third study, the ideas were an extension of those I developed in the second publication. Here the initial co-authors of study 1 again both provided feedback. The third study was an extension to the second study using the same ideas and a representative sample instead of a student sample and including the experimental factor of tag size, after feedback of my initial co-authors.

Introduction

Confirmation bias is the fundamental tendency to select, interpret and evaluate information in a way that confirms the pre-existing beliefs and attitudes of an individual (Hart et al., 2009; Klayman, 1995; Lord, Ross, & Lepper, 1979). This biased tendency is also prevalent in online environments (Liao, Fu, & Mamidi, 2015; Park, Konana, & Gu, 2013; Schwind & Buder, 2012; Schwind, Buder, Cress, & Hesse, 2012; Vydiswaran, Zhai, Roth, & Pirolli, 2012), and this biased, one-sided preference of information may lead to misconceptions and polarization of the individual's knowledge and attitudes (Lord et al., 1979), and consequently even polarization of social groups in online environments (Del Vicario, Scala, Caldarelli, Stanley, & Quattrociocchi, 2017; Muchnik, Aral, & Taylor, 2013).

This raises the question whether confirmation bias can be reduced, to mitigate negative effects on individual knowledge and attitude formation. To address this question in the context of online environments, as well as for specific content domains, first I lay out why individuals tend to confirm their pre-existing beliefs and their underlying motivations. According to a meta-analysis two essential motivations for confirmation bias are *accuracy motivation* on the one hand, and *defense motivation* on the other (Chaiken, Liberman, & Eagly, 1989; Hart et al., 2009; Kunda, 1990).

Individuals high in accuracy motivation strive to form accurate appraisals of stimuli, whereas individuals high in defense motivation strive to defend their existing beliefs and attitudes in the light of new information (Hart et al., 2009). Accuracy motivation promotes tendencies to form objective, open-minded appraisals. Individuals high in accuracy motivation aim for cues that highlight objective correctness and validity of information, such as information utility or outcome relevance (Hart et al., 2009).

In contrast to accuracy motivation, defense motivation promotes tendencies to avoid self-threat by avoiding stimuli that increase cognitive dissonance and conflict (Hart et al., 2009). To avoid cognitive conflict, individuals tend to choose information that is consistent with their existing attitudes, and they avoid attitude inconsistent information. Individuals avoid attitude inconsistent information even more, when cues emphasize the validity or correctness of attitude inconsistent information (Hart et al., 2009).

Accuracy motivation should increase in situations when information utility and outcome relevance are important (Hart et al., 2009). For example, in a health-related information search context, where searchers aim to learn about different treatment options for a mental

disorder, searchers will be particularly interested in the efficacy of each treatment option (outcome relevance), and whether information is addressing the efficacy of a treatment (information utility). So, in health-related information search, where the searcher choses among different treatment options, the searcher will be open toward treatments that may be attitude inconsistent but promise the best treatment outcome, given accuracy motivation is high.

Ideally, to determine the most relevant information leading to the best possible treatment outcomes, searchers browsing online *content aggregation platforms* should consider cues about the *online communities* who collect and share content. In these online contexts, an informative cue highlighting validity of information shared by the community can be majority consensus. It has been shown that when majorities within groups convey a message, recipients will often find it more believable (Crano & Prislin, 2006). There are many online cues available to indicate majority consent, such as the number of likes, retweets or the number of citations in scholar search engines. For the present work, I will investigate majority consensus cues as implied by tag size in social tag clouds, and expertise implied by banners on top of the pages of the tagging environment.

Social tag clouds arise as online communities provide verbal labels as metadata for content on the web (Mika, 2007; Trant, 2009). Tags can be provided for digital artefacts such as images (facebook.com), books (librarything.com), movies (movielens.org) or blog posts (wordpress.com). Tagging platforms aggregate tags provided by the community and display the most popular tags. More popular tags are larger than less popular tags (Cress, Held, & Kimmerle, 2013; Kuo, Hentrich, Good, & Wilkinson, 2007). This way, a form of implied or implicit majority consensus of the community arises. With the help of this majority consensus cues, information searchers have a visual cue of information validity.

To understand how searchers make use of consensus information of social tags, and thus of the implied information validity, the *extended information scent model* has been proposed (Cress & Held, 2013; Cress et al., 2013; Held, Kimmerle, & Cress, 2012). According to this model, individuals estimate which tag will lead to the desired information by estimating the expected gain of information with respect to the search goal. This estimate is referred to as information scent and depends on individual semantic associations in memory (Cress et al., 2013). When individual associations match semantically with the tag, the searcher assumes that the tag will lead to the desired information. The tag that elicits the strongest activation in memory will be clicked. Besides individual associations, information scent also depends on

collective associations inherent in the tag cloud. Collective associations are stronger when tags are more popular and displayed in a larger font size. The size of the tags linearly increases its selection rate, and therefore information searchers follow implied majority consent of a community (Cress & Held, 2013; Cress et al., 2013; Held et al., 2012).

In the context of health-related search with social tagging platforms, the question arises if besides individual semantic associations, also *individual* prior attitudes can influence tag-based navigation and confirmation bias. So, the present dissertation aims to fill the research gap of whether the extended information scent model could also consider prior attitudes to explain the phenomenon confirmation bias.

When investigating confirmation bias in a domain that increases accuracy motivation, individual evaluations about information should influence navigation and evaluation of content. In the present dissertation, besides *individual influences*, I also investigate *collective influences:* On the individual side, *prior knowledge, prior attitudes* and as moderating factor *confidence in prior attitudes*. As collective influences, I investigate *popularity of social tags* and *source credibility* of the tagging community. Individual evaluations should influence whether searchers perceive content as valid, and confidence in prior attitudes should determine whether searchers perceive their own beliefs and attitudes as valid, whereas collective cues should influence whether searchers perceive content provided by others as valid or not.

The role of *prior attitudes* as well as *prior knowledge* on confirmation bias, given searchers are accuracy motivated, is straightforward. If searchers evaluate a certain treatment (e.g. psychotherapy) more positively than another one (e.g. antidepressants) and have more favourable attitudes or favourable knowledge about the treatment, they should select the treatment more often and evaluate the treatment positively after reading about the treatment. It should be noted that according to *defense motivation*, when searchers with increased attitude confidence face attitude inconsistent information, they feel less threat and even show a tendency to favour attitude inconsistent content (Hart et al., 2009).

Next, I address the moderating role of *confidence in prior attitudes* in a context that should elicit accuracy motivation. The most recent review of confirmation bias has found that increased or decreased confidence (e.g. after feedback on the correctness of a choice) has consequences on confirmation bias that would be expected under defense motivation (Hart et al., 2009). That is, when confidence is increased, searchers feel reduced self-threat and therefore are more open to attitude-inconsistent information. In contrast to this, the present

dissertation I manipulate metacognitive confidence, which is confidence about the validity of one's own thoughts or attitudes (Petty, Briñol, & DeMarree, 2007). So, when metacognitive confidence is increased, this should act as an internal cue for the validity or correctness of one's attitudes. Consequently, when attitudes are perceived to be valid (high confidence), attitudes should guide information selection under high accuracy motivation, in contrast to when attitudes are not perceived to be valid, and metacognitive confidence is low (Briñol, Petty, & Barden, 2007; Petty, Briñol, & Tormala, 2002). To my knowledge, so far, no studies investigated the influence of metacognitive confidence on confirmation bias, thus the present work closes a gap in the literature here as well.

Besides individual evaluations and confidence as a moderator, I suggest that *collective cues*, such as t*ag popularity* act as external validity cue of information aggregated in social tagging platforms. More or less popular tags should linearly increase or decrease confirmation bias according to tag size (Cress et al., 2013). This is because tag popularity conveys implied majority consensus of the community, which serves as additional cue about information validity. It has been shown that when majorities convey a message, recipients will tend to believe it (Crano & Prislin, 2006). Moreover, tag popularity manifests in tag size, which acts on a basic visual perceptual level (Lohmann, Ziegler, & Tetzlaff, 2009) in guiding searchers. I do not expect that tag popularity interacts with attitude consistency or inconsistency. That is, tag popularity linearly increases and decreases the selection and evaluation of attitude consistent as well as attitude inconsistent tags and tag-related content (Cress et al., 2013).

Finally, *source credibility* in terms of expertise of the tagging community should affect confirmation bias in social tagging environments. Particularly for content aggregation platforms, where there are no gatekeepers who filter high quality information, the ability of the individual to distinguish high from low credibility is important (Gerjets & Kammerer, 2010). Research has shown that online information searchers do not always successfully recognize the degree of source credibility in online search, especially in the health context (Kammerer, Amann, & Gerjets, 2015; Kammerer, Bråten, Gerjets, & Strømsø, 2012). Moreover, the *platform type* plays a role in credibility evaluations. When searching on websites, for example, perceived expertise of content authors drives credibility evaluations, but when searching information on user-generated content platforms such as forums or blogs, homophily–high demographic similarity between searcher and content creator–increases credibility evaluations (Ma & Atkin, 2017). Social tagging platforms are a blend of these two types: aggregated information can represent user-generated content as well as webpages.

Finally, perceived credibility depends on the sample, particularly in health-related domains (Yang & Beatty, 2016). In student samples, there is a higher association between manipulated and perceived credibility, whereas in non-student samples, this relationship is weaker. For social tagging platforms, the searchers' perceived source credibility of the tagging community, and the link with confirmation bias have not been investigated yet. To close this gap, in the current dissertation I investigate whether source credibility of the community that collects content on social tagging platforms is adequately recognized by searchers, and if searchers' credibility evaluations have consequences on confirmation bias. I expect that high source credibility compared to low source credibility increases the selection and positive evaluation of content independent of prior attitudes.

Objectives of the Dissertation

The current dissertation has two main objectives which were explored in three experiments: a) in extension to existing literature, whether not only individual and collective associations, but also confirmation bias and related *individual evaluations,* as well as *the collective cues* would influence tag-based navigation in health-related search with social tags, and b) the assumption that accuracy motivation guides health-related information search.

As domain I chose a mental health-related search context, in which participants search for information about treatment options of depression. The two main treatment options were psychotherapy and antidepressants. There is evidence that different populations are biased to have a more positive opinion towards psychotherapy than towards antidepressants (Angermeyer, van der Auwera, Carta, & Schomerus, 2017; Jorm et al., 2005; Van Der Auwera, Schomerus, Baumeister, Matschinger, & Angermeyer, 2017). I expected that the preference for psychotherapy would lead to confirmation bias, so participants would select and evaluate psychotherapy more favourably than antidepressants. As mitigating factors of this confirmation bias, I investigated *individual evaluations* of the estimated validity of information: *1) prior knowledge, 2) prior attitudes, and* as a moderator *3) confidence in prior attitudes. As collective cues* for information validity I investigated the influence of *4) tag popularity, and 5) source credibility of the tagging community*. The procedure for all three experiments was similar. Searchers were asked to search for information about treatment options for depression, and then they navigated with a social tag cloud that included antidepressant and psychotherapy treatment tags. Clicking on tags guided the searchers to blog posts that either highlighted the efficacy of antidepressants or psychotherapy in the treatment of depression. Besides navigation, I measured attitude change in terms of treatment efficacy ratings presented prior to and after navigation.

The main research questions in study 1 are as follows:

Individual Evaluation: Does prior knowledge provided by experts influence confirmation bias?

Collective cue: Does popularity of social tags influence confirmation bias?

Study 1 addresses the question whether social tag clouds are a suitable interface to reduce confirmation bias. Given that individuals show a confirmation bias towards psychotherapy, the bias should be lower when attitude inconsistent (antidepressant > psychotherapy) tags are more popular, compared to balanced popularity of tags (psychotherapy = antidepressants).

Moreover, the bias should be higher when attitude consistent tags are more popular (psychotherapy > antidepressants) or popularity is balanced (psychotherapy = antidepressants). Besides tag popularity, study 1 addresses whether manipulated prior knowledge–provided by experts–influences confirmation bias. When attitude inconsistent prior knowledge is provided before search, participants should show decreased confirmation bias compared to when attitude consistent knowledge is provided.

The main research questions in study 2 are as follows:

Individual Evaluation: Does confidence change the influence of prior attitudes on confirmation bias?

Collective cue: Does source credibility lead to more openness independent of prior attitudes?

Study 2 addresses the question as to whether confidence moderates the influences of pre-existing prior attitudes on confirmation bias. Prior attitudes were inquired in the form of arguments for and against psychotherapy, before navigation. Attitude confidence was manipulated by having participants recall situations in which they were either confident or not about their own thoughts (Petty et al., 2002). Source credibility was manipulated by banners on top of the search environment and show that either experienced professionals or students collected and tagged the blog posts that are related to each social tag. For the study presented here, the sample consisted largely of university students as in study 1.

The main research questions in study 3 are as follows:

Collective cues: Does popularity of social tags influence confirmation bias?

Does source credibility lead to more openness independent of prior attitudes?

Individual Evaluation: Does confidence change the influence of prior attitudes on confirmation bias?

Study 3 addresses the same questions as the previous studies, however, in comparison to study 2, the experimental factor tag popularity was included as well, to investigate possible interaction effects with prior attitudes, confidence in prior attitudes and source credibility. Moreover, in this study I tried to replicate the previous findings with a *sample representative of the German population*.

Summary of the Main Findings

With the present dissertation I intended to investigate the phenomenon of confirmation bias with content aggregation platforms. The underlying assumption was that accuracy motivation is present in online health-related search. Given this, the aim was to test *individual evaluations and collective cues* influence confirmation bias. Accuracy motivation predicts that individuals seek to form objectively correct and valid appraisals (Hart et al., 2009). In contrast to this, defense motivation causes individuals to avoid dissonant cognitions and validate their self-concept by preferring attitude consistent information (Hart et al., 2009). In a health-related domain, I expected searchers to have high accuracy motivation, so they would search for the treatment options with the best possible outcome. I expected that with high accuracy motivation, the validity of information should have high importance when people navigate and evaluate content (Hart et al., 2009). As *individual evaluations* of information I investigated the influence of *1) prior knowledge, 2) prior attitudes,* and the moderator *confidence in prior attitudes.* As *collective cues* of information validity, I investigated the influence of *4) tag popularity, and 5) source credibility of the tagging community*.

Individual Evaluations

When it comes to navigating with tags and providing tags on social tagging platforms, research has extensively focussed on semantic associations in memory (Cress, 2013; Cress et al., 2013; W. Fu, Kannampallil, & Kang, 2009; Seitlinger, Ley, & Albert, 2015). The underlying idea is that semantic similarity between social tags and activated concepts in human memory determines which tags searchers click, and which tags users provide for content. If a searcher has a certain search goal (e.g. treatment of depression), this activates related concepts in the semantic memory of the searcher to a certain degree, depending on the strength of the connections between concepts, in other words, the weights of the associations. The searcher is more likely to click on tags that have high weights and thus are highly similar with the activated concepts in memory (Cress & Held, 2013; Cress et al., 2013; Held et al., 2012). This similarity matching principle has proven to be fruitful for research in selecting as well as providing tags for online resources (Cress et al., 2013; W. Fu et al., 2009; Seitlinger et al., 2015). In my dissertation I extend the focus from the weight of associations to evaluations of activated concepts, as this may be relevant for many search contexts such as the health-related context, where individuals may have positive and negative experiences or opinions towards certain topics or treatments which determine the searcher's individual evaluations about topics or treatments. For example, certain types of treatments for disorders might

consist of positively or negatively evaluated concepts in the memory of the searcher (e.g. negative: antidepressants - addictive, positive: psychotherapy – non-intrusive).

For the first study, I chose to manipulate individual prior knowledge about the degree of treatment efficacy, affecting the evaluations of the concepts in memory. Expert knowledge was provided in the form of blog posts that positively highlight the efficacy of either psychotherapy or antidepressants. This way, evaluative of aspects of each treatment were manipulated. Since associative weights were determined by the frequency by which participants encountered the concepts in the blog posts, and in total, two blog posts were provided, associative weights were kept constant across the experimental conditions. The first research question, the current dissertation addresses, is whether the manipulated individual evaluation about treatments in prior knowledge influences the selection of tags and tag-related content (i.e. blog posts), and content evaluation.

Prior Knowledge

So, in study 1, as *individual evaluation*, I manipulated *prior knowledge provided by experts*. Results showed that prior knowledge did not affect selection of tags, but the selection of tag-related blog posts, and content evaluation. This finding partially supports the notion that accuracy motivation was present in the health-related search context, since expert knowledge indicates highly valid information, and when expert knowledge was inconsistent with prior attitudes, participants reduced their bias towards psychotherapy and selected more attitude inconsistent content. If participants would have been defense motivated, after facing attitude challenging expert knowledge, participants would not have been willing to select more attitude challenging blog posts, but they would have tried to validate their attitudes by selecting more attitude consistent blog posts. Besides selection of attitude inconsistent blog posts, the same pattern of results was found for the evaluation of attitude inconsistent information. So, in study 1 there was partial support for the expectation that individual evaluations influence confirmation bias (selection and evaluation of tag-related content, but not selection of tags), and that accuracy motivation was driving information search and evaluation.

Prior Attitudes and Confidence in Prior Attitudes

The prior knowledge provided by experts that was manipulated in study 1, showed that individual evaluations and thus the evaluations of treatment concepts in memory can be changed in the short term. In contrast to this, naturally formed, long-term evaluations of individuals may be a much stronger indicator for the evaluation of concepts in memory, and

thus form a more concise individual evaluation that affects confirmation bias. So, in the second study, I inquired self-reported prior attitudes in an open format, to have an individualized predictor of the evaluation of concepts in memory.

Study 2 also introduces the moderator variable confidence in prior attitudes, which should moderate the influence of prior attitudes. A meta-analysis showed that high individual confidence decreases confirmation bias, as high confidence increases the validity of one's self-concept (Hart et al., 2009), but this finding is explained with defense motivation. Contrasting this, in the present study, in the health-related context, I expected searchers to be accuracy motivated, that is, participants were expected to aim for objectively valid information about different treatments of depression. When people are accuracy motivated and perceive their attitudes to be valid, because of their heightened confidence, this should increase favourable evaluations and selection of attitude consistent content and thus increase confirmation bias.

Using a student sample, I found that the influence of prior attitudes depended on confidence as predicted by accuracy motivation. When confidence in prior attitudes was high, implying that existing attitudes were perceived to be highly valid, participants selected more attitude consistent blog posts, and evaluated attitude consistent blog posts more favourably. However, the selection of tags was not associated with prior attitudes and manipulated confidence in prior attitudes (nor was prior knowledge associated with tag selection in study 1). Instead, I surprisingly found that the influence of attitudes on tag selection depended on source credibility. When source credibility was high, there was no influence of attitudes on tag selection, and when source credibility was low, prior attitudes were positively associated with the selection of attitude consistent treatment tags.

This finding is ambiguous with respect to the accuracy or defense motivation framework. On the one hand, high quality, attitude-inconsistent information poses a threat to the self-concept, so individuals show more confirmation bias under high source credibility. This was not the case. Under accuracy motivation, on the other hand, individuals should preferably select information with high source credibility independent of their prior attitudes (Hart et al., 2009), which was the case. But low source credibility was associated with confirmation bias, and there is no gain to be expected for information low in credibility; neither under accuracy nor defense motivation. One explanation of the increased confirmation bias with respect to the selection of treatment tags could be that under low source credibility, individuals tend to rely

on heuristics such as the confirmation bias, since there is a lack of cues for informational value in the environment.

In sum, the findings of study 2 again partially support the expectation that individual evaluations affect confirmation bias in social tagging platforms, and that accuracy motivation guided searchers.

Replication Attempt with a Representative Sample

For study 3, I intended to replicate the previous findings in a sample representative of the German population. Besides the general question of whether confirmation bias in the health context holds true in a representative sample, the study addressed whether student samples tend to consider credibility different than non-student samples (Yang & Beatty, 2016). Moreover, it could be possible that the interplay of prior attitudes and confidence could interact with source credibility and may lead to different results for confirmation bias in a representative sample.

The experimental manipulation of confidence in prior attitudes, however, failed, so I used a correlational design with the manipulation check scores instead of the experimental group as predictor variable. This showed that only low confidence in prior attitudes was positively associated with the selection of attitude consistent blog posts (but as in study 2, not tags). Importantly, this result stands in contrast to the expectation that accuracy motivation guides information searchers, and supports the notion that defense motivation drives information searchers. The defense motivation account would predict that when confidence is low, searchers aim to protect their self-concept, and thus they select more attitude consistent information. In line with the defense motivation account, in study 3, I found that searchers selected more attitude consistent blog posts under low confidence. If individuals had been accuracy motivated, low attitude confidence should have reduced the estimated validity of attitude consistent information, and searchers should have selected less, not more attitude consistent content. So, with respect to the selection of content, and in contrast to both previous studies, study 3 provides partial support for confirmation bias under the defense motivation. Again, only the selection of blog posts, but not tags was affected by the interaction of confidence in prior attitudes.

Regarding the evaluation of information content, results showed that the influence of individual and collective validity cues interacted in a way providing further support for the notion that defense motivation guided the representative sample of searchers in study 3. That is, only when source credibility was high, low confidence in prior attitudes led to more

favourable ratings of attitude consistent information. Likewise, a review has shown that high (vs. low) source credibility of attitude inconsistent information may pose a threat to the searcher, and consequently under defense motivation, searchers favour less attitude inconsistent information (Hart et al., 2009). In sum, the threat of highly credible, inconsistent information was present when at the same time confidence was low, and this provides further support for the notion that searchers were defense motivated. I attribute this finding to the representative sample that was used.

Taken together, with respect to individual evaluations, all three experiments showed that selection of tag-related blog posts and the evaluation of information depended on prior knowledge, prior attitudes and confidence in prior attitudes. Tag selection, however, was only related to the individual evaluation of prior attitudes in study 2, where prior attitudes interacted with source credibility. Studies 1 and 2 also supported the expectation that accuracy motivation did drive navigation and evaluation of information. Both studies were based largely on students. In contrast to this, study 3, which investigated a representative sample, showed a pattern of results that supports the notion of defense motivation. This interesting divergence is in line with research on the evaluation of credibility in online environments (Yang & Beatty, 2016), which shows that manipulated credibility is stronger associated in student than in non-student samples. This will be discussed in further detail below.

Collective Cues

Besides individual evaluations that should affect confirmation bias, I investigated the question whether collective cues play an important role in social content sharing platforms. Popularity of content and thus implied collective consensus is arguably one of the most important cues, searchers encounter on social platforms. When a majority endorses information online, for the searcher it becomes crucial to determine the authority of the majority, and the likelihood that the majority is knowledgeable about the content it shares. So, for this dissertation, I investigate tag popularity as well as source credibility as collective validity cues.

I expected both types of collective validity cues to influence the searcher independent of her or his prior attitudes. I expected that social tag clouds are a suitable tool to reduce the influence of prior attitudes, since tag clouds show implicit majority consent on a topic, and tag clouds guide tag selection on a basic visual perceptual level (Lohmann et al., 2009). For source credibility, I did not expect that the influence of prior attitudes on information search and evaluation would depend on source credibility, that is, I did not expect an interaction of prior attitudes and source credibility or tag popularity. So, *social tag clouds and source*

credibility were supposed to act as *collective cues*, independently and guiding searchers to find content that leads to high quality and valid information.

Tag Popularity

Collective consent or tag popularity in social tag clouds is shown via the size of each tag (for the tag clouds used in the present work, see Figure 1.1 in Appendix Publication 1). Participants were told, that the more users provide a tag for a certain treatment, the larger the treatment tag appeared in the tag cloud. To operationalize attitude-consistency, in study 1, I provided three versions of tag clouds: a) psychotherapy tags larger than antidepressant tags, b) both treatment tags with the same size, c) antidepressant tags larger than psychotherapy tags. Results showed that in study 1, tag popularity decreased the influence prior attitudes in a linear way. When psychotherapy tags were relatively popular, participants clicked on these tags most often in version a, less often in version b, and the least often in version c, the same was true for antidepressants. Tag popularity affected the selection and related evaluation of content in the same way. So, although searchers were overall biased towards psychotherapy according to their prior attitudes, tag clouds linearly reduced the influence of this bias on navigation as well as resulting evaluation of content.

In study 2 I did not manipulate tag popularity, since there was a linear effect of tag popularity on all dependent variables and tag popularity did not interact with prior knowledge. Also, because I wanted to focus on the condition of tags that were inconsistent with the prior attitudes (antidepressant tags > psychotherapy tags; version c). For study 3 however, I manipulated tag popularity again, and provided tag clouds where either psychotherapy or antidepressants were relatively popular (version a and version c) and dropped the balanced tag popularity condition in version b from study 1. The aim of the study was to replicate the findings of studies 1 and 2 with a representative sample, since the first two studies largely consisted of students.

The results of study 3 again showed a main effect of tag popularity on navigation the selection of tags and tag-related blog posts. Although, there was no direct effect of tag selection on content evaluation as in study 1, a mediation analysis showed, that the effect of tag popularity on content evaluation was mediated by tag selection. The effect sizes of tag popularity on navigation were large across studies 1 and 3, revealing the strongest effect in this dissertation.

In sum, in studies 1 and 3, I manipulated tag popularity, and both studies show that this manipulation reduced confirmation bias. Importantly, tag popularity did not interact with the individual factors prior knowledge and the interplay of prior attitudes and attitude confidence.

This shows that tag clouds may be highly effective tools to guide information selection or even nudge individuals toward certain options in information search, reducing the influence of their prior attitudes.

Source Credibility

In study 2, I manipulated *source credibility* in terms of expertise of the tagging community. In the high source credibility condition, I alleged that tagging community consists of experts who have been active in their field of research for years, whereas in the low credibility condition, I alleged that the tagging community consists of first semester students. I expected that high (vs. low) credibility would increase (vs. decrease) openness towards content independent of prior attitudes.

However, in study 2, I found that when credibility was low, participants selected tags in line with their prior attitudes, but when credibility was high, participants selected tags independent of their prior attitudes. These results partially support the assumption that searchers were guided by accuracy motivation. When source credibility of the community was high, participants did select content independent of their prior attitudes. If participants were defense motivated, high validity of attitude inconsistent information would have presented a threat of their self-concept and participants would have preferred attitude consistent information (Hart et al., 2009).

However, it was surprising that the influence of credibility could not be found for blog post selection or content evaluation (the influence of source credibility on blog post selection and evaluation was mediated via tags though, see Figure 2.8 in Appendix Publication 2). The finding that only in the low source credibility condition, participants selected more attitude consistent tags was interesting. One possible explanation could be that when credibility in tagging environments is low, and thus validity cues are not available in a context where validity cues are desired (accuracy motivation), in order to ease the decision-making process, participants could have employed simple heuristics such as the confirmation bias (Chaiken et al., 1989).

In study 3, using a representative sample, I conducted a correlational analysis, since the manipulation checks showed no significant difference between the high and low credibility groups. Here, I found a pattern of results showing that defense motivation may have guided information search. As mentioned above, with respect to the evaluation of information, individual and collective validity cues interacted. A positive association between attitudes and the evaluation of attitude consistent information was found only when confidence was low,

and source credibility was high. In contrast to the findings in study 1 and study 2, this finding supports the account of defense motivation in health-related search. When confidence is high, according to defense motivation, people would not be threatened by attitude inconsistent information, and they should be more willing to select attitude inconsistent information (Hart et al., 2009). Regarding the evaluation of content, the fact that this result was found only when source credibility was perceived relatively high, this provides further support for defense motivation. High credibility of attitude inconsistent information poses a threat under defense motivation. To avoid threat, searchers avoid attitude inconsistent information (Hart et al., 2009).

In sum, I manipulated source credibility in studies 2 and 3, and found mixed results. In study 2, I used a sample consisting mainly of students and the credibility manipulation was successful. In study 3, I used a representative sample, but the manipulation of credibility was not successful, so a correlational analysis was done. In study 2 credibility influenced navigation (selection of tags), but not evaluation of content. In study 3, credibility did affect evaluation but not selection of content. In both studies, credibility interacted with prior attitudes, in a way that supports the accuracy motivation account in study 2 and the account of defense motivation in study 3.

Accuracy and Defense Motivation in Health-Related Search

Study 1 supported the assumption that participants were accuracy motivated, and with studies 2 and 3, I found mixed support for the assumption that participants were accuracy motivated in health-related information search. Individual or subjective indicators such as prior knowledge (study 1), and prior attitudes moderated by confidence (study 2), guided navigation and evaluation of content. As collective indicators, study 1 and 3 show that tag popularity influenced navigation and subsequent content evaluation. The second collective indicator, source credibility of the tagging community affected tag selection in study 2. This is in line with predictions of accuracy motivation, but the results of source credibility in study 3 were more in line the defense motivation account for information search. This may be due to the differences in the sample as I will discuss below.

Strengths and Limitations

In the current dissertation, I found that affective processes guide selection and evaluation of tag-related content. This extends existing research in the field of social tagging, which exclusively focused on semantic associations in memory when explaining navigation (Cress, 2013; Cress & Held, 2013; Cress et al., 2013; Held et al., 2012). Moreover, in in all three studies, I found that confirmation bias was involved in the selection of tags, related content and content evaluation of searchers. This shows that besides semantic associations, also heuristic cognitive processes such as confirmation bias play a role tag-based navigation, which is a completely new finding in the line of research of social tagging.

Prior Attitudes

To operationalize the affective processes that play a role in tagging environments, I used prior attitudes in studies 2 and 3. I asked participants to state pro and contra arguments regarding the treatment efficacy of psychotherapy and antidepressants. In doing so, a broad and natural indicator of the affective concepts in memory was obtained. This also allowed to describe the precise opinions of individuals about treatments, which is particularly relevant for health-related research fields. For the representative sample in study 3, I also aggregated the arguments for and against each treatment, and presented the most common topics that were provided by the representative German sample. The drawback is that attitudes were not manipulated so attitudes may have been confounded with for example attitude certainty or knowledgeability of participants.

Prior attitudes had a consistent effect on information search, however, in study 2 there was an unexpected interaction with credibility with respect to tag selection, and in study 3 there was no effect of prior attitudes on the selection of tags, whereas both studies showed an effect on blog post selection. This may be explained in part because the tags represented topics (e.g. "psychoanalysis", "serotonin inhibitors"), and the tags were not evaluative themselves (e.g. a hypothetical tag "ineffective"). The impact of prior attitudes may have been higher, if also tags were evaluative.

Confidence in Prior Attitudes

By investigating the effects of metacognitive confidence on navigation and evaluation of online content, the current dissertation also entered uncharted territory. Whereas a great body of research investigated increased or decreased confidence as a result of feedback of the experimenters about correct or wrong decisions by participants (Hart et al., 2009), little work

has investigated the metacognitive aspect of confidence. That is, confidence manipulations that do not elicit self-threat as a result of negative feedback. Metacognitive confidence is the degree to which individuals estimate their knowledge or attitudes to be valid or correct. To manipulate metacognitive confidence, a manipulation that draws from the so-called Feynman technique (Holyoke & Feynman, 2010), is to have participants reflect on their ability to explain in a step-by-step manner, the causal functioning of objects (Johnson, Murphy, & Messer, 2016). In doing so, gaps of knowledge become apparent for the individual and confidence decreases.

I drew on a metacognitive confidence manipulation of a different line of research. Research about the reception of persuasive messages showed that when participants recalled situations in which they were confident or doubtful about their thoughts, high confidence increased attitude consistent processing of messages (Petty et al., 2002). The same authors used affective manipulations in other studies to manipulate confidence like writing about happy or sad personal experiences, or play a role based on happy or sad dramatic scripts (Briñol et al., 2007), but to my knowledge the authors or other researchers did not directly replicate their confidence manipulation. Therefore, studies 2 and 3 of the current dissertation provide an important replication attempt. And the replication attempt was only partially successful. In study 2 with a sample consisting largely of students, the replication was successful, however, with a representative sample of the German population in study 3, the replication did not work. Additionally, in study 2, the effect size was small to moderate, whereas the original research reported a large effect size of the manipulation (Petty et al., 2002).

Tag Popularity

The manipulation of social tag popularity has yielded the largest and a robust effect in the current dissertation. Manipulated tag size had consistent, large effect in studies 1 and 3, showing the capacity of tag clouds to counter heuristically guided search.

However, a drawback is that the tags themselves have been constant across all experiments. And the tags themselves have some limitations. A limitation with respect to the presentation of the tag clouds is the selection and arrangement of tags in the cloud. The treatment terms used as tags were selected to have the same number of characters across both treatments, but the antidepressant tags were slightly less representative of natural language ("serotonin inhibitors"), compared to their psychotherapy counterparts ("behavioral therapy"). Moreover, the arrangement of the tags in the cloud was not randomized due to technical restrictions,

which would have been desirable to counter positioning effects known for tag clouds (Lohmann et al., 2009).

Finally, antidepressant and psychotherapy tags may have had different base rates of selection, since psychotherapy tags might have been more representative of natural, everyday language (*psychotherapy*: cognitive therapy, gestalt therapy, interpersonal therapy, person-centered therapy, psychoanalysis; vs. *antidepressants*: MAO inhibitors, norepinephrine inhibitors, serotonin inhibitors, tetracyclic antidepressants, tricyclic antidepressants).

With all these limitations, however, a highly promising finding is that tag popularity had a strong impact on navigation and subsequent information evaluation. This shows the huge potential of social tag clouds to guide information searchers and to reduce their confirmation bias. Particularly in health-related search, where prior attitudes may be biased or misleading, social tags might provide a corrective, given that the majority of users in these systems consists of experts.

Source Credibility

Social tagging platforms or other content sharing platforms lack professional gatekeepers who critically screen and select high quality content. And the health-related domain is one in which the skill of discriminating between trustworthy and less reliable online resources is essential for the searcher, to obtain high quality information, and to make informed decisions. In study 2, when banners on top of the page indicated high (vs. low) expertise of the tagging community, the sample consisting largely of students recognized experts correctly as more knowledgeable then first semester students. The representative sample in study 3 however did not rate experts as more knowledgeable than first semester students. This is in line with a recent review on the perceived credibility in online environments (Yang & Beatty, 2016), and a highly important finding for the evaluation and design of online platforms, and this implies that the target audience needs to be considered, which will be discussed in the next section.

Accuracy and Defense Motivation

Taken together, studies 1 and 2 with samples consisting largely of university students do support the accuracy motivation account in the health-related search with social tagging platforms. However, the findings of study 3 using a representative sample supported the defense motivation account. All significant results seemed to support this interpretation, with one exception. In study 2, when credibility was low, participants selected more attitude consistent tags. When credibility was high, participants did not select more attitude consistent

tags, which is in line with the accuracy motivation account, since participants should aim for high quality of information when accuracy motivated. So why were participants guided by their attitudes when credibility was low? An explanation could be that accuracy and defense motivation are independent dimensions, and when credibility was low, accuracy motivation might have been suppressed, since participants could not gain valuable information in the tagging environment. So, they merely relied on the confirmation bias heuristic, and their prior attitudes. The question to which degree participants were accuracy or defense motivated could have been answered if both types of motivation had been measured directly, which is another limitation of the current dissertation.

Generalizability and External Validity

The main three factors limiting the external validity of the present study are the *samples* used, the *domain* and the *navigation platform type*. First, the external validity or generalizability of the present results is limited. In all three studies, I used the mental health-related search context, to be specific, the treatments of depression. Therefore, it would be desirable to investigate information search with social tagging platforms also in other health-related domains, and more generally, in other domains or contexts in which accuracy motivation is expected to occur. Potential contexts could be workplace learning or formal education.

Besides the domain, the results of the present dissertation were found with a specific content aggregation platform, which is mainly characterized by the usage of social tag clouds as navigation interface. It would be desirable to test the present results using different types of online platforms and content sharing or content aggregation platforms with different navigation interfaces, to test generalizability of the results to other online environments.

Besides the domain and the type of platform, the third restriction applies to the samples with which the results were found. In studies 1 and 2 the sample consisted largely of university students, whereas in study 3 the sample was representative of the German online population. Interestingly, with the representative sample, I could not replicate the pattern of results expected if participants were accuracy motivated.

Implications for Future Research and Practice

Although it is an important extension for the semantic focus on explaining navigation in tagging systems, the current dissertation remains exploratory when it comes to the formal representation of prior attitudes and their effects on navigation, whereas work on semantic associations in tagging systems is often highly formalized (Cress et al., 2013; W. Fu et al., 2009; Ley & Seitlinger, 2015; Trattner, Kowald, Seitlinger, Kopeinik, & Ley, 2016). Likewise, attitude research in general adopted ideas from connectionist models of human cognition quite late. According to Van Overwalle & Siebler (2005), the first formal model that draws on connectionist networks to explain attitude formation and change has been introduced by the authors in 2005. This could serve as a reference model to merge research on the field of semantic associations with the research on attitude formation in social tagging systems.

Nonetheless, the present dissertation breaks new ground compared to existing, formalized theories of social tagging systems, by adding the perspective of confirmation bias. That is the influence of individual prior attitudes on tag-based navigation and evaluation of tag-related information (Cress, 2013; Cress & Held, 2013; Cress et al., 2013; Held et al., 2012). This is particularly noteworthy since tags often reflect evaluations (Heckner, Neubauer, & Wolff, 2008) of users, and the motivation to provide tags can be social in nature. For example, it has been found that opinion expression, self-presentation, or more general, social presence are motifs when users provide tags (Ames & Naaman, 2007; Nov, Naaman, & Ye, 2008; Zollers, 2007).

Taken together, this dissertation potentially lays the ground for a broader view of tagging behavior, going beyond semantic associations towards the influence of attitudes and heuristics in social tagging. This may not only be true for tag-based navigation (Cress et al., 2013), but potentially also for the imitation of tags (W.-T. Fu, Kannampallil, Kang, & He, 2010; Seitlinger, 2012; Seitlinger et al., 2015).

Particularly for the health-related domain, a recent review has shown that perceived validity of information in online contexts varies with the platform that is being used (Yang & Beatty, 2016). On platforms presenting user generated content, participants' credibility evaluations are driven by homophily, that is the demographic similarity between content author or provider on the one hand, and the information searcher on the other. In contrast to this, on general purpose websites, credibility evaluations are driven by perceived expertise. In the

present work I found that credibility evaluations in studies 2 and 3 diverged. Whereas in study 2, for the student sample, credibility evaluations were driven by expertise, in study 3 with the representative sample for German online users, credibility evaluations could not be increased with higher alleged expertise.

It is conceivable that students may view themselves as more similar to experts who have a university degree, so that homophily might play a role here as well. This finding is in line with another recent review that found that manipulated and perceived credibility is stronger linked in student than in non-student samples (Yang & Beatty, 2016). In sum, these findings show that in the health-context, and possibly other context where accuracy motivation plays a role, confirmation bias depends on the sample and the platform being used.

In study 3, I discussed the distinction of tag selection and the selection and evaluation of tag-related content within the context of information foraging theory (Pirolli & Card, 1999), which distinguishes between-patch navigation (tag selection) and within patch navigation (blog post selection). In all studies of the present dissertation, I found that tag and blog post selection diverged to some degree. This divergence may be because tags reflected topics, whereas blog posts were evaluative in nature, positively highlighting the efficacy of a treatment. So, the exploratory (tag selection) and confirmatory (blog post selection) aspects of information were confounded with between- (tags) and within patch (blog posts) navigation. However, explicitly analyzing the distinction of between-patch and within-patch navigation may open a new perspective on the whole process of web-based information search, which has been widely neglected, although it is at the center of the information foraging theory, the most influential theory of human information search (Pirolli & Card, 1999).

The present dissertation also offers insights and implications for practitioners who provide and author content, and the construction of content aggregation platforms. The first takeaway is that it is worth considering one's target audience. Does the audience consist of the general population or people who are in an educational context such as university students? For students and in the health-related context I found accuracy motivation to be important. So, in case of university students, emphasizing the expertise of the source of community that collects content will likely lead to the recognition of expertise in students, however, in contrast to the general population, I found no impact of expertise on persuasion in student samples.

Moreover, with student samples it might also be promising to provide metacognitive scaffolds to reduce overconfidence and increase openness to content. For example, "are you confident

about your existing knowledge" similar to studies 2 and 3 presented here; or "carefully reflect on your ability to explain to an expert, in a step-by-step, causally-connected manner, with no gaps in your story how the object works" (Johnson et al., 2016).

In contrast to accuracy motivated *students*, if the target audience consists of the *general population*, I found defense motivation to more important. In study 3, the representative sample did not differentiate the source expertise of the tagging community in social tagging platform. When designing content aggregation platforms for such an audience, designers might consider that expertise does not play an important role when people estimate credibility of the content, but homophily or the similarity to the community collecting content might be more important. Also, in the health context with a representative sample, defense motivation seems to guide information search and evaluation. Therefore, attitude inconsistent content may be more persuasive if existing attitudes of the information searchers are acknowledged upfront, so that the information searcher does not become defensive in the face of attitude inconsistent information.

Conclusion

To conclude, the current dissertation extends the current view of the literature on social tagging behaviour (Cress et al., 2013), by showing that not only semantic associations, but also prior attitudes and consequent confirmation bias have an influence on tag-based navigation. And the current dissertation shows that individual evaluations as well as collective cues of information validity moderate the influence of confirmation bias on navigation and evaluation of information. Finally, in the context of health-related information search, it could be shown that different samples behave differently, providing interesting insights for the design of content aggregation platforms and authors of health-related content on the web.

References

Ames, M., & Naaman, M. (2007). Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 971–980). https://doi.org/10.1145/1240624.1240772

Angermeyer, M. C., van der Auwera, S., Carta, M. G., & Schomerus, G. (2017). Public attitudes towards psychiatry and psychiatric treatment at the beginning of the 21st century: a systematic review and meta-analysis of population surveys. *World Psychiatry*, *16*(1), 50–61. https://doi.org/10.1002/wps.20383

Briñol, P., Petty, R. E., & Barden, J. (2007). Happiness versus sadness as a determinant of thought confidence in persuasion: a self-validation analysis. *Journal of Personality and Social Psychology*, *93*(5), 711–27. https://doi.org/10.1037/0022-3514.93.5.711

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and Systematic Information Processing within and beyond the Persuasion Context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (pp. 212–252). Guilford.

Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology*, *57*, 345–74. https://doi.org/10.1146/annurev.psych.57.102904.190034

Cress, U. (2013). Making use of collective knowledge - A cognitive approach. In *Computer-Supported Collaborative Learning Conference, CSCL* (Vol. 1, pp. 129–136). Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84886554068&partnerID=tZOtx3y1

Cress, U., & Held, C. (2013). Harnessing collective knowledge inherent in tag clouds. *Journal of Computer Assisted Learning*, *29*(3), 235–247. https://doi.org/10.1111/j.1365-2729.2012.00491.x

Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers and Education*, *60*(1), 59–73. https://doi.org/10.1016/j.compedu.2012.06.015

Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2017). Modeling confirmation bias and polarization. *Scientific Reports*, *7*(December 2016), 1–9. https://doi.org/10.1038/srep40391

Fu, W.-T., Kannampallil, T., Kang, R., & He, J. (2010). Semantic imitation in social tagging.

ACM Transactions on Computer-Human Interaction, *17*(3), 1–37.
https://doi.org/10.1145/1806923.1806926

Fu, W., Kannampallil, T., & Kang, R. (2009). A semantic imitation model of social tag
choices. In *Proceedings of the 2009 International Conference on Computational Science
and Engineering - Volume 04 (CSE '09)* (pp. 66–73). Washington, DC: IEEE Computer
Society. https://doi.org/10.1109/CSE.2009.382 http://dx.doi.org/10.1109/CSE.2009.382

Gerjets, P., & Kammerer, Y. (2010). Topical relevance and information quality in cognitive
models of Web search behavior: Introducing epistemic scent into information foraging
theory. In O. & R. Catrambone (Ed.), *Proceedings of the 32nd Annual Conference of the
Cognitive Science Society* (pp. 2278–2283). Austin, TX: Cognitive Science Society.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009).
Feeling validated versus being correct: a meta-analysis of selective exposure to
information. *Psychological Bulletin*, *135*(4), 555–88. https://doi.org/10.1037/a0015701

Heckner, M., Neubauer, T., & Wolff, C. (2008). Tree, funny, to_read, google: What are tags
supposed to Achieve? A comparative analysis of user keywords for different digital
resource types. In *Proceedings of the 2008 ACM workshop on Search in social media*
(pp. 3–10). https://doi.org/10.1145/1458583.1458589

Held, C., Kimmerle, J., & Cress, U. (2012). Learning by foraging: The impact of individual
knowledge and social tags on web navigation processes. *Computers in Human Behavior*,
28(1), 34–40. https://doi.org/10.1016/j.chb.2011.08.008

Holyoke, T. C., & Feynman, R. P. (2010). What Do You Care What Other People Think? *The
Antioch Review*. https://doi.org/10.2307/4612049

Johnson, D. R., Murphy, M. P., & Messer, R. M. (2016). Reflecting on explanatory ability: A
mechanism for detecting gaps in causal knowledge. *Journal of Experimental Psychology:
General*, *145*(5), 573–588. https://doi.org/10.1037/xge0000161

Jorm, A. F., Nakane, Y., Christensen, H., Yoshioka, K., Griffiths, K. M., & Wata, Y. (2005).
Public beliefs about treatment and outcome of mental disorders: a comparison of
Australia and Japan. *BMC Medicine*, *3*. https://doi.org/10.1186/1741-7015-3-12

Kammerer, Y., Amann, D. G., & Gerjets, P. (2015). When adults without university education
search the Internet for health information: The roles of Internet-specific epistemic beliefs
and a source evaluation intervention. *Computers in Human Behavior*, *48*, 297–309.

https://doi.org/10.1016/j.chb.2015.01.045

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2012). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2012.10.012

Klayman, J. (1995). Varieties of Confirmation Bias. *Psychology of Learning and Motivation - Advances in Research and Theory*, *32*(C), 385–418. https://doi.org/10.1016/S0079-7421(08)60315-1

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2270237

Kuo, B. Y.-L., Hentrich, T., Good, B. M. ., & Wilkinson, M. D. (2007). Tag clouds for summarizing web search results. *Proceedings of the 16th International Conference on World Wide Web - WWW '07*, 1203. https://doi.org/10.1145/1242572.1242766

Ley, T., & Seitlinger, P. (2015). Dynamics of human categorization in a collaborative tagging system: How social processes of semantic stabilization shape individual sensemaking. *Computers in Human Behavior*, *51*(PA), 140–151. https://doi.org/10.1016/j.chb.2015.04.053

Liao, Q. V., Fu, W.-T., & Mamidi, S. S. (2015). It Is All About Perspective: An Exploration of Mitigating Selective Exposure with Aspect Indicators. In *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems* (Vol. 1, pp. 1439–1448). https://doi.org/10.1145/2702123.2702570

Lohmann, S., Ziegler, J., & Tetzlaff, L. (2009). Comparison of tag cloud layouts: Task-related performance and visual exploration. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5726 LNCS*(PART 1), 392–404. https://doi.org/10.1007/978-3-642-03655-2_43

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098–2109. https://doi.org/10.1037//0022-3514.37.11.2098

Ma, T. (Jennifer), & Atkin, D. (2017). User generated content and credibility evaluation of online health information: A meta analytic study. *Telematics and Informatics*, *34*(5), 472–486. https://doi.org/10.1016/j.tele.2016.09.009

Mika, P. (2007). Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, *5*(1), 5–15. https://doi.org/10.1016/j.websem.2006.11.002

Muchnik, L., Aral, S., & Taylor, S. J. (2013). Social influence bias: A randomized experiment. *Science*, *341*(6146), 647–651. https://doi.org/10.1126/science.1240466

Nov, O., Naaman, M., & Ye, C. (2008). What drives content tagging: the case of photos on Flickr. *Proceedings of the SIGCHI Conference on …*, 4–7. Retrieved from http://dl.acm.org/citation.cfm?id=1357225

Park, J., Konana, P., & Gu, B. (2013). Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems …*, *24*(4), 1050–1067. https://doi.org/10.1287/isre.2013.0492

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The Meta–Cognitive Model (MCM) of Attitudes: Implications for Attitude Measurement, Change, and Strength. *Social Cognition*, *25*(5), 657–686. https://doi.org/10.1521/soco.2007.25.5.657

Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology*, *82*(5), 722–741. https://doi.org/10.1037//0022-3514.82.5.722

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*(4), 643–675. https://doi.org/10.1037//0033-295X.106.4.643

Schweiger, S., Oeberst, A., & Cress, U. (2014) Confirmation bias in web-based search: a randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of Medical Internet Research*, *16*(3).

Schweiger, S., & Cress, U. (2019). Attitude confidence and source credibility in information foraging with social tags. *PLOS ONE*, *14*(1).

Schweiger, S., & Cress, U. (2019). How Confidence in Prior Attitudes, Social Tag Popularity, and Source Credibility Shape Confirmation Bias Toward Antidepressants and Psychotherapy in a Representative German Sample: Randomized Controlled Web-Based Study. *Journal of Medical Internet Research*, *21*(4), e11081.

Schwind, C., & Buder, J. (2012). Reducing confirmation bias and evaluation bias: When are preference-inconsistent recommendations effective - And when not? *Computers in*

Human Behavior, *28*(6), 2280–2290. https://doi.org/10.1016/j.chb.2012.06.035

Schwind, C., Buder, J., Cress, U., & Hesse, F. W. (2012). Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking? *Computers and Education*, *58*(2), 787–796. https://doi.org/10.1016/j.compedu.2011.10.003

Seitlinger, P. (2012). Implicit imitation in social tagging: familiarity and semantic reconstruction. In J. A. Konstan, E. H. Chi, & K. Höök (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)* (pp. 1631–1640). New York: ACM Press. Retrieved from http://dl.acm.org/citation.cfm?id=2208287

Seitlinger, P., Ley, T., & Albert, D. (2015). Verbatim and semantic imitation in indexing resources on the web: A fuzzy-trace account of social tagging. *Applied Cognitive Psychology*, *29*(1), 32–48. https://doi.org/10.1002/acp.3067

Trant, J. (2009). Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information*, *10*(1), 1–44.

Trattner, C., Kowald, D., Seitlinger, P., Kopeinik, S., & Ley, T. (2016). Modeling Activation Processes in Human Memory to Predict the Use of Tags in Social Bookmarking Systems. *Journal of Web Science*, *2*(1), 1–16. https://doi.org/10.1561/106.00000004

Van Der Auwera, S., Schomerus, G., Baumeister, S. E., Matschinger, H., & Angermeyer, M. (2017). Approval of psychotherapy and medication for the treatment of mental disorders over the lifespan. An age period cohort analysis. *Epidemiology and Psychiatric Sciences*, *26*(1), 61–69. https://doi.org/10.1017/S2045796015001134

Van Overwalle, F., & Siebler, F. (2005). A connectionist model of attitude formation and change. *Personality and Social Psychology Review : An Official Journal of the Society for Personality and Social Psychology, Inc*, *9*(3), 231–74. https://doi.org/10.1207/s15327957pspr0903_3

Vydiswaran, V., Zhai, C., Roth, D., & Pirolli, P. (2012). BiasTrust Teaching Biased Users About Controversial. In *Proceedings of the 21st ACM ...* (pp. 1905–1909). Retrieved from http://dl.acm.org/citation.cfm?id=2398541

Yang, Q., & Beatty, M. (2016). A meta-analytic review of health information credibility: Belief in physicians or belief in peers? *Health Information Management Journal*, *45*(2),

80–89. https://doi.org/10.1177/1833358316639432

Zollers, A. (2007). Emerging Motivations for Tagging: Expression (pp. 8–12).

Appendix

Publication 1

Schweiger, S., Oeberst, A., & Cress, U. (2014). Confirmation bias in web-based search: a randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of medical Internet research*, *16*(3).

Abstract

Background: The public typically believes psychotherapy to be more effective than pharmacotherapy for depression treatments. This is not consistent with current scientific evidence, which shows that both types of treatment are about equally effective.

Objective: The study investigates whether this bias towards psychotherapy guides online information search and whether the bias can be reduced by explicitly providing expert information (in a blog entry) and by providing tag clouds that implicitly reveal experts' evaluations.

Methods: A total of 174 participants completed a fully automated Web-based study after we invited them via mailing lists. First, participants read two blog posts by experts that either challenged or supported the bias towards psychotherapy. Subsequently, participants searched for information about depression treatment in an online environment that provided more experts' blog posts about the effectiveness of treatments based on alleged research findings. These blogs were organized in a tag cloud; both psychotherapy tags and pharmacotherapy tags were popular. We measured tag and blog post selection, efficacy ratings of the presented treatments, and participants' treatment recommendation after information search.

Results: Participants demonstrated a clear bias towards psychotherapy (mean 4.53, SD 1.99) compared to pharmacotherapy (mean 2.73, SD 2.41; t_{173}=7.67, P<.001, d=0.81) when rating treatment efficacy prior to the experiment. Accordingly, participants exhibited biased information search and evaluation. This bias was significantly reduced, however, when participants were exposed to tag clouds with challenging popular tags. Participants facing popular tags challenging their bias (n=61) showed significantly less biased tag selection ($F_{2,168}$=10.61, P<.001, partial eta squared=0.112), blog post selection ($F_{2,168}$=6.55, P=.002, partial eta squared=0.072), and treatment efficacy ratings ($F_{2,168}$=8.48, P<.001, partial eta squared=0.092), compared to bias-supporting tag clouds (n=56) and balanced tag clouds (n=57). Challenging (n=93) explicit expert information as presented in blog posts, compared to supporting expert information (n=81), decreased the bias in information search with regard to blog post selection ($F_{1,168}$=4.32, P=.04, partial eta squared=0.025). No significant effects were found for treatment recommendation (Ps>.33).

Conclusions: We conclude that the psychotherapy bias is most effectively attenuated—and even eliminated—when popular tags implicitly point to blog posts that challenge the widespread view. Explicit expert information (in a blog entry) was less successful in reducing

biased information search and evaluation. Since tag clouds have the potential to counter biased information processing, we recommend their insertion.

Introduction

In the last decade, patients' preferences have increasingly been taken into account when choosing a treatment for depression (SE & JP, 2013), which conforms to American Psychiatric Association guidelines (American Psychiatric Association, 2010). Previous research has demonstrated, however, that laypeople hold beliefs about depression treatment that are partly inconsistent with scientific evidence. They believe, for instance, that psychotherapy is a more effective treatment for depression than pharmacotherapy (Furnham, Pereira, & Rawles, 2001; Priest, Vize, Roberts, Roberts, & Tylee, 1996). In contrast to this, current scientific evidence demonstrates that pharmacotherapy and psychotherapy are nearly equally effective (Cuijpers et al., 2013; De Maat, Dekker, Schoevers, & De Jonghe, 2006). Consequently, the layperson's beliefs are biased.

This study investigates how biases like this one can be reduced. For our study, we chose the domain of depression treatment and made use of the psychotherapy bias. Specifically, we expected that laypeople's bias towards psychotherapy leads to a confirmation bias in information search and evaluation. The confirmation bias refers to the robust findings that individuals tend to process information in a manner that confirms their pre-existing beliefs. Therefore, a confirmation bias in searching for information is not only of interest for depression treatment or the comparison of psychotherapy and pharmacotherapy, but for health-related information search in general. Individual convictions lead to one-sided information processing. When these convictions are not justified by scientific evidence, people run the risk of being misinformed.

Therefore, we investigated two factors that might reduce one-sided information processing. One of the most reliable and objective information sources on the Web is expert information. We tested whether facing explicit expert information would reduce the bias. Moreover, we were interested if aggregated expert information presented in tag clouds would reduce the bias as well.

Blogs and Social Tagging

In the last decade, the Internet has become one of the most important sources for health-related information (Wang et al., 2012). This phenomenon created the need to investigate the communication between experts and laypeople (Kienhues, Stadtler, & Bromme, 2011). Blogs

have been among the most effective applications for disseminating and discussing health-related topics by experts and a general audience. Blogs are authored by and targeted at laypeople as well as health professionals (eg, New York Times Well Blog, Harvard Health Blog), and blogs often report current scientific studies, as well as the author's personal opinion, which can be discussed by the public in the comments section. Moreover, blogs are among the crucial starting points for health-related online information search (Atkinson, Saperstein, & Pleis, 2009).

In order to provide an overview of the relevant content of a blog and to organize related blog posts, popular blogging sites such as Technorati, WordPress, or Counselling Resource include tag clouds or tag lists (Batch, Yusof, & Noah, 2013). We focus on tag clouds (Figure 1.1) because tag clouds provide implicit information on the popularity of topics. Tag clouds display different tags in varying font sizes, according to tag popularity. In broad folksonomies (eg, del.icio.us), which allow not only creators, but also recipients to tag digital artifacts, many people search for the same tags or provide the same tag for numerous blog posts. These co-occurring tags can be displayed in a tag cloud with varying font size, according to the number of co-occurrences.
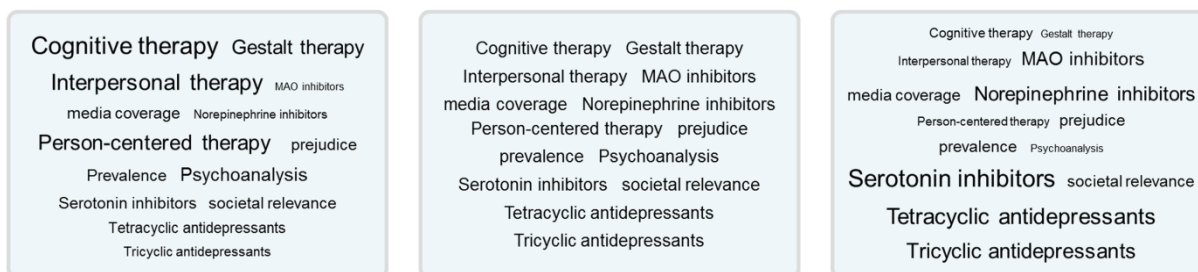
Figure 1.1 Tag cloud versions used in the study.

Tags have two important functions. First, tags organize content. When people provide the same tag for different blog posts, blog posts with a common topic are quickly found via a common tag (eg, the topic with the tag "health" on WordPress). Second, the font size of a tag reflects the popularity of the underlying concept. For example, Figure 1.1 demonstrates the three versions of a tag cloud with the same content, but different popularity of treatments for depressive disorders, used in the current study. Popular tags that represent treatments can be seen at a single glance (Peters & Stock, 2010).

Previous research on the perception of tag clouds has demonstrated that the popularity of tags (presented as tag size) influences information search and information evaluation (Cress, Held, & Kimmerle, 2013; Held, Kimmerle, & Cress, 2012). Popular tags in a tag cloud, for instance, are more frequently selected and their resources more often consulted (Cress & Held, 2013). Popular tags not only guide navigation behavior but also information evaluation. Concepts represented by popular tags are rated as more typical of a domain (Held et al., 2012). Moreover, people align their cognitive concepts to the concepts represented by popular tags. After navigating with tags, people remember more popular concepts compared to less popular concepts (Cress & Held, 2013; Seitlinger, 2012).

Confirmation Bias in Online Information Search

In order to investigate the confirmation bias in health-related information search, we chose the topic of depression treatment with pharmacotherapy or psychotherapy because previous research has demonstrated a discrepancy between laypeople's beliefs and scientific evidence. As mentioned, psychotherapy is viewed to be more effective (Furnham et al., 2001; Priest et al., 1996), whereas scientific evidence points to a comparable efficacy of both treatments (Cuijpers et al., 2013; De Maat et al., 2006). We refer to this misconception as psychotherapy bias. Bias in our conception thus differs from personal preference in that it represents a systematic deviation from scientific knowledge and it describes subjective weightings of information. We expected that users' information search is influenced by their belief that psychotherapy is more effective. Research from the confirmation bias has shown that people confirm their pre-existing beliefs by selecting information that supports those beliefs (Fischer, Schulz-Hardt, & Frey, 2008; Jonas, Schulz-Hardt, Frey, & Thelen, 2001; Kayhan, 2013; Schwind, Buder, Cress, & Hesse, 2012; Vydiswaran, Zhai, Roth, & Pirolli, 2012), for an overview, see (Hart et al., 2009).

The confirmation bias describes people's need to confirm their beliefs and attitudes when engaged in search for online information (Schwind et al., 2012; Vydiswaran et al., 2012). Regarding the psychotherapy bias of laypeople, we expected that when people search for information, they would prefer information about the efficacy of psychotherapy over information about the efficacy of pharmacotherapy. This preference in turn strengthens their prior belief that psychotherapy is effective in treating depression.

Accordingly, our first hypothesis is that the psychotherapy bias—the conviction that psychotherapy is more effective than pharmacotherapy—leads to a confirmation bias in online information search where people prefer to select psychotherapy-related tags and content (H1).

If this confirmation bias determines information search, the question arises as to how the bias can be reduced. Research has shown that people perceive expert information as credible (Kammerer, Bråten, Gerjets, & Strømsø, 2012), and this leads people to align subsequent information search behavior. Therefore, we hypothesized that prior expert information that challenges pre-existing efficacy evaluations, compared to prior expert information that supports pre-existing evaluations, decreases biased information search (ie, tag selection and blog post selection; H2). Likewise, biased information search was expected to decrease with the provision of tag clouds that challenge pre-existent efficacy evaluations. That is, being exposed to tag clouds that have antidepressants as popular tags should decrease the predominant selection of psychotherapy-related tags and blog posts, in comparison to balanced tag clouds and tag clouds with psychotherapy as popular tags (H3). The same bias-reducing effects of challenging (vs supporting) prior expert information (H4) and challenging (vs balanced or supporting) tag clouds (H5) were expected with regard to the evaluation of information. Furthermore, we expected challenging (vs supporting) prior expert information (H6) and challenging (vs balanced or supporting) tag clouds (H7) to lead to a more frequent recommendation of pharmacotherapy.

Results

Recruitment

Participants were recruited via two mailing lists, to which mostly university students from a broad range of disciplines had voluntarily enrolled. They were provided with a link that led them to a fully automated online survey. We reminded all participants twice via email to take part in the study. We did not use cookies or an IP (Internet protocol) check to detect or prevent multiple participation. However, all the provided email addresses were unique. There were no specific eligibility criteria with the exception of computer literacy as an implicit criterion. In order to have an 80% chance to detect a moderate effect (f=0.25), we would require 26 participants per group (a priori analysis of variance [ANOVA] power analysis conducted with G*Power 3.1.5; parameters set to f=0.25, 1–beta=.80, alpha=.05, numerator degrees of freedom=2, 6 groups; (Faul, Erdfelder, Lang, & Buchner, 2007)). The study was conducted within a period of 10 weeks from December 2012 until March 2013 and was stopped after planned sample size was reached in all conditions.

We outlined in the invitation mail that we were conducting a study on the treatment of depression, with the main task of rating short blog posts about different treatment options. We emphasized that participation would be voluntary, could be withdrawn at any point, and that

the study would not cause harm of any kind. We also assured anonymity and the option to withdraw the data at the end of the study without providing reasons. Participants were informed about the duration of the study and the possibility to win €25 or €50 Amazon gift certificates. They were informed that by clicking the next button, they would provide informed consent. Moreover, they were asked to contact the experimenter (email was provided) in case of questions or considerations of any sort. There was no institutional affiliation presented in the invitation mail, but during the online study (see upper left part of Figure 1.2). Ethical approval was provided by the Ethical Committee of the Knowledge Media Research Center (LEK 2012/023).
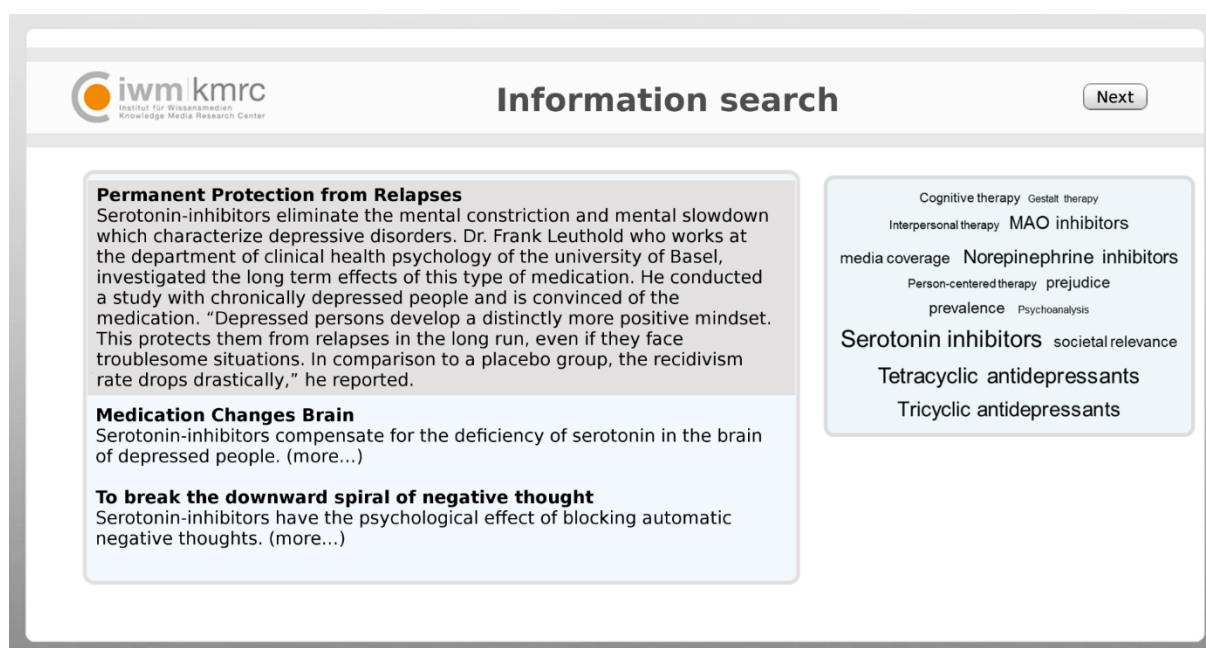
Figure 1.2 Screenshot of information search environment.

Design and Procedure

The study comprised a 2 (prior expert information: supporting, challenging) x 3 (tag popularity: psychotherapy, balanced, pharmacotherapy) between-subjects design. Participants were randomly assigned the following simple randomization procedures (computerized random numbers) to the different treatment groups, with the only restriction that a maximum of 35 individuals (who completed the study) were allowed per condition. We manipulated *prior expert information* by the content of the blogs that participants read before navigating in the tagging environment. Participants read either two blog posts highlighting the efficacy of

psychotherapeutic treatment (supporting) of depressive disorders, or two blog posts highlighting the efficacy of pharmacotherapy (challenging).

As a second factor, we manipulated *tag popularity* by the font size of tags in the tagging environment. In the case of tag popularity, it is not a single resource that explicitly provides a statement regarding the efficacy of a treatment. Rather, the size of the tags implicitly provides insight into the popularity of treatments, as it is seen by experts. Either psychotherapy tags were displayed with a larger font size compared to pharmacotherapy tags, or pharmacotherapy tags were larger, or tags of both types of treatment had the same size (Figure 1.1, middle panel). Importantly, the tag-related blog posts presented during information search were the same across all conditions.

After the first two pages where participants were informed about the study and provided informed consent, the algorithm randomly assigned participants to one of the six conditions and a series of online forms followed. Participants filled out demographic data, followed by questionnaires (eg, prior beliefs about treatment efficacy, cf. measures section).

In the first phase of the experiment, participants read two blog entries. Participants were randomly assigned to read either two blog posts emphasizing the efficacy of psychotherapy (supporting the bias, n=93) or to read two blog posts emphasizing the efficacy of pharmacotherapy (challenging the bias, n=81) in the treatment of depressive disorders. The first blog entry reported that a large global network of "neurologists and psychologists" (expert information) agree on the efficacy of either pharmacotherapy or psychotherapy in the treatment of depression. The second blog entry presented the positive results of a neuroimaging evaluation study, arguing for the respective interpretation. Prior information was held constant, so the reasoning in both conditions was exactly the same; we interchanged only the terms antidepressants and psychotherapy. Note that no comparison to other types of treatment was provided in the blog posts. After each blog post, participants rated its persuasiveness.

After the first phase, participants were informed about the nature of tags and tag clouds. It was stated that tags describe and categorize online content, and an example of a tag cloud was shown. Participants were told that experts provided the tags in the following task. The more often a certain tag had been provided by these experts, the larger the tag in the cloud appeared. Therefore, participants were aware that large tags described popular topics among experts.

In the second phase of the experiment, participants searched for treatment-related information. The task for participants was to find useful information to provide information to a hypothetical friend who suffered from major depressive disorder. After the instructions, the information search environment appeared. Participants were randomly assigned to one of the three versions of a tag cloud (Figure 1.1). The tag cloud either supported psychotherapy bias (psychotherapeutic treatments popular, n=56), or it was neutral with respect to treatment popularity (all treatments equally popular, n=57), or it challenged psychotherapy bias (pharmacological treatments popular, n=61). Participants navigated with the static tag cloud to search information for psychotherapy and pharmacotherapy treatments. When participants clicked on a tag, three short related blog posts appeared to the left of the tag cloud. Blog posts were constant across all experimental conditions. Therefore, all participants had access to the same information. A pilot study (n=32) had assured that blog posts did not differ in persuasiveness, in order to rule out material effects. Tags in the cloud represented different types of treatment, and tag-related blog posts described the efficacy of the respective treatment. After 5 minutes, a stop button appeared at the upper right part of the screen. From this moment, participants could freely choose when to end the information search task. The timer was implemented in order to assure sufficient amount of navigational data.

At the end of the study, all participants were thoroughly debriefed and informed about the fact that the presented materials were not genuine materials and that tag clouds thus did not reflect actual scientific knowledge but had been experimentally designed.

Materials

Content of Prior Expert Information

The two blog posts in the two different conditions of expert information contained matched main arguments for the efficacy of psychotherapy versus pharmacotherapy. Therefore, all blog posts in this study were fictitious. The first blog post in both conditions described the establishment of a database with scientific studies by an extensive and worldwide network of researchers. The second blog post in both conditions described the successful remediation of neuronal brain activity and brain structures, after treatment with either psychotherapy (supporting prior expert information) or pharmacotherapy (challenging prior expert information). Text length ranged from 98 to 118 words.

Tagging Environment

The tagging environment for information search consisted of two main sections (Figure 1.2). At the right side of the screen, 14 tags were presented. Five tags indicated psychotherapy, five tags indicated pharmacotherapy, and four tags were neutral with respect to treatment (media coverage, prejudice, prevalence, societal relevance; Figure 1.1). We varied tag popularity. In the *psychotherapy tags popular* condition, all psychotherapy tags were larger compared to pharmacotherapy tags. In the *pharmacotherapy tags popular* condition, all pharmacotherapy tags were larger compared to psychotherapy tags. In the *balanced tag popularity* condition, all tags had the same size.

At the left side of the screen in the tagging environment, for each tag, related blog posts were presented (Figure 1.2). Three blog posts were related to each tag. The content of the blog posts for pharmacotherapy (15 posts) and psychotherapy (15 posts) was held constant. We composed pairs of psychotherapy and pharmacotherapy blog posts, with the same main arguments and length (mean 76.8 words, SD 6.1) but different wording. Each post described a common symptom of depressive disorders (e.g., psychomotor impairment) and scientific studies reported by an expert. The alleged experts concluded that the studies showed the efficacy of treatment by successfully reporting a remediation of the symptoms. All reported studies referred only to the efficacy of the respective treatment. There was no information available on the comparability of efficacy between pharmacotherapy and psychotherapy. A pilot study (n=32) assured that the blog posts had equal readability and that the persuasiveness and quality of all arguments did not differ within the pairs of blog posts about pharmacotherapy and psychotherapy. Initially, only the headline and the first sentence of each blog post were presented. In order to read the full blog post, participants had to click on the first sentence to expand the blog post.

The tagging environment displayed in the Web browser (programmed with Adobe Flash Builder) was developed by software developers at the Knowledge Media Research Center. The tagging environment was used for the first time; there were no changes of functionality during the period of data collection. Personal information (email address, demographic data) was stored separate from the survey data on a local server.

Measures

Items of all the questionnaires were in fixed order; up to 7 items were displayed per screen. We implemented a completeness check so no items could be skipped by participants. Participants could not use a back button of the browser or within the survey. The measures are described in the order they appear in the experiment.

Prior Knowledge

Prior knowledge about depressive disorders was examined by 24 items regarding general knowledge (eg, false: "Women suffer from depressive disorders as often as men do"; true: "People suffering from diabetes are more likely to suffer also from depressive disorders compared to the general population") and symptoms of depressive disorders according to the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM IV) and the International Classification of Diseases (ICD) 10 (eg, true: "Depressive disorders are often characterized by heightened or lowered appetite"; false: "People with a depressive disorder show an obsessive need for cleanliness and order"). The answer format had the three categories: true/false/I don't know (Cronbach alpha=.72).

Evaluation

Efficacy ratings were inquired for all the treatments that were presented prior to and after the experimental manipulations (see pre- and posttest, Figure 1.3). Five pharmacotherapy treatments and five psychotherapy treatments were rated on a 7-point scale ranging from 1 (not effective) to 7 (highly effective). Prior to the experimental manipulation, we also provided an additional category "I don't know", in case participants were not knowledgeable about the treatment in question (which was coded as 4 on the 7-point scale). A rating bias score was derived by subtracting the sum score of pharmacotherapy from psychotherapy efficacy ratings. If participants did not click on a tag, the respective treatment rating was excluded. The tagging environment produced log files that coded every click in the environment and the respective time. For the posttest ratings, we analyzed only treatments that were viewed by participants for at least 10 seconds according to the log files.
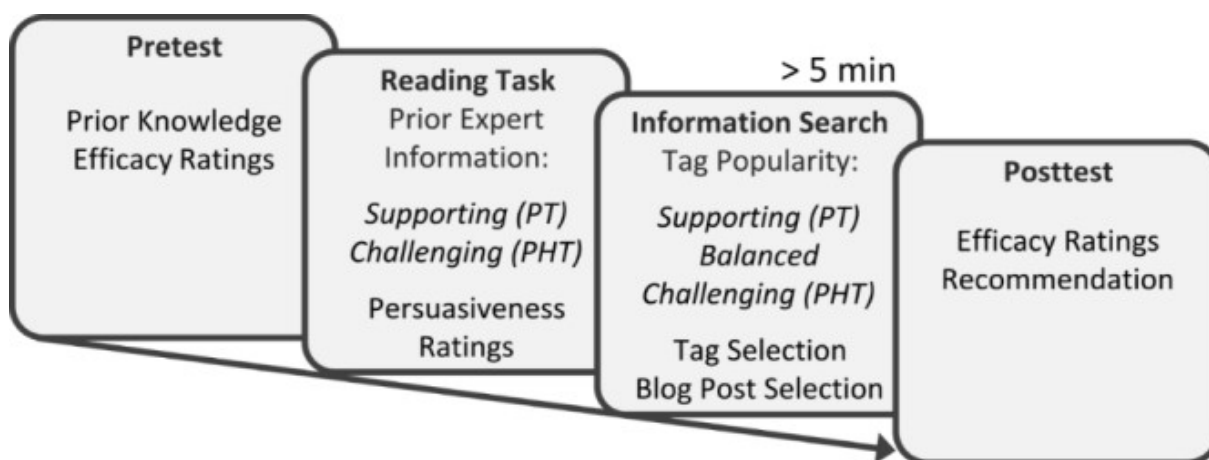

Figure 1.3 Study procedure.

Persuasiveness Ratings of Blog Posts

After reading each of the two prior blog posts, participants rated the degree to which each blog post stated the efficacy of the presented treatment (either psychotherapy or pharmacotherapy) on a 7-point Likert scale (1=I agree not at all, 7=I completely agree). This rating served to ensure that the texts in both prior expert information conditions were equally convincing.

Information Search

In order to analyze the psychotherapy bias in information search, the number of selected pharmacotherapy tags was subtracted from the psychotherapy tags. Thus a positive value represented a searching bias towards psychotherapy. The same procedure was applied to the number of blog posts that participants read.

Recommendation

After the experimental manipulations, participants were asked to provide a treatment recommendation for a hypothetical friend. They were instructed to give reasons for the recommendation in about five sentences. Recommendations were coded from 1-5 (5: recommendation for psychotherapy only, 4: psychotherapy preferred, 3: combination therapy, 2: pharmacotherapy preferred, 1: pharmacotherapy only). At the end of the study, participants had the opportunity to provide qualitative feedback through a feedback form.

Statistical Methods

In order to test our main hypotheses, we conducted a 2 (prior expert information: supporting, challenging) x 3 (tag popularity: psychotherapy, balanced, pharmacotherapy) ANOVA with planned contrasts for the factor tag popularity. With additional t tests, we examined whether participants in the challenging *tag popularity* condition demonstrated any bias in information search at all.

Results

Participants and Dropout Analysis

Initially, 440 individuals followed our invitation and started the online experiment. As can be seen in Figure 1.4, 33.6% (148/440) participants dropped out after the welcome page, and 24.3% (107/440) dropped out during the actual survey. The dropout during the survey is

comparable to other online surveys (Reips, 2002). In addition to these dropouts, we excluded a small number of participants 2.5% (11/440) due to excessive navigation times (see Figure 1.4). This was done in order to assure that the subsequent analysis of information search was not distorted by outliers. Excessive navigation times were detected using the conservative outlier labeling rule (Hoaglin & Iglewicz, 1987). In order to make sure that our results were not specific for the complete cases, we analyzed tag selection and blog post selection for all participants who had participated up to this point and regardless of their navigation duration (50.9%, 224/440). The pattern of results was identical, which argues for the robustness of our findings. Our subsequent report will be based on those participants who completed the study and did not exhibit excessive navigation times (39.5%, 174/440).
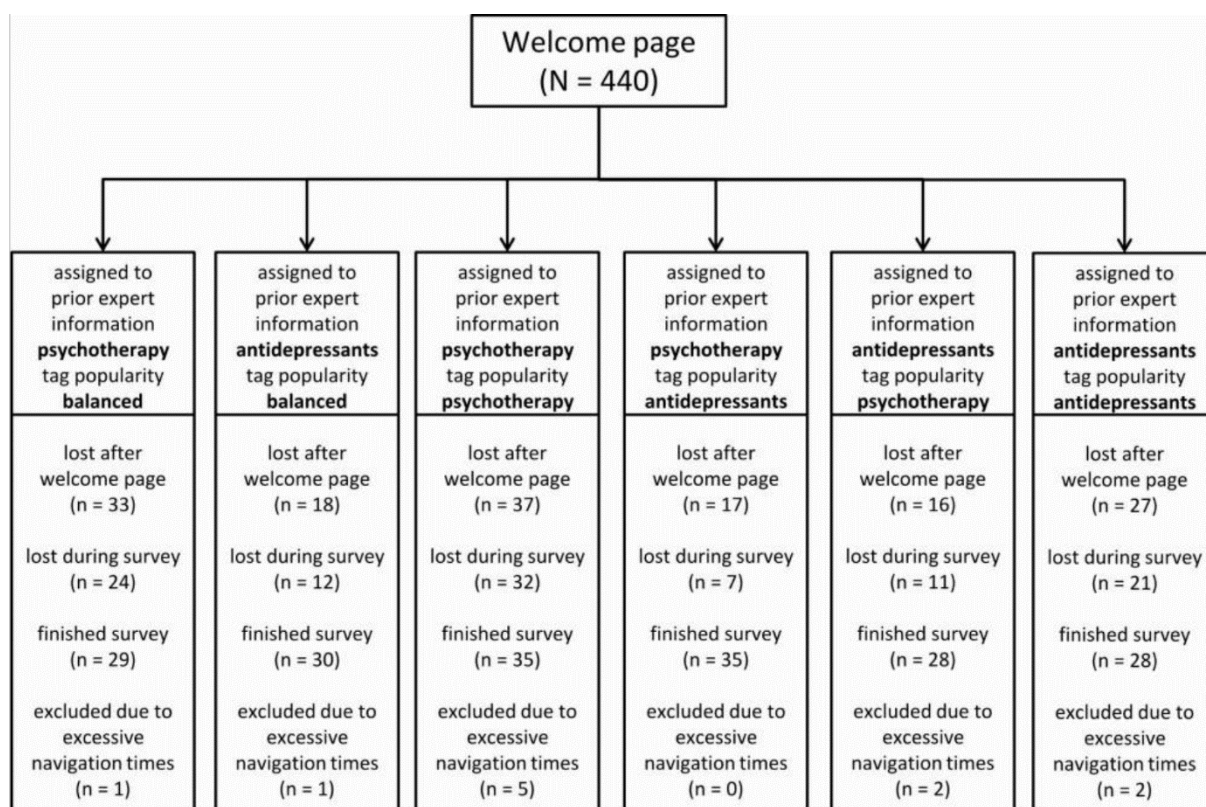
Figure 1.4 Participant flow diagram.

Table 1.1 details the demographics and baseline characteristics of participants. Ages ranged from 16-62 years (mean 23.8, SD 3.8); 74.7% (130/174) were women. Regarding familiarity with the applications under investigation, 44.8% (78/174) stated that they were familiar with social tags, 26.4% (46/174) had knowingly assigned social tags on the Web, 67.2% (117/174) were reading blogs, and 13.8% (24/174) had authored a blog. Most of them were students (74.7%, 130/174) of a non–health care related subject (72.4%, 126/174). A minor proportion had health care related background knowledge due to their field of study (21.3%, 37/174):

psychology, medicine, pharmacy, nursing care, molecular medicine, and neuroscience. It is noteworthy that we reran all analyses without participants from health care related subjects in order to test whether our results hold for laypeople, but the pattern of results was identical.

Table 1.1 Sample characteristics.

Characteristics		n	%
Education			
	Not yet graduated	130	74.7
	Graduated	43	24.7
	No higher education	1	0.6
Field of study, n			
	Health care related subject	37	21.3
	Non–health care related subject	126	72.4
	Not specified	11	6.3
Age			
	15-19	26	14.9
	20-24	97	55.7
	25-29	36	20.7
	30-39	10	5.7
	40-49	4	2.3
	62	1	0.6
Total		174	100

Assuring Equivalence of Groups

First, we checked the equivalence of groups regarding participants' *prior knowledge*. A 2 (prior expert information: supporting, challenging) x 3 (tag popularity: psychotherapy, balanced, pharmacotherapy) ANOVA showed no main effect of tag popularity ($F_{2,168}$=2.32, P=.102, partial eta squared=0.027), and no significant effect of prior expert information ($F_{1,168}$=3.63 P=.06, partial eta squared=0.021). Prior knowledge was not significantly related

to any of the dependent variables (tag selection: $r=-.04$, $P=.62$; blog post selection: $r=-.03$, $P=.66$, efficacy rating: $r=.03$, $P=.66$, recommendation: $r=.06$, $P=.47$), nor was it a significant covariate, nor did prior knowledge as a covariate change the pattern of significance for each dependent variable in separate ANCOVAs. Therefore, we did not include prior knowledge as a covariate in the following analyses.

In order to assure equivalent *treatment intensity of prior expert information*, participants rated persuasiveness of both blog posts on a 7-point scale (1=I don't agree, 7=I completely agree). There was no difference of the persuasiveness ratings between the prior pharmacotherapy expert information group (mean 5.86, SD 1.03) and the prior psychotherapy expert information group (mean 5.91, SD 1.11; $t_{173}=0.27$, $P=.79$, $d=0.08$).

Psychotherapy Bias

In the following analyses, we investigated whether participants showed a psychotherapy bias regarding pre-existent beliefs (H1). To this end, we analyzed efficacy ratings of psychotherapy and pharmacotherapy that had been assessed prior to the information search. Efficacy ratings on a scale ranging from 1-7 showed that participants expressed strong superiority of psychotherapy (mean 4.53, SD 1.99) over pharmacotherapy (mean 2.73, SD 2.41; $t_{173}=7.67$, $P<.001$, $d=0.81$) with regard to the treatment of depression. Thus the participants of our study clearly demonstrated a psychotherapy bias (Furnham et al., 2001; Priest et al., 1996). In the following sections, we will show how the bias influenced information processing and what factors affected the bias.

Information Search

We first tested whether the psychotherapy bias emerges in information search (H1). This hypothesis was confirmed, since participants generally selected more psychotherapy tags (mean 4.66, SD 2.28) compared to pharmacotherapy tags (mean 3.87, SD 3.35; $t_{173}=2.83$, $P=.005$, $d=0.25$). Further support was provided by the fact that participants selected more psychotherapy blog posts (mean 7.02, SD 4.47) compared to pharmacotherapy blog posts (mean 4.21, SD 3.97; $t_{173}=6.47$, $P<.001$, $d=0.66$).

Beyond demonstrating the biased information search behavior, we hypothesized that the psychotherapy bias is reduced by providing prior expert information (H2) and popular tags (H3) that challenge the psychotherapy bias. We will report two separate 2 (prior expert information: supporting, challenging) x 3 (tag popularity: psychotherapy, balanced, pharmacotherapy) ANOVAs for tag selection on one hand, and blog post selection on the

other. With regard to *tag selection*, the analysis did not yield a significant main effect of *prior expert information* ($F_{1,168}$=.32, P=.57, partial eta squared=0.002). There was no tendency of participants to prefer either pharmacotherapy or psychotherapy tags when prior expert information challenged or supported psychotherapy bias (Figure 1.5, left panel). There was, however, a significant main effect of *tag popularity* ($F_{2,168}$=10.61, P<.001, partial eta squared=0.112). A polynomial contrast analysis showed that there was a linear trend of selection bias across the tag popularity conditions (P<.001; Figure 1.5, left panel). Psychotherapy tag selection was higher in the condition with psychotherapy tags being popular compared to the balanced condition (Cohen's d=0.49) and the pharmacotherapy popular condition (Cohen's d=0.85). The interaction between prior expert information and tag popularity ($F_{2,168}$=.02, P=.98, partial eta squared<0.001) was not significant.
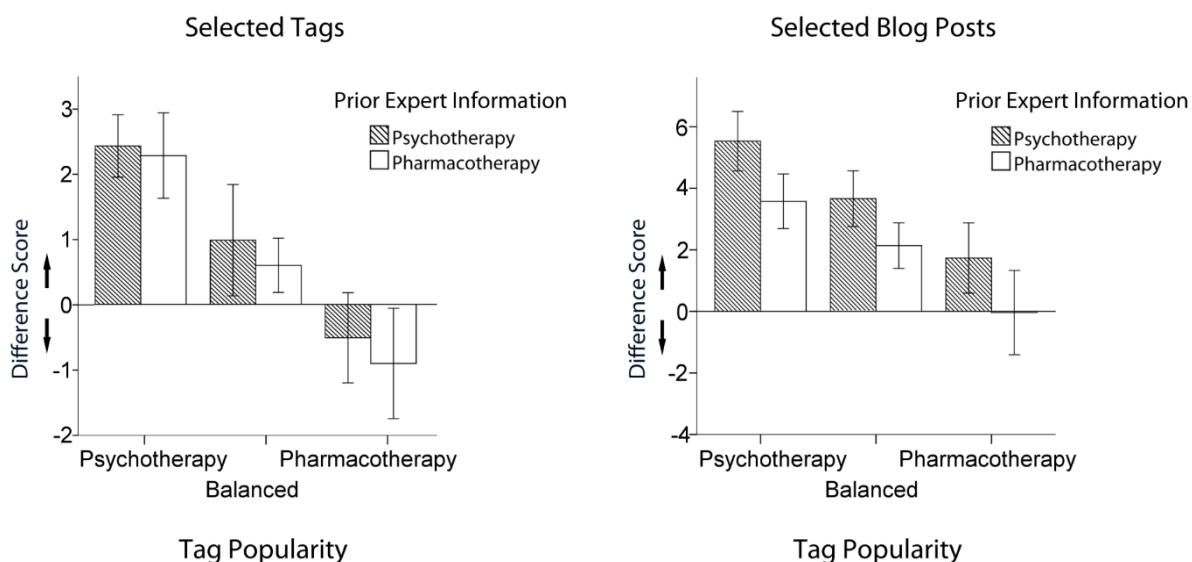
Figure 1.5 Information search bias (pharmacotherapy scores subtracted from psychotherapy scores; positive scores indicate a preference for psychotherapy over pharmacotherapy; negative scores indicate a preference of pharmacotherapy over psychotherapy).

With regard to the second dependent measure of information search, *blog post selection*, a separate 2 x 3 ANOVA revealed a significant effect of *prior expert information* ($F_{1,168}$=4.32, P=.04, partial eta squared=0.025). Reading a prior blog post that challenged the psychotherapy bias led participants to read more pharmacotherapy blog posts during their navigation in the tag cloud (Cohen's d=0.30; Figure 1.5, right panel). The ANOVA also

showed a main effect of *tag popularity* on biased *blog post selection* ($F_{2,168}$=6.55, *P*=.002, partial eta squared=0.072). A polynomial contrast analysis showed that there was a linear trend of selection bias across the tag popularity conditions (*P*<.001; Figure 1.5, right panel). Psychotherapy blog post selection was higher in the psychotherapy tags popular condition compared to the balanced condition (Cohen's *d*=0.49) and the pharmacotherapy popular condition (Cohen's *d*=0.85). The interaction of prior expert information and tag popularity was not significant ($F_{2,168}$=.02, *P*=.98, partial eta squared<0.001).

In an additional analysis, we exploratively examined whether participants in the challenging *tag popularity* condition exhibited any bias in information search at all. As indicated by *t* tests, this was not the case. Neither *tag selection* nor *blog post selection* were significantly biased: *P*s>14.

Taken together, we found evidence for a confirmation bias with participants selecting significantly more resources that were consistent with their previously held beliefs that psychotherapy is more effective. Our results also demonstrate, however, that this biased information selection can be significantly reduced. Whereas prior expert information reduced the biased selection of blog posts (but not of tags), tag popularity affected both measures of information search. Being exposed to a tag cloud that contained pharmacotherapy tags as the most popular ones did not only significantly decrease the biased search, but eventually eliminated the confirmation bias in that participants selected as many tags and resources of both treatment types. Hence, challenging tag clouds led to a balanced (ie, unbiased) information search.

Evaluation of Information

With regard to information evaluation, we hypothesized that prior expert information (H4) that challenges the psychotherapy bias decreases biased evaluation of information, compared to prior expert information, which confirms psychotherapy bias. We also expected popular tags (H5) that challenge psychotherapy bias to reduce biased evaluation of information, compared to balanced tag popularity and even more compared to popular tags that support the bias. In order to analyze both hypotheses, we conducted a 2 (prior expert information: supporting, challenging) x 3 (tag popularity: psychotherapy, balanced, pharmacotherapy) ANOVA, with efficacy ratings as the dependent measure. The main effect of *prior expert information* (H4) on biased *efficacy rating* failed to reach conventional significance levels ($F_{1,168}$=2.93, *P*=.09, partial eta squared=0.017). Prior expert information that challenged

psychotherapy bias failed to significantly decrease biased information evaluation compared to prior expert information that confirms the bias.

Popularity of tags challenging psychotherapy bias, in contrast, decreased biased information evaluation as indicated by a significant main effect of *tag popularity* on *evaluation of information* ($F_{2,168}$=8.48, P<.001, partial eta squared=0.092). A polynomial contrast analysis showed that there was a linear trend of evaluation bias across the tag popularity conditions (P<.001; Figure 1.6). Psychotherapy bias in treatment evaluation was higher in the psychotherapy tags popular condition compared to the balanced condition (Cohen's d=0.35) and the pharmacotherapy popular condition (Cohen's d=0.77). The interaction of prior expert information and tag popularity was not significant ($F_{2,168}$=.18, P=.84, partial eta squared=0.002).
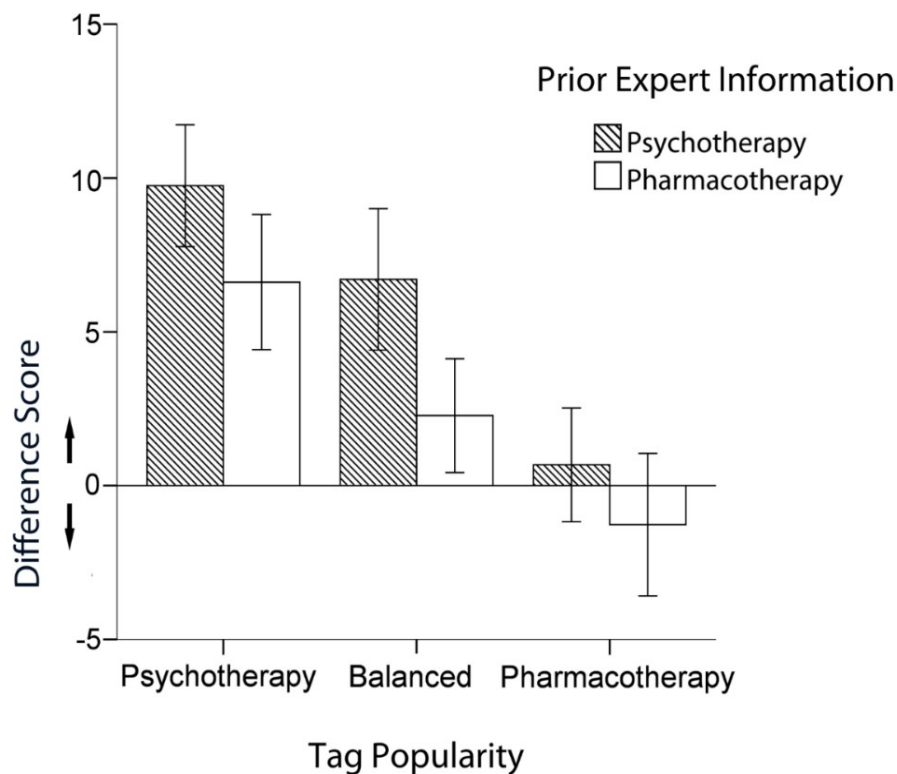

Figure 1.6 Efficacy ratings of blog posts (pharmacotherapy scores subtracted from psychotherapy scores; positive scores indicate a preference for psychotherapy over pharmacotherapy; negative scores indicate a preference of pharmacotherapy over psychotherapy).

Further explorative analyses supported what can be derived from Figure 1.6 already. Efficacy ratings after the information search task were no longer biased in the challenging *tag popularity* condition (t_{33}=0.37, *P*=.72 in the supporting *prior expert information* condition and t_{25}=0.55, *P*=.59 in the challenging *prior expert information* condition).

In sum, our interventions were differentially successful in reducing the confirmation bias with regard to the evaluation of information. Whereas prior expert information failed to exert a significant influence, tag clouds with tags that challenged the psychotherapy bias not only reduced biased information evaluation, but eventually eliminated any bias. Efficacy ratings in this condition were thus eventually in line with scientific evidence.

Recommendation

Beyond information selection and evaluation, we expected that *prior expert information* (H6), as well as *tag popularity* (H7) that challenges the psychotherapy bias, to decrease biased treatment recommendation for a hypothetical friend. We conducted an additional 2 (prior expert information: challenging, supporting) x 3 (tag popularity: psychotherapy, balanced, pharmacotherapy) ANOVA with treatment recommendation as the dependent variable. The results showed neither a significant main effect of tag popularity ($F_{2,168}$=.22, *P*=.81, partial eta squared=0.003) nor a significant main effect of prior expert information ($F_{1,168}$=.97, *P*=.33, partial eta squared=0.006). The interaction was also not significant ($F_{1,168}$=.08, *P*=.92, partial eta squared=0.001). Overall, prior expert information and tag popularity had no effect on recommendation. Figure 1.7 shows that most of the participants recommended psychotherapy.
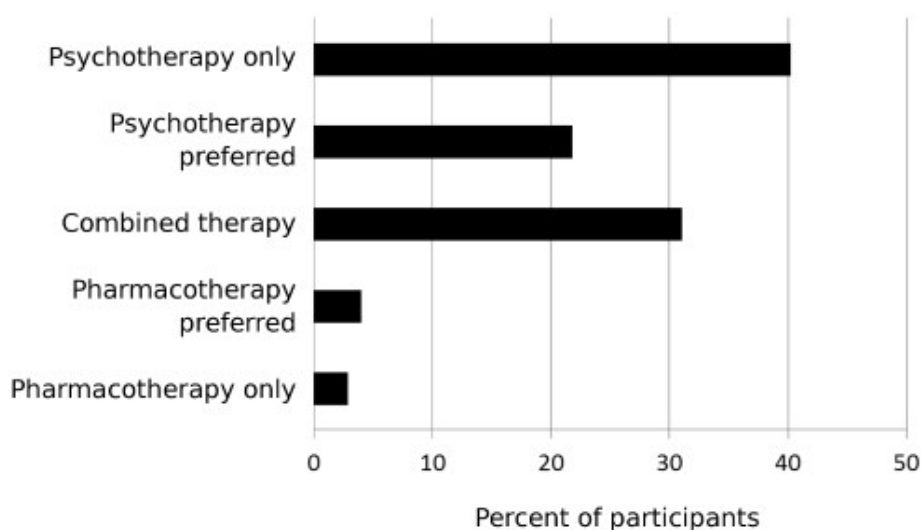

Figure 1.7 Treatment recommendation after experiment.

We conducted an exploratory qualitative analysis of the reasons for the treatment recommendation. Most of the participants did not provide any reasons but among those who did, the most frequently mentioned aspects regarded etiology or negative consequences of antidepressants. Specifically, 16.7% (29/174) participants argued for psychotherapy because they were convinced that biographical and social causes are crucial for the causation and treatment of depression. Another 10.3% (18/174) participants mentioned side effects, and 6.3% (11/174) reasoned that antidepressants are addictive. Finally, 4.6% (8/174) revealed that they believed that overcoming depression is an act of will or a personal responsibility.

Discussion

This study investigated potential measures to decrease biased beliefs and their influence on information selection and information evaluation. To this end, we made use of laypeople's (erroneous) convictions that psychotherapy is more effective in treating depression and examined whether this conviction guides online information search. In line with prior findings, participants did believe in the superiority of psychotherapeutic treatment and thus exhibited a psychotherapy bias. When searching for information online about the treatment of depressive disorders, participants showed a general bias towards selecting psychotherapy treatments compared to pharmacotherapy treatments.

We took two measures to reduce biased information processing. First, we exposed participants to expert information explicitly challenging the superiority of psychotherapy, by demonstrating the effectiveness of pharmacotherapy. This manipulation led participants to select fewer blog posts that were related to psychotherapy compared to the presentation of expert information supporting the effectiveness of psychotherapy. It did not affect, however, tag selection, and there was only a trend for it to exert an influence upon subsequent efficacy ratings. Hence, explicit expert information was only partially successful in reducing biased information processing.

Second, we attempted to decrease biased information processing by presenting participants with tag clouds in which the most popular tags referred to pharmacotherapy (vs psychotherapy). Consistent with our hypotheses, participants in the pharmacotherapy condition selected these popular pharmacotherapy tags more frequently and read more of the underlying blog posts. Moreover, treatment efficacy ratings were affected. In contrast to our expectations, however, we did not find any effects on treatment recommendations.

Although both manipulations had an impact upon search behavior and efficacy evaluation, the manipulations did not exert an impact on providing recommendations to other people. The gap between the efficacy ratings and treatment recommendations might be due to other beliefs people have with regard to both therapies, such as side effects (Althaus, Stefanek, Hasford, & Hegerl, 2002; Angermeyer, Däumer, & Matschinger, 1993). Participants might be convinced that pharmacotherapy is effective, but they might still have feared detrimental side effects. The reasoning of participants supported this notion, as they frequently referred to side effects and even addictiveness of antidepressants when justifying their recommendation. This might indicate that even if a part of the beliefs changed (ie, the efficacy beliefs), other beliefs (eg, about side effects) still have a strong impact on the overall evaluation of a treatment. This is likely to be based on multiple aspects with efficacy being only one of them. Nevertheless, because our primary aim was to reduce laypeople's misconceptions and to counter their biased information processing, we had primarily focused on treatment efficacy. After all, their beliefs had been shown to stand in contrast to scientific evidence. And it was due to this focus that all of our materials concerned treatment efficacy only. With regard to this misconception, however, our findings clearly argue for a success. Tag clouds with challenging popular tags were able to not only reduce biased information search and evaluation, but eventually led to an unbiased search and evaluation. That is, we were able to completely eliminate laypeople's bias regarding treatment efficacy.

Theoretical Implications

Previous research on confirmation bias has shown that people's prior beliefs influence their information search in a way that they seek to confirm their beliefs (Hart et al., 2009; Schwind & Buder, 2012; Schwind et al., 2012). The present study showed that implicit presentation of expertise is even more effective than the explicit one. Earlier research (Seitlinger, 2012) showed that tag semantics and popularity determine individual information processing behavior. Likewise, previous studies successfully showed that social tags influence information selection, evaluation, incidental learning (Cress et al., 2013; Held et al., 2012), and conceptual memory representations (Cress & Held, 2013; Seitlinger, 2012). The findings of the current study extend existing evidence by showing that expert information exerts an even larger influence on users' beliefs, if it is presented implicitly such as in tag clouds compared to explicit presentations as in blog posts alone. This finding has some practical implications.

Practical Implications

In order to make people more aware of expert information and to overcome their individual biases, it seems to be useful to provide them with tag clouds. If these tag clouds challenge their subjective beliefs, users are motivated to select more popular tags (that are inconsistent with their own beliefs) and to read more information challenging their own views. This leads to a reduced confirmation bias, not just with regard to information search, but also with regard to evaluation.

A "correction" of subjective biases can only be achieved, however, if the information provided is not also biased. Thus, whether the effect that tag clouds have is really positive depends on the quality of tags and resources: does tag popularity really represent the scientific knowledge about a topic? In order to ensure that, it is important that people with high expertise provide the resources and tags. The provision of such expert information could be fostered if experts were encouraged to publish scientific studies in a style suitable for a broad audience, as this is already sufficient to reduce biased attitudes.

Limitations

In the current study, we carefully balanced the quality of arguments for both types of treatment. We therefore provided information only about the efficacy of treatments, not about other aspects such as side effects, which would be specific for each treatment. For future studies, it may be desirable to test this in more depth by including diagnostic information with respect to relative efficacy of both treatment types (eg, information on treatments that are less effective compared to others or placebo), as well as providing information on side effects or other treatment-specific information.

Second, it must be pointed out that the present sample consisted mainly of university students or persons with a degree in higher education. Some of our participants had a health care related background. Our analyses showed, however, that the pattern of results was identical when these more knowledgeable participants were excluded. Hence, our findings should be valid with regard to laypeople. Nevertheless, future studies should also include participants without a higher education, as well as older persons.

Conclusion

Our major aim in this study was to investigate whether people exhibit a biased online information search behavior that is guided by biased beliefs. We examined the biased perception of laypersons that psychotherapy is more effective than pharmacotherapy, when it comes to the treatment of depression (Furnham et al., 2001; Priest et al., 1996). We do not

believe that our results are limited to the topic of depression or the pharmacological or psychological treatments. Rather, we would suggest that for any health-related issue involving different accounts or treatments, information challenging users' prior knowledge and attitudes may increase their understanding of the topic in question (Angermeyer et al., 1993; Bientzle, Cress, & Kimmerle, 2013).

References

Althaus, D., Stefanek, J., Hasford, J., & Hegerl, U. (2002). [Knowledge and attitude of the general public regarding symptoms, etiology and possible treatments of depressive illnesses]. *Nervenarzt*.

American Psychiatric Association. (2010). American Psychiatric Association Practice Guidelines for the Treatment of Patients With Major Depressive Disorder. *Practice*, *150*(October), 1–28. https://doi.org/10.1176/appi.books.9780890423387.654001

Angermeyer, M. C., Däumer, R., & Matschinger, H. (1993). Benefits and risks of psychotropic medication in the eyes of the general public: results of a survey in the Federal Republic of Germany. *Pharmacopsychiatry*, *26*(4), 114–20. https://doi.org/10.1055/s-2007-1014354

Atkinson, N. L., Saperstein, S. L., & Pleis, J. (2009). Using the internet for health-related activities: Findings from a national probability sample. *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.1035

Batch, Y., Yusof, M. M., & Noah, S. A. (2013). ICDTag: a prototype for a web-based system for organizing physician-written blog posts using a hybrid taxonomy-folksonomy approach. *Journal of Medical Internet Research*, *15*(2), e41. https://doi.org/10.2196/jmir.2353

Bientzle, M., Cress, U., & Kimmerle, J. (2013). How students deal with inconsistencies in health knowledge. *Medical Education*, *47*(7), 683–90. https://doi.org/10.1111/medu.12198

Cress, U., & Held, C. (2013). Harnessing collective knowledge inherent in tag clouds. *Journal of Computer Assisted Learning*, *29*(3), 235–247. https://doi.org/10.1111/j.1365-2729.2012.00491.x

Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers*

and Education, *60*(1), 59–73. https://doi.org/10.1016/j.compedu.2012.06.015

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry*, *12*(2), 137–148. https://doi.org/10.1002/wps.20038

De Maat, S., Dekker, J., Schoevers, R., & De Jonghe, F. (2006). Relative efficacy of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Psychotherapy Research*, *16*(5), 566–578. https://doi.org/10.1080/10503300600756402

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17695343

Fischer, P., Schulz-Hardt, S., & Frey, D. (2008). Selective exposure and information quantity: how different information quantities moderate decision makers' preference for consistent and inconsistent information. *Journal of Personality and Social Psychology*, *94*(2), 231–44. https://doi.org/10.1037/0022-3514.94.2.94.2.231

Furnham, A., Pereira, E., & Rawles, R. (2001). Lay theories of psychotherapy: perceptions of the efficacy of different "cures" for specific disorders. *Psychology, Health and Medicine*, *6*(1), 77–84. https://doi.org/10.1080/13548500020021946

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–88. https://doi.org/10.1037/a0015701

Held, C., Kimmerle, J., & Cress, U. (2012). Learning by foraging: The impact of individual knowledge and social tags on web navigation processes. *Computers in Human Behavior*, *28*(1), 34–40. https://doi.org/10.1016/j.chb.2011.08.008

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association*. https://doi.org/10.1080/01621459.1987.10478551

Jonas, E., Schulz-Hardt, S., Frey, D., & Thelen, N. (2001). Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social*

Psychology, *80*(4), 557–71. Retrieved from
http://www.ncbi.nlm.nih.gov/pubmed/11316221

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2012). The role of Internet-specific
epistemic beliefs in laypersons' source evaluations and decisions during Web search on a
medical issue. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2012.10.012

Kayhan, V. O. (2013). Seeking health information on the web: Positive hypothesis testing.
International Journal of Medical Informatics.
https://doi.org/10.1016/j.ijmedinf.2012.12.004

Kienhues, D., Stadtler, M., & Bromme, R. (2011). Dealing with conflicting or consistent
medical information on the web: When expert information breeds laypersons' doubts
about experts. *Learning and Instruction*.
https://doi.org/10.1016/j.learninstruc.2010.02.004

Peters, I., & Stock, W. G. (2010). "Power tags" in information retrieval. *Library Hi Tech*,
28(1), 81–93. https://doi.org/10.1108/07378831011026706

Priest, R. G., Vize, C., Roberts, a, Roberts, M., & Tylee, a. (1996). Lay people's attitudes to
treatment of depression: results of opinion poll for Defeat Depression Campaign just
before its launch. *BMJ (Clinical Research Ed.)*, *313*(7061), 858–859.
https://doi.org/10.1136/bmj.313.7061.858

Reips, U.-D. (2002). Internet-Based Psychological Experimenting: Five Dos and Five Don'ts.
Social Science Computer Review, *20*(3), 241–249.
https://doi.org/10.1177/089443930202000302

Schwind, C., & Buder, J. (2012). Reducing confirmation bias and evaluation bias: When are
preference-inconsistent recommendations effective - And when not? *Computers in
Human Behavior*, *28*(6), 2280–2290. https://doi.org/10.1016/j.chb.2012.06.035

Schwind, C., Buder, J., Cress, U., & Hesse, F. W. (2012). Preference-inconsistent
recommendations: An effective approach for reducing confirmation bias and stimulating
divergent thinking? *Computers and Education*, *58*(2), 787–796.
https://doi.org/10.1016/j.compedu.2011.10.003

SE, W., & JP, B. (2013). Should treatment for depression be based more on patient
preference? *Patient Preference and Adherence*, *7*, 1047–1057.
https://doi.org/http://dx.doi.org/10.2147/PPA.S52746

Seitlinger, P. (2012). Implicit imitation in social tagging: familiarity and semantic reconstruction. In J. A. Konstan, E. H. Chi, & K. Höök (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)* (pp. 1631–1640). New York: ACM Press. Retrieved from http://dl.acm.org/citation.cfm?id=2208287

Vydiswaran, V. G. V., Zhai, C., Roth, D., & Pirolli, P. (2012). BiasTrust: teaching biased users about controversial topics. In *Proceedings of the 21st ACM international conference on Information and knowledge management - CIKM '12*. https://doi.org/10.1145/2396761.2398541

Wang, L., Wang, J., Wang, M., Li, Y., Liang, Y., & Xu, D. (2012). Using internet search engines to obtain medical information: A comparative study. In *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.1943

Publication 2

Schweiger, S. & Cress, U. (2019). Attitude confidence and source credibility in information foraging with social tags. *PLoS ONE*, 14(1).

Publication 2

Abstract

There is growing concern that online information searchers are overconfident and therefore largely search for information which reinforces their prior attitudes, blinded by confirmation bias. This study tests if this effect can be reduced in content aggregation platforms, when social tag clouds show popular topics among experts. We manipulated (1) confidence in prior attitudes, and (2) the credibility of the expert community that tagged the content. We found that both factors influence navigation in different ways. First, attitude confidence moderated the influence of prior attitudes when choosing how much attitude-consistent content in blog posts to read. When attitude confidence was high, prior attitudes were positively associated with selection of blog posts, when low, not positively associated. After navigation, when confidence was high, the content of attitude-consistent blog posts was more favourably evaluated, whereas when confidence was low, attitude inconsistent blog posts were more favourably evaluated. Second, source credibility moderated the influence of prior attitudes on tag selection. When source credibility was low, prior attitudes did guide tag selection, when high, they did not. With low source credibility, people selected more attitude-consistent content. The findings advance social tagging theories by showing that not only semantic associations, but also attitudes play a role when people select and process tags and related content. The findings also show that credibility and confidence have a different impact on different stages of information selection and evaluation. Whereas credibility is more important when switching among pages, attitude confidence is more important when reading and evaluating the content of one page.

Introduction

In online environments, it has been suggested that we often find ourselves in a filter bubble or echo chamber, where we only receive and attend to information that is consistent with our views and prior attitudes (Hart et al., 2009; Klayman, 1995; Knobloch-Westerwick, Sharma, Hansen, & Alter, 2005; Liao & Fu, 2013; Liao, Fu, & Mamidi, 2015; Nickerson, 1998; Schweiger, Oeberst, & Cress, 2014; Schwind, Buder, Cress, & Hesse, 2012; Vydiswaran, Zhai, Roth, & Pirolli, 2012; White & Horvitz, 2015). This is called confirmation bias (Klayman, 1995; Nickerson, 1998), and its robustness has been extensively documented by literature see (Klayman, 1995; Nickerson, 1998) for an overview and (Hart et al., 2009) for a meta-analysis). Confirmation bias is particularly prevalent and pertinent in online content platforms (Liao & Fu, 2013; Liao et al., 2015; Schweiger et al., 2014; Schwind et al., 2012;

Vydiswaran et al., 2012), and it fundamentally shapes our search for, and evaluation of information (Klayman, 1995; Klayman et al., 1987). Confirmation bias might possibly set us on a particularly harmful track, for example, when we search for health-related information online (Schweiger et al., 2014).

Online search is so important that people still prefer to search online for an answer to a question even if they have an answer at hand (Ferguson, McLean, & Risko, 2015). This also suggests that searchers have less confidence in their own knowledge than in information they find online (Ferguson et al., 2015). But people are also exceedingly confident about their knowledge in spite of actual limitations to that knowledge. Such confidence has been exhibited in a large range of domains (Zell & Krizan, 2014). These findings show the central role confidence plays in online information search. Thus, the question of the consequences of confidence on confirmation bias is a fundamental one. In the study presented here, we investigate whether confirmation bias increases when searchers are highly confident and perceive their own attitudes as highly valid.

Besides confidence, perceived validity of one's own attitudes and perceived validity of the online community's knowledge could also influence confirmation bias. However, this implies that searchers are equipped with the skill of recognizing whether a source is credible or not. There is mixed evidence from different online platforms whether people succeed in taking source credibility adequately into account. A recent review found in the context of health-related search that homophily drives credibility evaluations of user-generated content (Ma & Atkin, 2017). That is, blog posts and health forum entries are evaluated as more credible when authors of blog posts or forum entries have demographics similar to those of the searcher (Ma & Atkin, 2017). But in social media or content aggregation platforms, the creator of the content often differs from the person or community which collects and shares content. We test whether searchers evaluate credibility correctly in terms of expertise of the community. And we ask the question whether the perception that the community is highly credible makes the searcher more open to content in spite of their prior attitudes.

Finally, we explore how confidence and credibility shape confirmation bias in different stages of information search. To do this, we draw on the most influential theory of human information search, the information foraging theory (Pirolli & Card, 1999), as well as the extended information scent model (Cress, Held, & Kimmerle, 2013). Information foraging theory distinguishes between the breadth and depth of navigation. There are two fundamental search processes: the *navigation among* information patches (e.g., different websites) and the

uptake of information within one information patch (e.g., a single website). With respect to confirmation bias, this has far reaching implications: Individuals with higher attitude confidence may be prone to explore more attitude-consistent information patches, and dwell on attitude-consistent information within each patch. On the other hand, when searchers perceive the community that created and tagged the content as more credible and knowledgeable, they may be open to move among patches and read more content within single patches, independently of their prior attitudes. Interestingly, to our knowledge, no studies thus far have investigated confirmation bias with respect to both of these online search activities. In addition, we explore whether factors such as credibility or attitude confidence influence both information search activities.

In sum, our expectations are the same for both selection activities (between-patch and within-patch). As for the health-related context, we expect people to try to find the objectively best possible treatment, not a treatment that validates their self-concept (Hart et al., 2009). So, when searchers doubt the objective correctness of their prior attitudes because of lowered attitude confidence, they should be more open to attitude-inconsistent information. When searchers hold their prior attitudes with high confidence, they should select more attitude-consistent information. On the other hand, when searchers doubt the objective correctness of the information provided by a tagging community where source credibility may be low, they should select less information in general, independently of prior attitudes. In the following section we will review literature relevant to these issues.

Attitude confidence and confirmation bias

People are, in general, exceedingly confident in the correctness of their knowledge and attitudes a confidence which is often inconsistent with their actual abilities. This is a phenomenon found in a wide range of domains, such as academic, intellectual, vocational, athletic, or even medical domains (Dunning, Heath, & Suls, 2004, 2011; Kruger & Dunning, 1999; Moore & Healy, 2008; Zell & Krizan, 2014). At the same time, people don't trust the confidence in their own knowledge enough, instead preferring to search online for an answer to a question, even when they already know the answer (Ferguson et al., 2015). Moreover, attitude confidence varies in different situations and domains. For example, when searching for health-related information, people may find themselves becoming uncertain when they must cope with diagnosis or a disease (Mishel, 1997; Neville, 2003). Having online information at hand even appears to increase overconfidence, as information searchers mistake access to information as having knowledge (Fisher, Goddu, & Keil, 2015). So, the

question arises, how does heightened or lowered confidence diminish or mitigate confirmation bias?

Confidence, in general, has manifold consequences for selection and judgment of information. For example, people who are highly confident will select information in line with their prior attitudes (Kruglanski & Webster, 1996). On the other hand, a meta-analysis has shown that high confidence tends to decrease confirmation bias, since people don't experience any threat to their own point of view (low defense motivation (Hart et al., 2009)). It should be noted, however, that the meta-analysis also presented a number of studies where high confidence did not decrease but instead increased confirmation bias (Hart et al., 2009).

We expect that high attitude confidence would increase confirmation bias. This is because when we manipulate the metacognitive aspect of attitude confidence, we assume that defense motivation would not increase (Hart et al., 2009), but that the individual's estimation of the validity of their own prior attitudes would increase. When individuals view their prior attitudes to be valid, confirmation bias should increase. When they estimate their attitudes to be invalid, confirmation bias should decrease. This is in line with the accuracy motivation theory, which in the health-related context, implies that information searchers should be motivated to search for objectively correct, valid information.

To illustrate this, there is only one study that used the same manipulation of attitude confidence as we have used it (Petty, Briñol, & Tormala, 2002). So in this respect, the study presented here is a replication attempt. In the original study using manipulation of confidence, students first read a strong or weak version of a persuasive message in favor of comprehensive exams. In response to this, they were asked to provide their thoughts about the message. Then, for an alleged unrelated study, students were instructed to think back to situations in which they experienced doubt or confidence about their own thoughts. They were subsequently instructed to think back about the thoughts they had had in response to the persuasive message they had just experienced. Students who had recalled previous situations in which they were highly confident were more confident about their thoughts regarding the message than students who had recalled previous situations in which they were doubtful. The degree of confidence also had consequences for the degree of persuasion. Participants who were highly confident about their thoughts in response to the message were also more persuaded in line with their thoughts. This study showed that high metacognitive confidence led to more confirmation bias with respect to content evaluation. Moving beyond this result,

we go further here to test if high metacognitive confidence will lead to more attitude consistent information selection.

To conclude, as a result of undergoing metacognitive confidence manipulation in a health-related context, searchers' information selection and evaluation should change as a function of attitude confidence. Highly confident searchers should believe in the validity of their prior attitudes. In contrast, less confident searchers should lose confidence in the validity of their attitudes and thus should look for more attitude inconsistent information and evaluate attitude inconsistent information more favorably.

Source credibility in online information search

Social tagging platforms and other Web 2.0 environments like social networks are often characterized by the absence of professional gatekeepers who critically filter and select high quality content (Rowland, 2011). Therefore, it lies in the hands of the information searcher to critically evaluate sources of information themselves. In the following section, we will review the findings regarding whether people succeed in considering source credibility adequately, both in the online health-related context, and in online content aggregation platforms, such as social tagging environments.

A recent meta-analysis has shown that in the health context, it makes a difference whether people search for information on general websites or on sites that deliver user-generated content (Ma & Atkin, 2017). For websites, expertise was the decisive factor in credibility evaluations. But for user-generated content, homophily was decisive in credibility evaluations, so that searchers perceived laypersons who created content to have more credibility when their demographics were similar (Ma & Atkin, 2017).

Also in the health-related context, but using web search engines, only some people succeeded in considering source credibility (Kammerer, Bråten, Gerjets, & Strømsø, 2013). Among laypersons who searched for two competing treatments of Bechterew's disease, those participants who viewed the web as a highly (vs. less) reliable, accurate source of knowledge failed to verbally reflect on the credibility of the information source. Therefore, they visually inspected URLs for shorter periods, and they were less likely to select results at the bottom of the search engine results (thereby failing to consider differences between search results; (Kammerer et al., 2013)). In a follow-up study, those participants who viewed the web as a highly (vs. less) reliable source of knowledge, spent less time on pages with objective information, and were less likely to base their treatment recommendation and evaluation on

objective pages (Kammerer, Amann, & Gerjets, 2015). In this case, a short intervention improved their source evaluation behavior.

With respect to the visual presentation of search results, a study found that when search results were presented in a gridlike interface, similar to tag clouds, source evaluation became more important than when the presentation was made in list format, where also the influence of list position decreased while participants selected more trustworthy search results (Kammerer & Gerjets, 2014). So, there is evidence that people succeed in critically evaluating information sources (Winter & Krämer, 2012, 2014), but this depended on personal characteristics as well as the visual characteristics of the navigation interface (Kammerer et al., 2013).

Regarding the source evaluation of health-related information in social tagging systems, we found a single study which investigated whether people select tags that indicate the credibility of the tag-related source (O'Grady et al., 2012). When participants who had been diagnosed with diabetes browsed a tag cloud with information on their condition, only one-third used at least one credibility-related tag. When explicitly asked to search for highly credible content, about 90% used at least one credibility-related tag, and only 23% exclusively used credibility-related tags. So, even when explicitly asked to evaluate source credibility, participants only partially evaluated credibility. In contrast to this, the study presented here does not focus on tags indicating credibility, but on the credibility of the community that provides the tags.

It is our aim to assess whether searchers adequately consider the source credibility of the tagging community. We varied the expertise of the community providing tags and expected that high expertise would relate to a perception that the community was more credible as a source. Blog post authors and cited sources within blog posts were constant and tested in a prior study. In line with the accuracy motivation theory for confirmation bias (Hart et al., 2009), we expected that perception of a community as highly credible should increase openness to selecting tags and blog posts, and increase openness towards persuasive messages from blog posts.

Information foraging in tagging environments

In its description of the search for information, the information foraging theory is based on an analogy to food-foraging strategies in behavioral ecology (Pirolli & Card, 1999). One of the basic claims is that cognitive systems aim to maximize gains of valuable information in relation to search cost, analogous to animals' food search. The information searcher (predator)

browses different web sites (different patches in the environment) by estimating which links lead to highly valuable information (high yield patches, with lots of high caloric prey).

There are two fundamental activities in information foraging: between-patch exploration, describing the breadth of information search, and within-patch exploitation describing the depth of search and consumption of information. A searcher will switch between patches if the next patch promises higher information value than the current patch. In contrast to between-patch exploration, within-patch exploitation is the activity of processing information within a website. To operationalize navigation between and within information sources, we use a tagging environment where searchers navigate between patches (separate sites) via social tags. The tags lead searchers to patches that list several blog post headlines on which they can click to read the full blog post.

When users navigate between patches via tags (Figure 2.1), they need to estimate which navigation path will lead them to a patch with valuable information, or in other words, searchers follow the information scent. Information scent in social tagging systems depends on a user's individual semantic associations as well as on the tag popularity of the tagging community (collective associations; (Cress & Held, 2013; Cress et al., 2013; Held, Kimmerle, & Cress, 2012)). Individual associations are activated in the searcher's memory and will guide her or him to select tags that match with her or his associations. Collective associations in the form of popular tags will guide the searcher to select more popular tags which are displayed with a larger font size. Individual and collective associations are the fundamental building blocks of the extended information scent model (Cress et al., 2013), and literature on tagging presents individual and collective associations as the main factors that determine navigation between patches in social tagging systems (Cress & Held, 2013; Cress et al., 2013; Held et al., 2012).
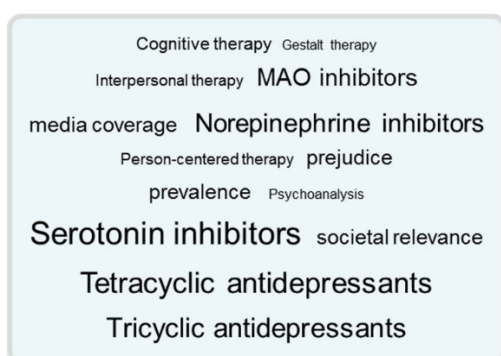

Figure 2.1 Tag cloud used in this study.

Extending these findings, we raise the question as to whether, besides individual semantic associations, prior attitudes and thus confirmation bias also influence between-patch navigation in social tagging systems. A preliminary study showed that prior attitudes guided between-patch navigation (tag selection) as well as subsequent within-patch navigation (selection of blog posts on separate sites) (Schweiger et al., 2014). In terms of information foraging, this result suggests that information scent depends on prior attitudes. So, regarding between-patch navigation, tags that were in line with prior attitudes were selected more often. Moreover, within-patch navigation, or the selection of tag-related blog posts, was related to prior attitudes. As an extension of these preliminary findings, we test in this study if the influence of prior attitudes on navigation can be enhanced or weakened by attitude confidence. We also test whether high credibility of the community circumvents the influence of the participants' prior attitudes and increases their selection rate between and within patches independently of prior attitudes.

We align our expectations with existing literature on social tagging systems, where individual and collective factors independently influenced navigation and processing of information (Cress & Held, 2013; Cress et al., 2013; Schweiger et al., 2014). We assume that the searchers' expected gain of desired information and therefore the information scent would depend on prior attitudes. Expected gain would be high when information is attitude consistent, and searchers would switch to more attitude-consistent patches by selecting respective tags. For blog posts, higher gain would be experienced from reading blog posts within attitude-consistent patches, and subsequent evaluation of blog posts should be more favourable when posts are attitude consistent. These effects should be stronger when attitude confidence is high compared to low. When source credibility is high, the expected gain should be higher for attitude-consistent as well as attitude- inconsistent patches, so overall, searchers should navigate more between and within patches.

This study's domain

To structure an investigation of a health-related information search scenario using the treatment of depression disorders, we made use of participants' prior attitudes. In many countries, psychotherapy is thought to be more effective than antidepressants (Angermeyer, Breier, Dietrich, Kenzine, & Matschinger, 2005; Jorm et al., 2005), although recent meta-analyses show equal efficacy, on a moderate level, for both treatments (Cuijpers et al., 2013; De Maat, Dekker, Schoevers, & De Jonghe, 2006). In Germany, for example, we previously

found that antidepressants were considered to be moderately effective, whereas psychotherapy was even considered to be moderately to highly effective (Schweiger et al., 2014). However, it should be noted that this issue is probably not highly controversial, as a combination of both therapies was recommended by a significant proportion of participants in our previous study (Schweiger et al., 2014).

To measure prior attitudes regarding the domain, and as part of the attitude confidence manipulation, we asked participants to provide arguments for and against psychotherapy and antidepressants. We propose that when a participant offers more arguments in favor of therapy and fewer against therapy, the overall evaluation is then positive, reflecting a positive attitude (Gawronski & Bodenhausen, 2011). So, to measure prior attitudes, we built an index with the sum of pro and contra arguments for psychotherapy and antidepressants. This should ensure that treatment relevant attitudes are measured in a broad and personal sense, without topical constraints.

Finally, for the health-related domain of treatment for depression, we assume that individuals are accuracy-motivated when searching for information. A recent review has highlighted two different motivations that fuel confirmation bias (Hart et al., 2009). First, accuracy-motivated individuals seek to select objectively correct information. That is, when accuracy motivation is high, individuals are more open to attitude-inconsistent information. This is because accuracy-motivated searchers are more influenced by cues that indicate objective correctness or perceived validity of information (Hart et al., 2009). In contrast to this, defense motivation leads to selecting information that confirms prior attitudes and thus defends one's self concept. For this study, our hypotheses are in line with the accuracy motivation theory, as we expect that for health-related searches, obtaining accurate information is more important to the searcher than protecting one's self-concept.

Overview and hypotheses

We used tags in a social tag cloud where antidepressant tags were larger than tags that related to psychotherapy. Both kinds of treatment are discussed in the context of health. Since there is a preference for psychotherapy over antidepressants in the population (Angermeyer, van der Auwera, Carta, & Schomerus, 2017; Van Der Auwera, Schomerus, Baumeister, Matschinger, & Angermeyer, 2017), we aimed to counter this preference to increase the likelihood of equal amounts of clicking on both treatments (Schweiger et al., 2014). Before browsing the tag cloud, however, participants indicated their prior attitudes by providing arguments and efficacy ratings for antidepressants and psychotherapy. We began manipulation of the

participants' attitude confidence by having them recall experiences where they either felt confident or unconfident about their thoughts (see (Petty et al., 2002)). Subsequently, we presented a tag cloud which came from one of two communities that differed in alleged expertise in the domain. Via tag clouds, participants navigated among patches, and they were able to select multiple blog posts within each patch. All blog posts highlighted the efficacy of either psychotherapy or antidepressants in a positive way. We did not present any blog posts that compared both types of therapy, or any blog posts that presented studies with negative findings. To measure the searchers' content evaluation as a consequence of navigation, we had them rate treatment efficacy again, after navigation. Finally, we conducted a recognition test for the blog posts participants had read.

We expected the following:

First, we expected our replication of the confidence manipulation (Petty et al., 2002) to be successful (H1a). When attitude confidence is high, people will select more attitude-consistent tags (H1b) and blog posts (H1c), compared to when attitude confidence is lower. That is, to the degree that people favour or disfavour treatments, confidence should moderate the influence of prior attitudes on attitude-consistent selection of tags and blog posts. Content evaluation in terms of treatment efficacy ratings after navigation will change accordingly (H1d).

Second, participants will recognize credibility of the tagging community (H2a). If a highly credible community (compared to a less credible community) provides content, tags (H2b) and blog posts (H2c) of this community should be selected more often. In the same vein, content gathered by a highly credible community will increase change of treatment efficacy ratings more compared to a less credible community (H2d).

A possible interaction effect between attitude confidence and source credibility remains an open question, since to the best of our knowledge, there are no background studies that have manipulated attitude confidence in combination with source credibility in online or offline information selection tasks. A second open question to be explored is whether knowledge acquisition will also be affected.

Results

Recruitment and participants

Participants were contacted via a mailing list. As an incentive, participants were offered the opportunity to take part in a lottery with 50 Euro Amazon gift certificates. Ethical approval

was provided by the Ethical Committee of the Knowledge Media Research Center (LEK 2014/006). 138 participants out of a total of 331 persons who accessed the survey finished it. Five participants retracted their data. We dropped participants from the analysis who completed the study twice (n=2), who did not provide prior attitudes(n=2) or did not click on tags (n=2). We also excluded outliers who scored in the attitude confidence manipulation check with a median absolute deviation greater than three (n=2; (Leys, Ley, Klein, Bernard, & Licata, 2013)). 125 participants were included in the analysis. In the final sample, the age ranged from 19 to 64 years old (M = 24.74, SD = 6.27), and 90 participants were female (72%).

Materials

For browsing through treatments, we provided a tagging environment that comprised two main sections (see Figure 2.2). The right part of the screen displayed 14 tags. Five tags represented psychotherapy and five tags represented antidepressants. Four tags were not relevant with respect to treatment (prejudice, media coverage, societal relevance, prevalence). Antidepressant tags were larger than psychotherapy tags, as can be seen in Figure 2.1 and Figure 2.2.

On the left part of the screen, blog posts were presented for each tag (Figure 2.2). Three blog posts were related to each tag. The content of the blog posts for antidepressants (15 posts) and psychotherapy (15 posts) was held constant with respect to number of arguments and length (mean 76.8 words, SD 6.1). Each post described a common symptom of depression disorders and scientific studies on the efficacy of the respective treatment. In a pilot study, we had made certain that the blog posts were equal in readability and credibility, and that there was no difference in the persuasiveness or quality of all of the arguments within the pairs of blog posts about antidepressants and psychotherapy. Initially, only the headline and the first sentence of each blog post were presented. In order to read the full blog post, participants had to click on the first sentence to expand the blog post.
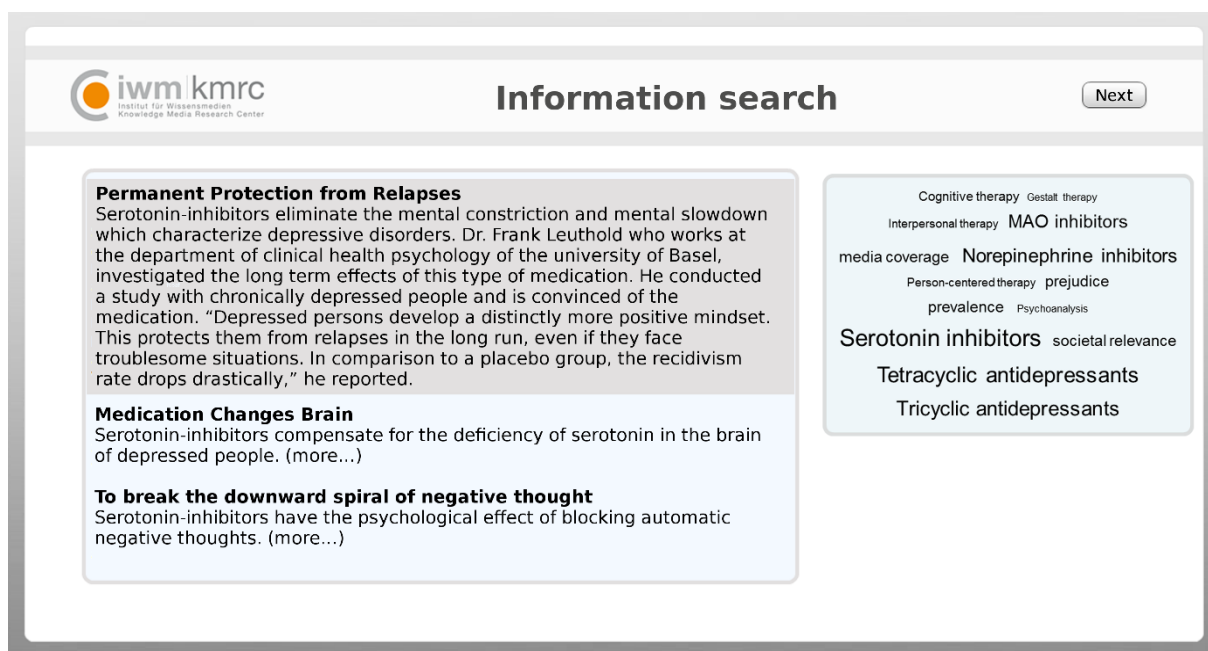
Figure 2.2 Social tagging environment used in the study.

Design and procedure

The study comprised a 2 (attitude confidence: high, low) x 2 (source credibility: high, low) between-subjects design. Participants were randomly assigned to one of the four experimental conditions. On the first pages of the online survey, we welcomed participants, provided a brief review of the procedure of the study, asked participants to provide consent by clicking on the "continue" button, and requested basic demographic data. Then we asked them to state pro and contra arguments regarding antidepressants and psychotherapy (see box pretest in Figure 2.3). We also asked them to rate the efficacy of antidepressants and psychotherapy. After this, we manipulated *attitude confidence.* For an alleged unrelated study, we asked them to recall situations in which they had felt either confident or unconfident about their own thoughts (Petty et al., 2002). After they had recalled such situations, we asked participants to think back to their arguments for and against psychotherapy and antidepressants. They rated how confident they felt about the arguments they had provided at the beginning of the study. This rating served as a manipulation check for attitude confidence.
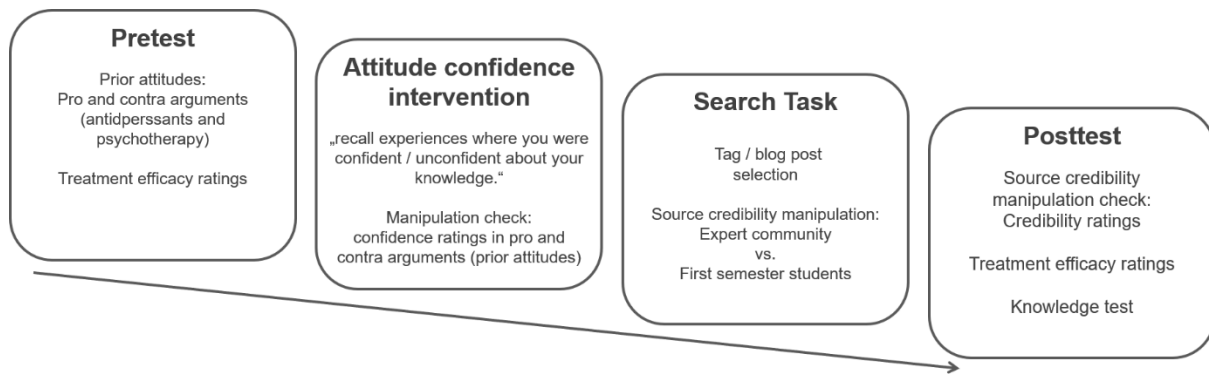
Figure 2.3 Experimental procedure.

In the search task, participants browsed a tag cloud that visualized treatments for depression. Tags related to antidepressants were larger than tags related to psychotherapy (Figure 2.2). We asked participants to search for information about the efficacy of treatments for depression, to provide advice for a hypothetical friend. The treatments were either psychotherapy or antidepressants. The tagging environment appeared for at least five minutes, after which participants could freely decide to browse further or to stop browsing tags and tag-related blog posts. After the navigation task, participants rated source credibility, which provided a manipulation check. Again, participants were asked to rate the efficacy of antidepressants and psychotherapy. Finally, participants filled out a retention test on the blog posts they had read. At the end, participants were debriefed and given the opportunity to leave comments on the study.

Independent variables

Prior attitudes

As a measure of prior attitudes, we built an index of the sum of *pro* (positive value) and *contra* (negative value) *arguments* separately for psychotherapy and antidepressants. Each argument was counted by one rater, who coded each proposition for and against both treatments. (for examples, against antidepressants: "antidepressants are addictive"; for psychotherapy: "it helps when someone listens to your problems"). We validated the arguments against the treatment efficacy rating scale prior to navigation and found a correlation between the arguments and efficacy ratings for antidepressants ($r = .18$, $p < .05$), but not for psychotherapy ($r = .06$, ns.).

Attitude confidence

To manipulate attitude confidence, we adapted the experimental procedure used by Petty and colleagues (2002) to our study. After participants had provided their arguments for and against psychotherapy and antidepressants, we asked participants to recall situations in the past where they had felt either confident or unconfident about their own thoughts. The task was presented for a minimum of 5 minutes in which participants were asked to enter the situations in 5 input boxes.

Source credibility

We presented banners which implied that either college students (low expertise) or domain experts (high expertise) had collected and tagged resources. At the top of the page in the tagging environment (above the visible space in Figure 2.2), either a banner of an online student forum (Figure 2.4), or a banner of alleged federal expert association (Figure 2.5) was displayed.
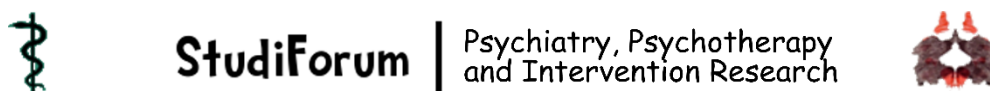
Figure 2.4 Translated version of the banner for the low source credibility group.

Figure 2.5 Translated version of the banner for the high source credibility group.

Dependent variables

Selection of tags and blog posts

As indicators of participants' search behavior, we assessed their clicks on tags and blog posts. The use of both measures as dependent variables allowed for an analysis of participants' search behavior at various levels of elaboration: Whereas clicks on tags might indicate a general interest in the sources linked to this tag, they do not allow for any elaborate reception (i.e. reading) of related content. Therefore, we also analyzed the frequency of clicks on blog posts, which indicated participants' interest in the posts' specific content. We only counted

clicks on those blog posts which were displayed for at least five seconds, which suggested that participants had spent more time reading those posts more thoroughly.

Treatment efficacy ratings

We calculated a score with subjective treatment efficacy ratings for antidepressants and psychotherapy. Participants rated the degree to which they agreed with the statements about the efficacy of both treatments, on a scale ranging from 1 (completely disagree) to 7 (completely agree). As an example: "There is scientific evidence that clearly demonstrates the efficacy of psychotherapy/antidepressants." Efficacy ratings were assessed prior to navigation (see S2 Table; antidepressants Cronbach's α = .79, psychotherapy Cronbach's α = .88), and after navigation (antidepressants Cronbach's α = .88, psychotherapy Cronbach's α = .91).

Knowledge acquisition score

After navigation, participants filled out a multiple-choice test (1 target, 2 foils for each browsed blog post). For targets and foils, participants responded on a scale ranging from 1 ("completely wrong") to 5 ("completely correct"). For example, the test questions for the blog post in Figure 2.2 is as follows: Target: "People who are treated with serotonin inhibitors show lowered risk of relapse in comparison to a control group." Foil: "Serotonin inhibitors show a high tolerability." We only analysed targets and foils that were related to a blog post that participants had clicked on. We recoded hits as 1 point and deleted items unrelated to blog posts that had been read, whereas the middle category (3: "I don't know") was coded as zero points. We calculated the total score separately for psychotherapy and antidepressants.

Results

All of the following analyses were conducted with the statistical software R (v 3.5.0; (R Core Team, 2018)), including the packages "glmmADMB" (v 0.8.3.3; (Skaug et al., 2018)), "tidyverse" (v 1.2.1; (Wickham, 2017)). A full reproducible analysis script and data are available in the supplementary materials (see S1 Statistical Analyses).

Prior attitudes

As expected (Angermeyer et al., 2017; Van Der Auwera et al., 2017), prior to navigation, participants evaluated psychotherapy more positively than antidepressants on the treatment efficacy rating scales (antidepressants: M = 4.23, SD = 0.93; psychotherapy: M = 5.69, SD = 0.76; t_{124} = 17.02, p < .001, d = 1.73). The same, but somewhat weaker tendency towards psychotherapy could be observed when analyzing the number of arguments regarding

antidepressants (pro: M = 2.06, SD = 1.21; contra: M = 2.69, SD = 1.36) and psychotherapy arguments (pro: M = 2.91, SD = 1.39; contra: M = 1.92, SD = 1.32). These results showed a moderate tendency in favor of psychotherapy compared to antidepressants (pro arguments: t_{124} = 6.70, p < .001, d = 0.65; contra arguments: t_{124} = 6.12, p < .001, d = 0.57).

We calculated an index for prior attitudes by subtracting arguments in favor of each treatment from arguments against the respective treatment, so a negative prior attitudes score reflected a negative evaluation for the respective treatment, and a positive score a positive evaluation.

Treatment efficacy ratings were measured by the comparison of pretest ratings (before navigation) and posttest ratings (after navigation). Figure 2.6 and S2 Table provide an overview of the treatment efficacy rating items and response distribution.
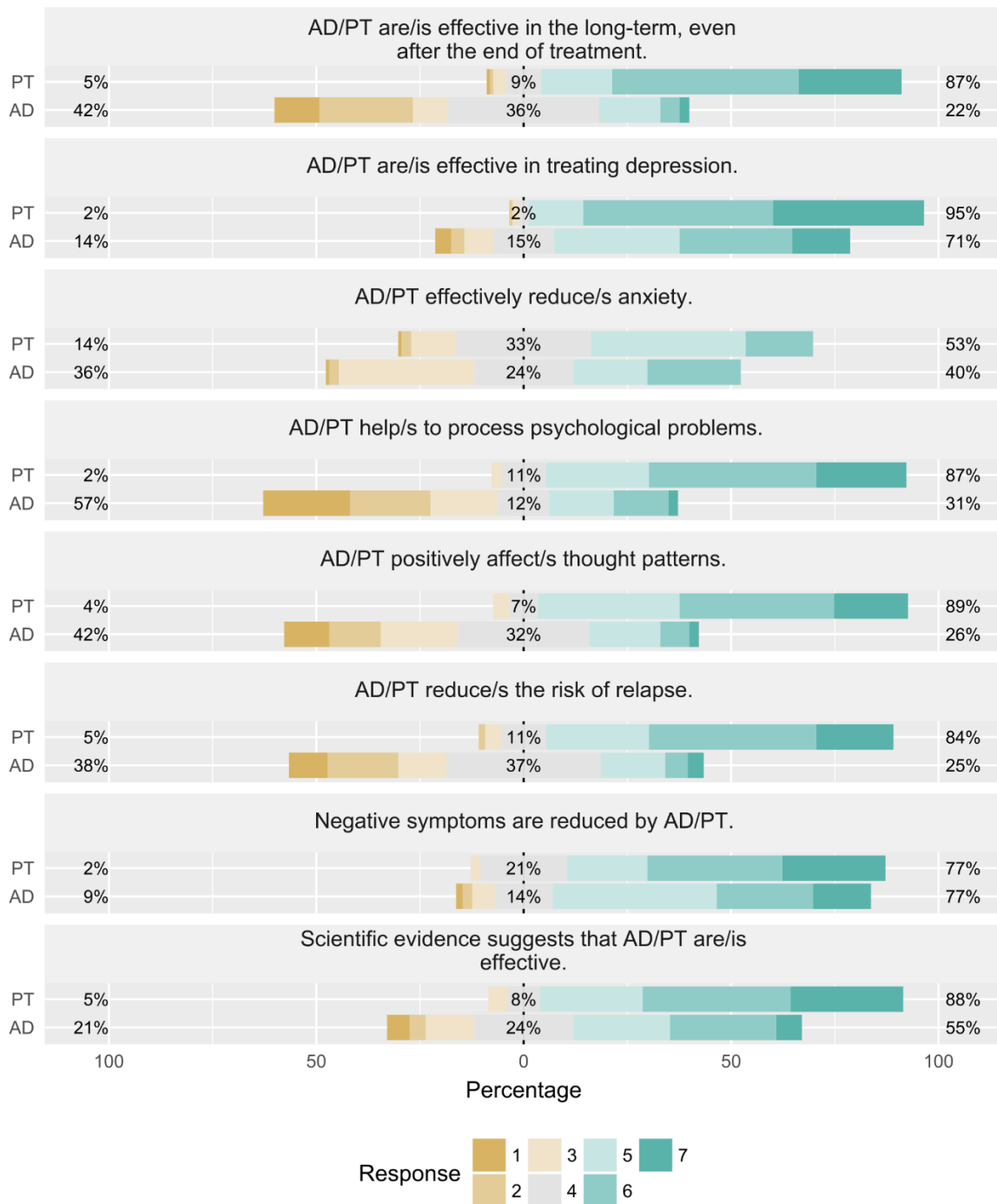
Figure 2.6 Treatment efficacy ratings for each item before navigation.

Manipulation checks

Attitude confidence

After participants recalled situations in which they had felt either confident or unconfident, they rated confidence in their own arguments for or against a depression treatment on a scale

ranging from 1 (not at all) to 7 (highly). They also rated the degree to which the following qualities described their arguments: obvious, dubious, justified, credible, factual, well-founded, persuasive and objective (Cronbach's α = .86). Those participants who during the manipulation had recalled situations in which they had felt confident in their own thoughts were also confident about their arguments about treatment (M = 5.10, SD = 0.75). In contrast, participants were less confident in their arguments after having recalled situations in which they had been unconfident (M = 4.81, SD = 0.74; t_{123} = 2.16, p = .03, d = 0.39). Our attitude confidence manipulation was thus shown to be effective and hypothesis H1a was supported.

Source credibility

After navigation, participants rated the credibility of the source of information on a scale ranging from 1 (not at all) to 7 (highly). They rated the degree to which the following qualities described the group who had compiled the blog posts: informed, conscientious, trustworthy, credible, and competent (Cronbach's α = .89). Participants rated the source as more credible when the group consisted of alleged experts (M = 5.23, SD = 0.95), compared to students in their first semester who were rated as less credible (M = 4.88, SD = 0.87; t123 = 2.19, p = .04, d = 0.25). Thus, participants were able to recognize high and low expertise of the tagging community, and hypothesis H2a was supported.

Navigation analyses

We analyzed navigation behavior with generalized linear mixed models (Bolker et al., 2009), using the "glmmADMB" (Skaug et al., 2018) package, to model the dependent variables counts of clicks on tags [H1b], and blog posts [H1c]). We first compared the empirical versus theoretical quantiles of tag and blog post counts visually, which suggested that the count data followed the Poisson distribution, so we marked? the models with a Poisson distribution with a log link.

As fixed factors, we included the independent variables credibility and confidence (both 0 = low, 1 = high) and the prior attitudes index. For the model predicting blog post counts, we included the number of clicks on tags as a covariate. Continuous predictors were standardized. To account for the within-subjects measure treatment (Magezi, 2015), we included random by-participant and by-treatment intercepts.

For both models, visual inspection of the residual plots did not reveal any obvious deviations from homoscedasticity or normality. The model comparison strategy was to include all main (fixed) effects and random intercepts, and to compare this to the model including only the

fixed intercept and the random intercepts (Tables 2.1-2.2, Step 1). Next, we included the hypothesized interaction term between attitude confidence and prior attitudes and compared this to the main effect model from Step 1 (Tables 2.1-2.2, Step 2). For exploratory purposes, we separately tested whether the model fit would further be improved by including the other two-way interaction terms (confidence x credibility, prior attitudes x credibility), and the three-way interaction term (prior attitudes x confidence x credibility). Significance of effects was obtained by means of a likelihood ratio test comparing the full model with the effect in question to the model without the effect in question (Bolker et al., 2009).

Attitude confidence and navigation

Tag selection

We hypothesized that when attitude confidence is high (vs. low), people should select more attitude-consistent tags (H1b), depending on their prior attitudes. Results showed that there was no relationship between confidence and tag selection when including only main effects. It was shown that this model did not fit the data any better than the intercept-only model (Table 2.1, Step 1), and including the interaction term between confidence and prior attitudes did not improve the model fit either (Table 2.1, Step 2). Additionally, we checked for possible interaction effects among all predictors. Including the 2-way interaction between confidence and credibility (Table 2.1, Step 3) and the 3-way interaction (Table 2.1, Step 5) did not improve the model fit compared to step 1, whereas the model fit with main effects and the interaction between prior attitudes and confidence did improve mdel fit compared to step 1 (Table 2.1, Step 2). However, including the 2-way interaction between prior attitudes and source credibility did show a significant improvement of model fit compared to Step 2 (Table 2.1, Step 4), which will be discussed in the section below (Source credibility and navigation). Therefore, the hypothesis that attitude confidence would moderate the influence of prior attitudes on tag selection was not supported (H1b).

Table 2.1 Tag selection.

Step 1	b	SE	p
Intercept	.95	.10	<.001
Prior attitudes	.03	.04	.52
Attitude confidence	-.02	.10	.83
Source credibility	.24	.10	.02
($\chi^2(3) = 5.48$, p = .14, R^2 = .03)			
Step 2			
Prior attitudes X attitude confidence	.01	.08	.87
($\chi^2(1) = 0.25$, p = .87, ΔR^2 = .00)			
Step 3			
Attitude confidence X source credibility	.14	.20	.51
($\chi^2(1) = 0.44$, p = .51, ΔR^2 = .03)			
Step 4			
Prior attitudes X source credibility	-.21	.08	.01
($\chi^2(1) = 6.47$, p = .01, ΔR^2 = .00)			
Step 5			
Prior attitudes X attitude confidence X source credibility	.02	.17	.88
($\chi^2(1) = 0.02$, p = .88, ΔR^2 = .03)			

b, Beta coefficients with standardized continuous predictors; SE, Standard Error; R^2, Nagelkerke R^2 for fixed effects only.

Blog post selection

We hypothesized that when attitude confidence is high (vs. low), people should select more attitude consistent blog posts (H1c), depending on their prior attitudes. For blog post selection (H1c, H2c), we used the same procedure for model comparison and the same independent variables as in the analysis of tag selection but additionally included the number of selected tags as a covariate. As Table 2.2 shows, tag selection predicted blog post selection (Step 1), and we found the hypothesized interaction between prior attitudes and attitude confidence (Step 2).

Table 2.2 Blog post selection.

Step 1	b	SE	p
Intercept	.29	.14	.04
Tags selected	.53	.05	<.001
Prior attitudes	.05	.06	.36
Attitude confidence	.05	.13	.71
Source credibility	.15	.13	.25
($\chi^2(4) = 128.24$, p = <.001, R^2 = .33)			
Step 2			
Prior attitudes X attitude confidence	.17	.08	.03
($\chi^2(1) = 4.57$, p = .03, ΔR^2 = 0.01)			
Step 3			
Attitude confidence X source credibility	-.47	.25	.07
($\chi^2(1) = 3.41$, p = .407, ΔR^2 = .02)			
Step 4			
Prior attitudes X source credibility	-.05	.08	.54
($\chi^2(1) = 0.37$, p = .53, ΔR^2 = .00)			
Step 5			
Prior attitudes X attitude confidence X source credibility	.07	.17	.68
($\chi^2(1) = 0.17$, p = .68, ΔR^2 = .02)			

b, Beta coefficients with standardized continuous predictors; SE, Standard Error; R^2, Nagelkerke R^2 for fixed effects only.

To disentangle the interaction between prior attitudes an attitude confidence we did a subgroup analysis separately for the high (b = .13, SE = .06, p = .04) and low (b = -.02, SE = .05, p = .69) confidence groups, and plotted the predicted number of selected tags depending on prior attitudes (see Figure 2.7, middle panel). This finding supported our expectations that prior attitudes are only associated with the selection of blog posts when confidence is high (H1c).
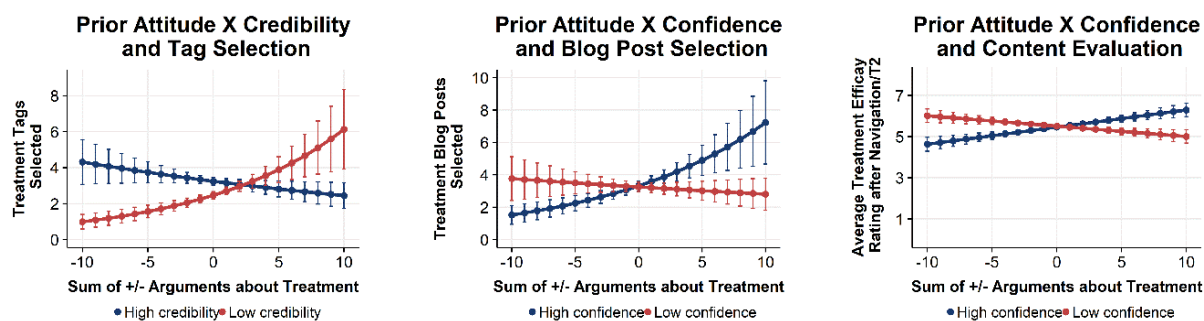
Figure 2.7 Overview of interaction effects.

Source credibility and navigation

Tag selection

We expected that participants would select more tags when source credibility was high, independent of their prior attitudes (H2b). As described above, the main effect model including the effect of source credibility did not fit the data. However, the model including the interaction between prior attitudes and source credibility did fit. To disentangle this interaction effect, we did a subgroup analysis for the high ($b = -.05$, $SE = .05$, $p = .32$) and low credibility ($b = .16$, $SE = .06$, $p = .02$ groups (see Figure 2.7, left panel), and plotted the predicted number of selected tags against prior attitudes. The result shows that only only the low credibility group selected tags which indicated attitude consistency with prior attitudes.

Blog post selection

We also expected that if the source credibility of the tagging community is high (vs. low), blog posts would be selected more often, independently of prior attitudes (H2c). Controlling for tag selection, however, we did not find any effect of source credibility on blog post selection (Table 2.3).

Content evaluation analyses

To investigate evaluation of content, we ran linear mixed models (Bolker et al., 2009), using the "lme4" (Bates, Maechler, Bolker, & Walker, 2015) package, with the standardized, dependent variable treatment efficacy ratings after navigation. We first compared the empirical versus theoretical quantiles of tag and blog post counts visually, which suggested that the data followed a normal distribution.

As fixed factors, we included the independent variables source credibility and attitude confidence (both 0 = low, 1 = high) and the prior attitudes index. As covariates we included the treatment efficacy ratings before navigation and the number of clicks on blog posts. To control for the within-subjects factor treatment type, we included random intercepts for each participant and both treatment types. The continuous predictors tags and prior attitudes as well as the dependent variable treatment efficacy ratings were standardized (Aiken & West, 1991).

For both models, visual inspection of the residual plots did not reveal any obvious deviations from homoscedasticity or normality. The model comparison strategy was to first include all main (fixed) effects and random intercepts (Table 2.1, Step 1), and to compare this to the model including only the fixed intercept and the random intercepts. Next, we included the fixed, hypothesized interaction term between attitude confidence and prior attitudes, and compared this to the main effect model from the first phase (Table 2.1, Step 2). For exploratory purposes, we separately tested whether the model fit would further be improved by including the other two-way, fixed interaction terms (confidence x credibility, prior attitudes x credibility), and the three-way, fixed interaction term (prior attitudes x confidence x credibility). Significance of effects was obtained by means of a likelihood ratio test comparing the full model with the effect in question to the model without the effect in question (Bolker et al., 2009).

Attitude confidence and content evaluation

Treatment efficacy ratings

With H1d we expected that participants would keep their prior attitudes only when attitude confidence was high. When attitude confidence is low, participants should become more open to content of attitude inconsistent blog posts. The model including the interaction between prior attitudes and confidence did fit the data (Table 2.1, step 2). To disentangle this interaction effect, we did a subgroup analysis for the high ($b = .17$, $SE = .06$, $p < .01$) and low ($b = -.11$, $SE = .05$, $p = .05$) confidence groups, and plotted the predicted treatment ratings against prior attitudes. In support of H1d, the results showed that after controlling for blog post selection, only with high confidence were prior attitudes and attitude after navigation associated (Figure 2.7, right panel).

Table 2.3 Treatment efficacy ratings after navigation.

Step 1	b	SE	p
Intercept	.03	.07	
Blog posts selected	.22	.04	<.001
Treatment efficacy ratings before navigation	.72	.04	<.001
Prior attitudes	.02	.04	.57
Attitude confidence	-.02	.09	.78
Source credibility	-.03	.09	.72
($\chi^2(4) = 190.92$, p < .001, $R^2 = .64$)			
Step 2			
Prior attitudes X attitude confidence	.22	.07	.002
($\chi^2(1) = 9.45$, p = .002, $\Delta R^2 = .01$)			
Step 3			
Attitude confidence X source credibility	.12	.17	.47
($\chi^2(1) = 0.52$, p = .47, $\Delta R^2 = .01$)			
Step 4			
Prior attitudes X source credibility	.04	.07	.57
($\chi^2(1) = 0.32$, p = .57, $\Delta R^2 = .00$)			
Step 5			
Prior attitudes X attitude confidence X source credibility	-.01	.14	.93
($\chi^2(1) < 0.01$, p = .99, $\Delta R^2 = .01$)			

b, standardized Beta coefficients; SE, Standard Error; R^2, fixed effects only.

Source credibility and content evaluation

We expected that a high (vs. lower) source credibility would increase change of treatment efficacy ratings for attitude-consistent as well as attitude-inconsistent information (H2d). As shown in Table 2.3, there was no effect of source credibility on content evaluation measured by treatment efficacy ratings after navigation.

Knowledge acquisition analyses

Additionally, we explored knowledge acquisition with linear regression. The goal was to evaluate whether knowledge acquisition would take place in a way consistent with prior attitudes. The criterion *correct responses* was entered as a dependent variable, and blog post selection, prior attitudes, attitude confidence, and source credibility were entered as predictors. As a random factor we included a by-participant intercept, whereas predictors were centered. The main effect model fit the data ($\chi^2(4) = 98.99$, p < .001). The number of selected antidepressant blog posts predicted the knowledge acquisition score (standardized b = .68, SE = .06, p < .001). However, the remaining predictors or including their interactions did not predict learning (all ps > .08). Our experimental manipulations of attitude confidence and source credibility were thus shown not to lead to more or less attitude-consistent knowledge acquisition.

Overview and path analysis

To provide an overview on how attitude confidence and source credibility affect different stages of the navigation process, we conducted a confirmatory path analysis (B Shipley, 2000), using the R software with the packages "lme4" (v 1.1-17; (Bates et al., 2015)), and "piecewiseSEM" (v 2.0.2; (Lefcheck, 2016)). We used the piecewise structural equation modelling approach, as it allows for including models for Poisson distributed count data (B Shipley, 2000). The approach makes it possible to evaluate multiple causal hypotheses simultaneously within a single network of connected nodes (Lefcheck, 2016). Therefore, we evaluated separate component models in the form of regression equations, as reported above (Table 2.1, Step 4; Table 2.2, Step 2; Table 2.3, Step 2). However, tests of directed separation showed that 2 paths of tag selection on treatment efficacy ratings T2, and treatment efficacy ratings T1 for blog post selection were missing, which were included in the final model. The final model did fit the data, and no significant paths were omitted (C = 1.50, P = .83; (Bill Shipley, 2013); see Figure 2.8).
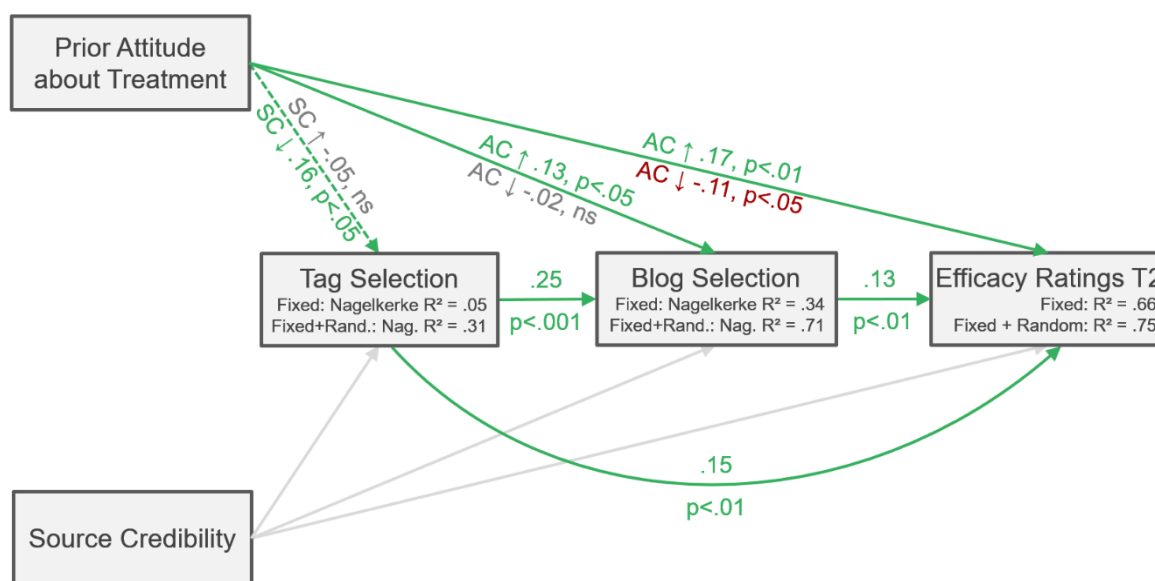
Figure 2.8 Path analysis.

Note. AC = Attitude Confidence, SC = Source Credibility; Solid arrows = hypothesized effect; Dashed arrows = not hypothesized effect; Grey Arrows = not significant effects. Effects of the covariate treatment efficacy ratings T1 on blog post selection (b = .11, p < .01) and treatment efficacy ratings T2 (b = .71, p < .001) not shown. Subgroup analyses were conducted to provide separate group parameters when there was a significant association with prior attitudes.

Discussion

In this study, our aim was to investigate how social tag clouds would influence a pre-existing confirmation bias. Searchers would be expected to be more open to social tags and blog posts that challenged their prior attitudes if they had little confidence in their prior attitudes. We also aimed to test whether searchers would correctly evaluate source credibility, and if in turn they would explore more content and be more open to content when they perceived the credibility of the tagging community to be high. We found that searchers exhibited more favourable attitudes towards psychotherapy compared to antidepressants and that searchers correctly recognized source credibility in terms of high and low community expertise. Results also showed that our replication of the attitude-confidence manipulation was successful,

making it possible to rely on its effects. In the navigation process, source credibility had influence only on the early stage of tag selection, whereas confidence had consequences later, when participants were selecting blog posts and subsequently evaluating content.

We acknowledged that people tend to confirm their prior attitudes. However, specifically when prior attitudes were pronounced, and attitude confidence was high (vs. low), confirmation bias increased in blog post selection, but not in tag selection. Moreover, when attitude confidence was high, prior attitudes were positively associated with evaluation of attitude-consistent content. When attitude confidence was low, the effect was even reversed, that is, attitude-inconsistent content was evaluated more favourably.

We also found that when credibility was high, the influence of prior attitudes on tag and blog post selection was eliminated. By contrast, when source credibility was low, prior attitudes guided selection of tags and blog posts, showing positive association with the selection of attitude-consistent tags.

Attitude confidence

In general, people tend to be overly confident about their own knowledge (Dunning et al., 2004; Moore & Healy, 2008; Zell & Krizan, 2014), but particularly in the context of health-related information search, confidence in one's prior attitudes may vary. Having chosen a health-related domain, we therefore used manipulation that had the potential to counter overconfidence. In this scenario, we expected that the metacognitive aspect of attitude confidence, that is, confidence in the validity of one's own arguments, would have consequences for the search and cognitive processing of health-related information.

We had expected that the influence of prior attitudes would depend on confidence in information search and content evaluation. When they possessed high confidence, searchers would be expected to select and favourably evaluate attitude-consistent information. We found this influence of confidence only for the evaluative stage of information search, or in terms of information foraging (Pirolli & Card, 1999), the activity of within-patch exploitation. This included the selection of blog posts as well as the subsequent content evaluation. When confidence was high, participants showed a tendency to select and favourably evaluate attitude-consistent information, but when confidence was low, there was no influence of prior attitudes on navigation.

In the light of the self-validation hypothesis, from which the original attitude confidence manipulation was derived (Petty et al., 2002), the findings seem to support the

assumption that searchers in the tagging environment were motivated primarily to look for accuracy. The self-validation hypothesis states that only thoughts that are perceived as valid determine attitudes and related processing of information. If confidence is high, one's own thoughts or arguments are perceived as valid, and consequently one's own arguments should have a high impact on navigation and content evaluation, in contrast to when confidence is low. In support of this theory, we found an influence of high attitude confidence on the evaluative stage of search, in which content was selected and evaluated. More surprisingly, but in line with this train of thought, when attitude confidence was low, the effect of prior attitudes on content evaluation was even reversed, and participants rated attitude-inconsistent content more favourably.

Finally, to the best of our knowledge, this is the first study that has directly replicated the metacognitive confidence manipulation (Petty et al., 2002). It is therefore worthwhile to compare the effects of the manipulation in both studies. Interestingly, the original study showed a large effect of the manipulation. Our replication, however, showed a small to moderate effect on the same scale. Moreover, both the high and the low attitude confidence groups were above the midpoint of the scale. This finding is not surprising, considering that most people tend to be overly confident in a wide range of domains (Kruger & Dunning, 1999; Zell & Krizan, 2014).

Evaluation of source credibility

We had expected that source credibility would influence information search and the resulting content evaluation, independently of prior attitudes. Information search consisted of two consecutive processes: First, selection of topics (tags), and second, selection of in-depth, valenced information (blog posts). The topic-oriented tag selection illustrated the between-patch activity in terms of information foraging theory (Pirolli & Card, 1999), where information foragers switch among information environments (patches). Source credibility only affected this exploratory, uncertain information selection process where the searcher needed to estimate which tag would lead to the most valuable information. We also observed an interplay of source credibility and prior attitudes. That is, prior attitudes had no influence on tag selection only when source credibility was high, and when source credibility was low, participants with increasingly positive attitudes towards a treatment selected that respective treatment tag more often.

In terms of a confirmation bias, the finding that under high source credibility prior attitudes were not associated with the selection of attitude consistent tags supports the accuracy

motivation theory (Hart et al., 2009). That is, participants valued objectively correct information, which they would more likely get from a highly credible source. In contrast to the accuracy motivation theory, the defense motivation theory would predict that prior attitudes would be negatively associated with the selection of attitude inconsistent tags when source credibility is high, since high credibility of attitude- inconsistent information would pose a threat. For the high source credibility condition, the finding is in line with our assumption that participants would be likely to be motivated by a desire for accuracy, due to the health-related context.

But for the low source credibility condition, searchers might have been acting less out of accuracy or defense motivation. If searchers were highly defense-motivated, participants would have avoided attitude-consistent information under low source credibility, as this potentially could have made them aware of their own questionable position. If searchers were highly accuracy-motivated, prior attitudes should not have affected tag selection, as these searchers would want to aim for high quality information (Hart et al., 2009). A possible explanation could be that if searchers in the low credibility source condition were low in both accuracy as well as defense motivation, they might simply have been guided by their confirmation bias.

Finally, participants successfully identified tagging communities as more credible when high (vs. low) community expertise was indicated by banners. This is an interesting finding, since for platforms offering user-generated content such as blog posts or forums, users now generally evaluate content as more credible if other users with similar demographics (not experts) generated content (Ma & Atkin, 2017). This shows the potential of highlighting the expertise and credibility of the community for the searcher on social tagging platforms.

Limitations and future work

In this study, tags were related to a set of different blog posts, so at first sight, tags were ambiguous with respect to related content. If one encountered the tag "antidepressant", it was not clear what the related documents were about, or if these documents supported or refuted claims about the efficacy of antidepressants. In fact, purely semantic processing of tags was forced by the selection of the tags in our tagging environment. So when prior attitudes were in line with tag selection in this study, it implies that people might have been showing that the testing strategy documented in confirmation bias literature is a good one to use (Klayman et al., 1987). But tags that are provided in real tagging systems are also evaluative, motivational and social in nature (Ames & Naaman, 2007; Heckner, Neubauer, & Wolff, 2008). We think

it would be desirable to pursue studies that investigate the evaluative nature of processing of tags, in order to include evaluative processes in tagging theories as well (e.g. (Fu, Kannampallil, Kang, & He, 2010)).

Moreover, issue involvement has been shown to be an important moderator with respect to attitude-consistent information processing (Hart et al., 2009). In this study, we did not measure involvement, which would be interesting for future studies.

As already mentioned, all of the blog posts people had access to were positive about the efficacy of either psychotherapy or antidepressants as treatments for depression. There were no blog posts that refuted the efficacy of either type of therapy. Future studies should include also neutral and negative statements about efficacy.

Outlook

We found that tag clouds may offer a way to counter confirmation bias in online health-related tagging environments. However, the extent of confirmation bias also depends on individual cognitive processes, such as confidence in one's attitudes, and on the credibility of the community providing the information. Manipulating attitude confidence offers an effective and uncomplicated intervention to reduce bias among individuals. Highlighting the credibility of a source helps to increase the impact of health-related online information and also reduce bias. With respect to confirmation bias, there is a concern that online aggregation mechanisms act as echo chambers, reinforcing people's attitudes, trapping them in filter bubbles (e.g. (Rowland, 2011)). Although some expect the effect of the bubbles to be large (Rowland, 2011), others expect the effect to be small (Hosanagar, Fleder, Lee, & Buja, 2014). With the findings of the study presented here, however, we hope to contribute to a more nuanced discussion about this topic.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression. Multiple regression: Testing and interpreting interactions*. https://doi.org/10.1080/00031305.1972.10478934

Ames, M., & Naaman, M. (2007). Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 971–980). https://doi.org/10.1145/1240624.1240772

Angermeyer, M. C., Breier, P., Dietrich, S., Kenzine, D., & Matschinger, H. (2005). Public attitudes toward psychiatric treatment: An international comparison. *Social Psychiatry*

and Psychiatric Epidemiology, *40*(11), 855–864. https://doi.org/10.1007/s00127-005-0958-x

Angermeyer, M. C., van der Auwera, S., Carta, M. G., & Schomerus, G. (2017). Public attitudes towards psychiatry and psychiatric treatment at the beginning of the 21st century: a systematic review and meta-analysis of population surveys. *World Psychiatry*, *16*(1), 50–61. https://doi.org/10.1002/wps.20383

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.1177/009286150103500418

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, *24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Cress, U., & Held, C. (2013). Harnessing collective knowledge inherent in tag clouds. *Journal of Computer Assisted Learning*, *29*(3), 235–247. https://doi.org/10.1111/j.1365-2729.2012.00491.x

Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers and Education*, *60*(1), 59–73. https://doi.org/10.1016/j.compedu.2012.06.015

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry*, *12*(2), 137–148. https://doi.org/10.1002/wps.20038

De Maat, S., Dekker, J., Schoevers, R., & De Jonghe, F. (2006). Relative efficacy of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Psychotherapy Research*, *16*(5), 566–578. https://doi.org/10.1080/10503300600756402

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest, Supplement*, *5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Dunning, D., Heath, C., & Suls, J. M. (2011). Flawed Implications for Health , Education , and the Workplace. *Psychological Science*, *5*(3), 69–106.

Ferguson, A. M., McLean, D., & Risko, E. F. (2015). Answers at your fingertips: Access to the Internet influences willingness to answer questions. *Consciousness and Cognition*, *37*, 91–102. https://doi.org/10.1016/j.concog.2015.08.008

Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for explanations: How the internet inflates estimates of internal knowledge. *Journal of Experimental Psychology: General*, *144*(3), 674–687. https://doi.org/10.1037/xge0000070

Fu, W.-T., Kannampallil, T., Kang, R., & He, J. (2010). Semantic imitation in social tagging. *ACM Transactions on Computer-Human Interaction*, *17*(3), 1–37. https://doi.org/10.1145/1806923.1806926

Gawronski, B., & Bodenhausen, G. V. (2011). *The associative-propositional evaluation model. Theory, evidence, and open questions*. *Advances in Experimental Social Psychology* (1st ed., Vol. 44). Elsevier Inc. https://doi.org/10.1016/B978-0-12-385522-0.00002-0

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–88. https://doi.org/10.1037/a0015701

Heckner, M., Neubauer, T., & Wolff, C. (2008). Tree, funny, to_read, google: What are tags supposed to Achieve? A comparative analysis of user keywords for different digital resource types. In *Proceedings of the 2008 ACM workshop on Search in social media* (pp. 3–10). https://doi.org/10.1145/1458583.1458589

Held, C., Kimmerle, J., & Cress, U. (2012). Learning by foraging: The impact of individual knowledge and social tags on web navigation processes. *Computers in Human Behavior*, *28*(1), 34–40. https://doi.org/10.1016/j.chb.2011.08.008

Hosanagar, K., Fleder, D., Lee, D., & Buja, A. (2014). Will the global village fracture into tribes recommender systems and their effects on consumer fragmentation. *Management Science*, *60*(4), 805–823. https://doi.org/10.1287/mnsc.2013.1808

Jorm, A. F., Nakane, Y., Christensen, H., Yoshioka, K., Griffiths, K. M., & Wata, Y. (2005). Public beliefs about treatment and outcome of mental disorders: a comparison of Australia and Japan. *BMC Medicine*, *3*. https://doi.org/10.1186/1741-7015-3-12

Kammerer, Y., Amann, D. G., & Gerjets, P. (2015). When adults without university education search the Internet for health information: The roles of Internet-specific epistemic beliefs

and a source evaluation intervention. *Computers in Human Behavior*, *48*, 297–309. https://doi.org/10.1016/j.chb.2015.01.045

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*, *29*(3), 1193–1203. https://doi.org/10.1016/j.chb.2012.10.012

Kammerer, Y., & Gerjets, P. (2014). The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid Interface. *International Journal of Human–Computer Interaction*, *30*(3), 177–191. https://doi.org/10.1080/10447318.2013.846790

Klayman, J. (1995). Varieties of Confirmation Bias. *Psychology of Learning and Motivation - Advances in Research and Theory*, *32*(C), 385–418. https://doi.org/10.1016/S0079-7421(08)60315-1

Klayman, J., Ha, Y., Edwards, W., Gnepp, J., Hoch, S., Hogarth, R., … Tra-, T. (1987). Confirmation , Disconfirmation , and Information in Hypothesis Testing. *Psychological Review*, *94*(2), 211–228. https://doi.org/10.1037/0033-295X.94.2.211

Knobloch-Westerwick, S., Sharma, N., Hansen, D. L., & Alter, S. (2005). Impact of Popularity Indications on Readers' Selective Exposure to Online News. *Journal of Broadcasting & Electronic Media*, *49*(3), 296–313. https://doi.org/10.1207/s15506878jobem4903_3

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10626367

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "seizing" and "freezing". *Psychological Review*, *103*(2), 263–283. https://doi.org/10.1037/0033-295X.103.2.263

Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579. https://doi.org/10.1111/2041-210X.12512

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use

standard deviations around the mean, do use the median absolute deviation around the median. *Journal of Experimental ...*, *49*(4), 764–766.

Liao, Q. V., & Fu, W.-T. (2013). Beyond the filter bubble. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (p. 2359). https://doi.org/10.1145/2470654.2481326

Liao, Q. V., Fu, W.-T., & Mamidi, S. S. (2015). It Is All About Perspective: An Exploration of Mitigating Selective Exposure with Aspect Indicators. In *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems* (Vol. 1, pp. 1439–1448). https://doi.org/10.1145/2702123.2702570

Ma, T. (Jennifer), & Atkin, D. (2017). User generated content and credibility evaluation of online health information: A meta analytic study. *Telematics and Informatics*, *34*(5), 472–486. https://doi.org/10.1016/j.tele.2016.09.009

Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: An introductory tutorial and free, graphical user interface (LMMgui). *Frontiers in Psychology*, *6*(JAN), 1–7. https://doi.org/10.3389/fpsyg.2015.00002

Mishel, M. H. (1997). Uncertainty in acute illness. *Annual Review of Nursing Research*, *15*, 57–80.

Moore, D. A., & Healy, P. J. (2008). The Trouble With Overconfidence. *Psychological Review*, *115*(2), 502–517. https://doi.org/10.1037/0033-295X.115.2.502

Neville, K. L. (2003). Uncertainty in Illness. *Orthopaedic Nursing*. https://doi.org/10.1097/00006416-200305000-00009

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

O'Grady, L., Wathen, C. N., Charnaw-Burger, J., Betel, L., Shachak, A., Luke, R., … Jadad, A. R. (2012). The use of tags and tag clouds to discern credible content in online health message forums. *International Journal of Medical Informatics*, *81*(1), 36–44. https://doi.org/10.1016/j.ijmedinf.2011.10.001

Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social*

Psychology, *82*(5), 722–741. https://doi.org/10.1037//0022-3514.82.5.722

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*(4), 643–675. https://doi.org/10.1037//0033-295X.106.4.643

R Core Team. (2018). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.r-project.org/

Rowland, F. (2011). *The Filter Bubble: What the Internet is Hiding from You. portal: Libraries and the Academy* (Vol. 11). Penguin Press. https://doi.org/10.1353/pla.2011.0036

Schweiger, S., Oeberst, A., & Cress, U. (2014). Confirmation bias in web-based search: A randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of Medical Internet Research*, *16*(3), e94. https://doi.org/10.2196/jmir.3044

Schwind, C., Buder, J., Cress, U., & Hesse, F. W. (2012). Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking? *Computers and Education*, *58*(2), 787–796. https://doi.org/10.1016/j.compedu.2011.10.003

Shipley, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modelling*, *7*(2), 206–218.

Shipley, B. (2013). The AIC model selection method applied to path analytic models compared using a d-separation test. *Ecology*, *94*(3), 560–564. https://doi.org/10.1890/12-0976.1

Skaug, A. H., Fournier, D., Nielsen, A., Magnusson, A., Bolker, B., & Bolker, M. Ben. (2018). Package ' glmmADMB .'

Van Der Auwera, S., Schomerus, G., Baumeister, S. E., Matschinger, H., & Angermeyer, M. (2017). Approval of psychotherapy and medication for the treatment of mental disorders over the lifespan. An age period cohort analysis. *Epidemiology and Psychiatric Sciences*, *26*(1), 61–69. https://doi.org/10.1017/S2045796015001134

Vydiswaran, V., Zhai, C., Roth, D., & Pirolli, P. (2012). BiasTrust Teaching Biased Users About Controversial. In *Proceedings of the 21st ACM …* (pp. 1905–1909). Retrieved from http://dl.acm.org/citation.cfm?id=2398541

White, R. W., & Horvitz, E. (2015). Belief Dynamics and Biases in Web Search. *ACM Transactions on Information Systems*, *33*(4), 1–46. https://doi.org/10.1145/2746229

Wickham, H. (2017). tidyverse: Easily Install and Load the "Tidyverse." Retrieved from https://cran.r-project.org/package=tidyverse

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. https://doi.org/10.1111/j.1083-6101.2012.01596.x

Winter, S., & Krämer, N. C. (2014). A Question of Credibility - Effects of Source Cues and Recommendations on Information Selection on News Sites and Blogs. *The European Journal of Communication Research*, *39*, 435–456.

Zell, E., & Krizan, Z. (2014). Do People Have Insight Into Their Abilities? A Metasynthesis. *Perspectives on Psychological Science*, *9*(2), 111–125. https://doi.org/10.1177/1745691613518075

Publication 3

Schweiger, S., & Cress, U. (2019). How Confidence in Prior Attitudes, Social Tag Popularity, and Source Credibility Shape Confirmation Bias Toward Antidepressants and Psychotherapy in a Representative German Sample: Randomized Controlled Web-Based Study. *Journal of medical Internet research*, *21*(4), e11081.

Publication 3

Abstract

Background: In health-related, Web-based information searches, people should select information in line with expert (vs nonexpert) information, independent of their prior attitudes and consequent confirmation bias.

Objective: This study aimed to investigate confirmation bias in mental health–related information searches, particularly (1) if high confidence worsens confirmation bias, (2) if social tags eliminate the influence of prior attitudes, and (3) if people successfully distinguish high and low source credibility.

Methods: In total, 520 participants of a representative sample of the German Web-based population were recruited via a panel company. Among them, 48.1% (250/520) participants completed the fully automated study. Participants provided *prior attitudes* about antidepressants and psychotherapy. We manipulated (1) *confidence* in prior attitudes when participants searched for blog posts about the treatment of depression, (2) *tag popularity* — either psychotherapy or antidepressant tags were more popular, and (3) *source credibility* with banners indicating high or low expertise of the tagging community. We measured *tag* and *blog post* selection, and *treatment efficacy ratings* after navigation.

Results: Tag popularity predicted the proportion of selected antidepressant tags (beta=.44, SE 0.11; *P*<.001) and blog posts (beta=.46, SE 0.11; *P*<.001). When confidence was low (−1 SD), participants selected more blog posts consistent with prior attitudes (beta=−.26, SE 0.05; *P*<.001). Moreover, when confidence was low (−1 SD) and source credibility was high (+1 SD), the efficacy ratings of attitude-consistent treatments increased (beta=.34, SE 0.13; *P*=.01).

Conclusions: We found correlational support for defense motivation account underlying confirmation bias in the mental health–related search context. That is, participants tended to select information that supported their prior attitudes, which is not in line with the current scientific evidence. Implications for presenting persuasive Web-based information are also discussed.

Introduction

Do people attend to information independent of their prior attitudes, and do they distinguish expert from non-expert sources on the web? To address these important questions (Del Vicario, Scala, Caldarelli, Stanley, & Quattrociocchi, 2017; Kimmerle et al., 2017; Liao & Fu, 2013), we investigate confirmation bias, the tendency to favourably select and evaluate

attitude consistent information (Del Vicario et al., 2017; Klayman, 1995; Nickerson, 1998; White & Horvitz, 2015).

A comprehensive meta-analysis identified two major motivational factors that moderate confirmation bias (Hart et al., 2009). First, when we face information that suggests our point of view is wrong, we try to maintain our prior attitudes by choosing and believing attitude consistent information, which is called *defense motivation* (Hart et al., 2009; Leary, Terry, Allen, & Tate, 2009; Wood, 2000).

In contrast to this, in some situations we may be genuinely interested in acquiring objectively correct and accurate information (Hart et al., 2009; Kitchin, 2005; Wood, 2000). This *accuracy motivation* can guide our information search, even when information is not consistent with our prior attitudes (Hart et al., 2009). Particularly in the health-context, we should form attitudes independent of our defense mechanisms and base evaluations on objectively correct information. In the following, we outline three factors that may reduce confirmation bias, given that we are accuracy motivated when searching for mental health-related information.

Confidence and Confirmation Bias

First, low confidence should decrease confirmation bias (Hart et al., 2009). However, people tend to be overly confident in prior attitudes and knowledge (Koriat, Lichtenstein, & Fischhoff, 1980; Moore & Healy, 2008) in a large range of domains, such as academic, intellectual, vocational, athletic, and medicine (Zell & Krizan, 2014). When people are overly confident in their prior attitudes confirmation bias increases (Park, Konana, & Gu, 2010).

For the mental health-related context, it is important that confidence varies for people with different mental disorders (Rouault, Seow, Gillan, & Fleming, 2018). For example, individuals who experience anxious and depressive symptoms, show less than average confidence (but average accuracy) in decision-making tasks (Rouault et al., 2018), which suggests that they could be even less prone to confirmation bias.

Looking at how to influence confidence, overconfidence can be reduced when participants reflect on their ability to describe, in a step-by-step manner, the causal functioning of objects to experts (Johnson, Murphy, & Messer, 2016). We draw on a manipulation that focussed on people recalling situations where they were either confident or doubtful about their own thoughts (Petty, Briñol, & Tormala, 2002). When participants recalled situations in which they had been confident (vs. doubtful), and subsequently provided arguments about a controversial topic, they were more (less) confident about their arguments (Briñol & Petty,

2009; Petty et al., 2002). For the current study, one goal was to replicate the manipulation (Petty et al., 2002) with a representative sample, in the mental health context.

A recent review has shown that shown that confidence manipulations tend to increase confirmation bias, which is explained by the defense motivation account (Hart et al., 2009). According to defense motivation, when people have low confidence, they aim to defend their self-concept by selecting information that is in line with their attitudes. In contrast to this, we draw on a metacognitive manipulation of confidence, that aims to make people perceive their current thoughts as less valid (Petty et al., 2002), and consequently they should perceive their attitudes as less valid (independent of their self-concept), and confirmation bias should decrease, given that searchers aim for valid information.

We expect that when prior attitudes are held with high confidence, participants preferably select and evaluate attitude consistent information. If participants were defense motivated, high (vs. low) confidence would make them less (vs. more) threatened by attitude inconsistent information, and they would select more attitude inconsistent information, and evaluate it more favourable (Hart et al., 2009).

Social Tags as Signposts to Popular Information

The second influence on confirmation bias occurs when people face cues from socially aggregated information on the web (Cress, Held, & Kimmerle, 2013; Ley & Seitlinger, 2015; Muchnik, Aral, & Taylor, 2013a; Schweiger, Oeberst, & Cress, 2014; Seitlinger, Ley, & Albert, 2012, 2015). Cues indicating socially aggregated information include star ratings, likes, retweet counts, or social tags. In the case of tagging, tag clouds arise when users label or tag content on the web, such as videos, images or documents (Figure 3.1; (Bateman, Gutwin, & Nacenta, 2008; Gligorov, Hildebrand, van Ossenbruggen, Schreiber, & Aroyo, 2011)). When tags from the tagging community are aggregated and presented in tag clouds, the tags represent the consent of a majority of people, and guide information searchers (Muchnik, Aral, & Taylor, 2013b; Schweiger et al., 2014). High majority consent or high tag popularity translates into large tags, which attract more attention than smaller tags with less social consent.

We suggest that social tag clouds are particularly non-intrusive and therefore highly suited to circumvent the influence of prior attitudes, since larger tags are visually dominating, and it has been shown that people who primarily attend to large tags (Bateman et al., 2008; Kuo, Hentrich, Good, & Wilkinson, 2007; Lohmann, Ziegler, & Tetzlaff, 2009), are more likely to click on large tags (Cress & Held, 2013; Held, Kimmerle, & Cress, 2012; Schweiger et al.,

2014), even when large tags are inconsistent with activated associations in memory (Cress & Held, 2013; Held et al., 2012), or prior attitudes (Schweiger et al., 2014). Moreover, social consent elicits behaviour that conforms to the majority in offline settings (Bohner, Dykema-Engblade, Tindale, & Meisenhelder, 2008; Darke et al., 1998).

Moreover, people select more trustworthy results when facing a grid-like (vs. list-like) arrangement of search results, similar to social tag clouds (Kammerer & Gerjets, 2014). In sum, tag clouds should be suited to decrease the influence of prior attitudes in information search and reduce confirmation bias.
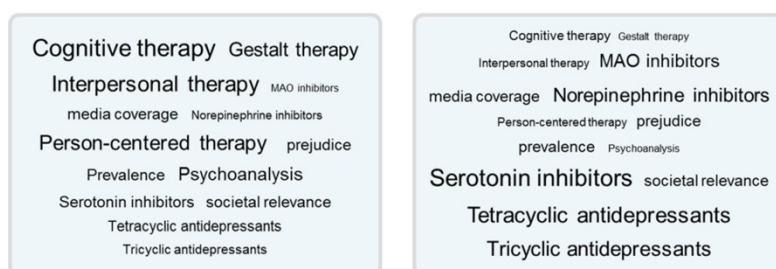

Figure 3.1 The tag clouds used in the present study. Either psychotherapy (left), or antidepressants (right) were more popular.

Source Credibility of the Tagging Community

People do not always successfully consider high quality information (Ma & Atkin, 2017; Yang & Beatty, 2016), particularly not when browsing user generated content (Ma & Atkin, 2017). Two meta analyses concluded that personal characteristics (Yang & Beatty, 2016), as well as platform characteristics (Ma & Atkin, 2017) play an important role. The relationship between manipulated source credibility and perceived information credibility is higher for student samples (vs. non-student samples) (Yang & Beatty, 2016), and user generated content that is presented on common websites (vs. blogs and discussion boards) (Ma & Atkin, 2017).

Besides education, epistemic beliefs can influence how people perceive source credibility. For example, when searching information on two competing therapies for Bechterew's disease, not all participants took source credibility into account (Kammerer, Bråten, Gerjets, & Strømsø, 2013). Particularly participants who viewed the web as a reliable (vs. not reliable) source of accurate knowledge did not reflect upon source credibility, they viewed URLs for a shorter time, and selected less search results at the bottom of the page.

To our knowledge, there is a single study using tag clouds to investigate source credibility in the health context (O'Grady et al., 2012). People searched for information on how to manage

diabetes on a health forum with tag clouds (O'Grady et al., 2012). In the first task, they searched for information that was of general interest, and in a second task, they searched for credible information. The tag cloud consisted of 12 filler tags (e.g. glucose, diet, exercise), and six tags indicating source credibility of content (author, date, quote, reference, statistics, testimonial). When participants browsed for general interest, only one third used at least a source credibility tag. When explicitly asked to take source credibility into account, 90% used at least one source credibility tag.

It remains an open question, whether people in a representative sample take the source credibility on a social tagging platform into account. In line with the accuracy motivation account, we expect that if information searchers recognize high source credibility, they will select more tags and related blog posts in total, regardless of whether attitude consistent or attitude inconsistent tags are more popular in the social tag cloud. If, on the other hand, people showed defense motivation (Hart et al., 2009), they would avoid attitude-inconsistent tags and blog posts with high source credibility, and evaluate it less favourable.

Prior Attitudes towards Antidepressants and Psychotherapy

With respect to the treatment of depressive disorders, people clearly favour psychotherapy over antidepressants (Angermeyer, Breier, Dietrich, Kenzine, & Matschinger, 2005; Angermeyer, Matschinger, & Schomerus, 2013; A. F. Jorm, 2000; Anthony F. Jorm, Christensen, & Griffiths, 2006; Anthony F Jorm et al., 2005; Schomerus et al., 2012; Van Schaik et al., 2004). Attitudes of laypeople manifest in estimated treatment efficacy as well as treatment recommendations (Angermeyer, van der Auwera, Carta, & Schomerus, 2017; Anthony F. Jorm et al., 2006; Anthony F Jorm et al., 2005; Schweiger et al., 2014; Van Der Auwera, Schomerus, Baumeister, Matschinger, & Angermeyer, 2017). People believe antidepressants to be little to moderately effective, whereas psychotherapy is believed to be moderately to highly effective (Schweiger et al., 2014). Since literature shows about equal, moderate efficacy of both types of treatment (Amick et al., 2015; Cuijpers et al., 2013; De Maat, Dekker, Schoevers, & De Jonghe, 2006), people's attitudes and recommendations are biased.

We expect more positive prior attitudes towards psychotherapy than towards antidepressants in the German population, and with the current study we aim to describe the magnitude of the psychotherapy preference and present the arguments which shape these biased attitudes.

Hypotheses

First, we expect that people's attitudes (H1a) and efficacy ratings (H1b) prior to navigation are more favourable for psychotherapy than for antidepressants.

We expect that high (vs. low) *confidence* leads to a more pronounced confirmation bias, an increased selection of attitude consistent tags (H2b) and blog posts (H2c), and this will strengthen the attitudes people already had before navigation (H2d). So, when prior attitudes favour psychotherapy, and confidence is high, participants prefer psychotherapy tags and blog posts, and change their attitudes even more towards psychotherapy. If confidence is low, prior attitudes should not be related to selection of tags and blog posts, and attitude change.

Tag popularity should circumvent the influence of prior attitudes, so participants select popular tags more frequently than less popular tags (H3a) and blog posts (H3b). Consequently, attitudes change in line with tag popularity (H3c).

Participants distinguish high from low *source credibility* (H4a). When tags and blog posts were collected by experts (vs. novices), participants click on more tags (H4b), and blog posts (H4c) overall, independent their prior attitudes, and people should show more attitude change for both treatments (H4d).

Results

Participants

A representative sample with respect to age and gender was randomly drawn from a pool of a panel company. 520 (100%) participants started the fully automated online study, 250 (48%) completed it, 7 (1%) withdrew their data, further 17 (3%) participants were dropped as they did not provide responses (Figure 3.2). Age of the remaining 226 (43%) participants ranged from 18 to 60 years (mean = 40.36, SD = 12.17), 113 (50%) were female (Table 2.1). With respect to familiarity of the technology used in the study, 56 (25%) stated they were familiar with the term "tag cloud", 83 (37%) stated they had already clicked on single tags to navigate the web. Ethical approval was granted by the Ethical Committee of the Knowledge Media Research Center (LEK 2014/006).

Table 3.1 Sample characteristics.

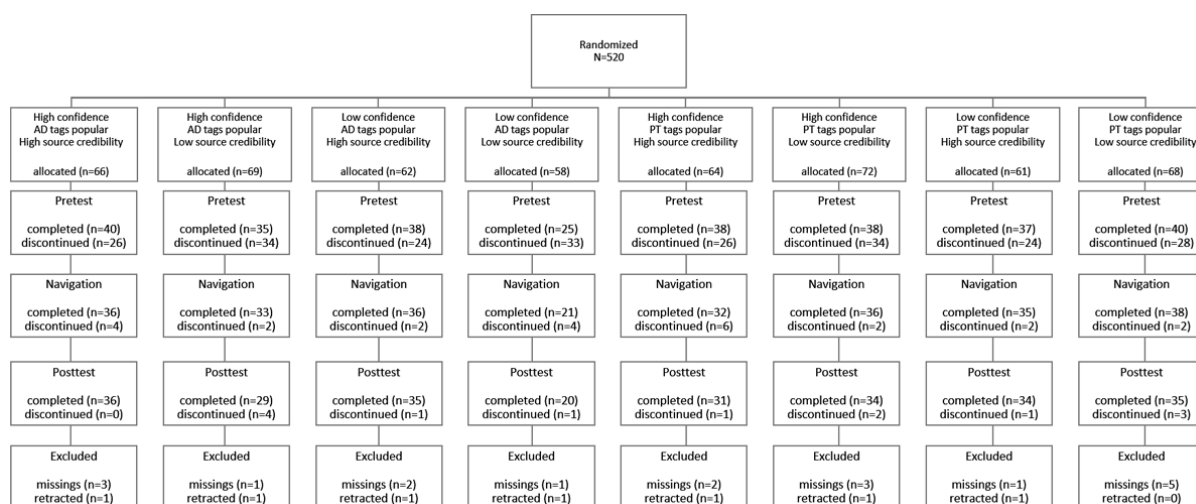| Characteristic		N	%
Education			
	Qualified job	18	8.0
	Abitur certificate	102	45.1
	University degree	53	23.5
	Other	53	23.5
Age			
	18-19	9	4.0
	20-29	46	20.4
	30-39	45	19.9
	40-49	65	28.8
	50-59	52	23.0
	60	9	4.0
Gender			
	Male	113	50.0
	Female	113	50.0
Total		226	100

Figure 3.2 Participant flow diagram.

Procedure and Design

The study comprised a 2 (confidence: high, low) x 2 (tag popularity: antidepressants high, psychotherapy high) x 2 (tagging source credibility: high, low) between-subjects design. Participants enrolled via online portal of a private panel company (respondi AG, Cologne, Germany; ISO 26362 certified), which linked to our survey, and participants were offered 4 € to complete it. First, participants were welcomed and informed that they could withdraw participation at any point. Participants were granted anonymity and asked to provide informed consent by clicking the button to start the study, after which they were randomly assigned to one of the six experimental conditions by a computerized random number procedure. Then, for *prior attitudes*, we asked participants to state pro and contra arguments regarding antidepressants and psychotherapy (Pretest tasks I, Figure 3.3). Next, they rated the efficacy of antidepressants and psychotherapy on scales. Then they provided responses for an allegedly unrelated pilot study, which served to manipulate *confidence* (Petty et al., 2002). Participants were asked to recall situations in which they had either felt confident or doubtful about their own knowledge (Petty et al., 2002). After this, they were asked to think back to their arguments regarding psychotherapy and antidepressants and they rated how confident they were about the arguments they had provided before. This rating served as a manipulation check for confidence.

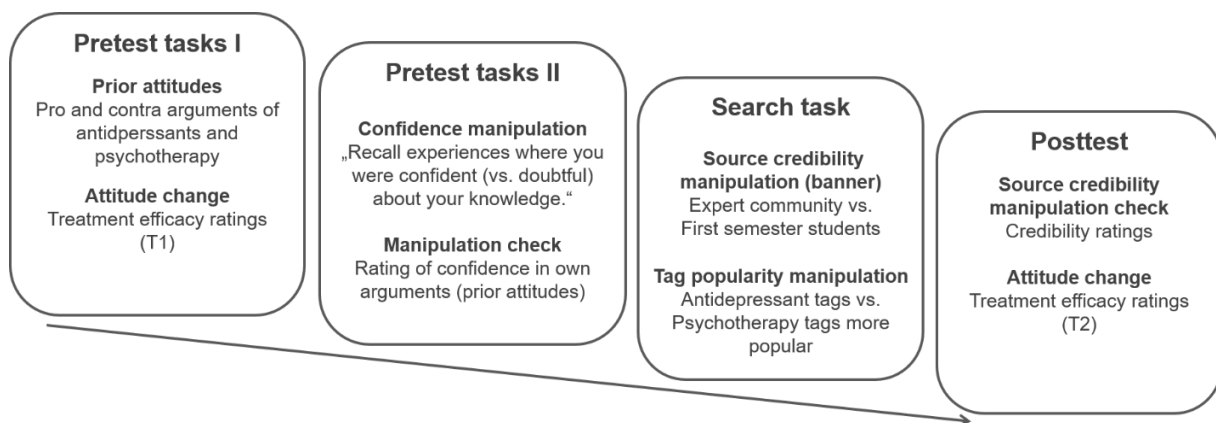**Figure 3.3** Experimental procedure.

Next, participants searched for information about treatment efficacy to provide treatment advice for a hypothetical, closely related person. To manipulate *source credibility*, we informed them that the blog post and responding tag had been gathered by a community of either experts in the field such as experienced psychiatrists and by psychotherapists (high

source credibility condition), or by psychology students and medical students in their first semester (low source credibility condition). To manipulate *tag popularity*, either psychotherapy or antidepressant tags were larger (Figure 3.1). They could also provide tags for blog posts. After five minutes of browsing in the tagging environment, a "Next" button appeared and from then on participants could decide, when to stop browsing tags and related blog posts. After navigation, participants rated source credibility (manipulation check) and provided efficacy ratings again. At the end of the study participants could provide feedback in a text box.

Materials

For the information search task, we provided a tagging environment. We presented 14 tags on the right side of the screen. Five tags represented psychotherapy and five tags represented antidepressant treatments. Four filler tags were irrelevant for treatment (prejudice, media coverage, societal relevance, prevalence). Depending on the experimental condition either psychotherapy-related tags or antidepressants-related tags were larger (i.e., more popular). When participants clicked on a tag, three blog posts were presented to the left. Each blog post described a symptom of depressive disorders and scientific studies on the efficacy of the treatment. In a pilot study, we had assured that the blog posts had equal persuasiveness. First only the headline and the first sentence of each of the three related blog posts was shown. To read the full post, participants clicked on "(more…)".

Independent Variables

Prior Attitudes. As index of pro and contra arguments for psychotherapy and antidepressants, we subtracted the number of arguments favouring antidepressants (contra psychotherapy, pro antidepressants) from the number of arguments favouring psychotherapy (pro psychotherapy, contra antidepressants). Positive values of this index thus indicate a preference for psychotherapy. Arguments were rated by two raters ($r = .78$, $P < .001$), differences were resolved by agreement.

Confidence. We adapted the experimental procedure by Petty and colleagues (Petty et al., 2002), and participants recalled situations in which they had felt confident or doubtful about their own knowledge, using five input text boxes for five minutes.

Tag Popularity. For the psychotherapy popular group, psychotherapy tags were larger, and for the antidepressant popular group, antidepressants tags were larger (Figure 3.1).

Source Credibility. On top of the page, banners showed that either alleged college students (low source credibility; Figure 3.4) or domain experts (high source credibility; Figure 3.5) had collected and tagged the blog posts. After the search task, participants rated the source credibility of the information on a scale from 1 (not at all) to 7 (highly).

Confidence Ratings (Manipulation check). After participants listed situations in which they had been (un)confident, they rated confidence in their own arguments regarding prior attitudes on a scale from 1 (not at all) to 7 (highly). They were asked how the following words described their arguments: obvious, dubious, justified, credible, factual, well-founded, persuasive and objective (Cronbach alpha = .88).

Source Credibility Ratings (Manipulation check). Participants rated the degree to which the following words described the tagging community: informed and competent (r = .70, P < .001).

Figure 3.4 Banner for the low source credibility condition.

Figure 3.5 Banner for the high source credibility condition.

Dependent Variables

Efficacy Ratings (Attitude Change). Participants agreed to statements on the efficacy of psychotherapy and antidepressants on a scale from 1 (completely disagree) to 7 (completely agree), before (antidepressants Cronbach alpha = .89, psychotherapy Cronbach alpha = .92), and after navigation (antidepressants Cronbach alpha = .94, psychotherapy Cronbach alpha = .95). To predict attitude change with respect to treatment preference, we derived a difference index score, subtracting the antidepressant from psychotherapy treatment ratings.

Besides attitude change in terms of treatment preference, we analysed pooled attitude change by taking the sum of efficacy ratings for both treatments prior to and after navigation (divided it by the number of items for interpretability).

Tag and Blog Post Selection. To measure attitude consistent navigation, we recorded the number of tags and blog posts selected for each treatment category (0 = psychotherapy, 1 = antidepressants).

Results

All analyses presented were conducted with the R software (Version 3.3.4).

Prior Attitudes

As expected in H1a we found that participant's prior attitudes favour psychotherapy over antidepressants. Participants stated more arguments for psychotherapy (mean = 1.69, SD = 1.77) than for antidepressants (mean = 1.06, SD = 1.51; t(225) = 5.30, P < .001, d = .26), and they stated more arguments against antidepressants (mean = 1.51, SD = 1.53) than against psychotherapy (mean = 0.73, SD = 1.54; t(225) = 8.13, P < .001, d = .34). We also descriptively analysed arguments and pooled them into qualitative categories (Figure 3.6).
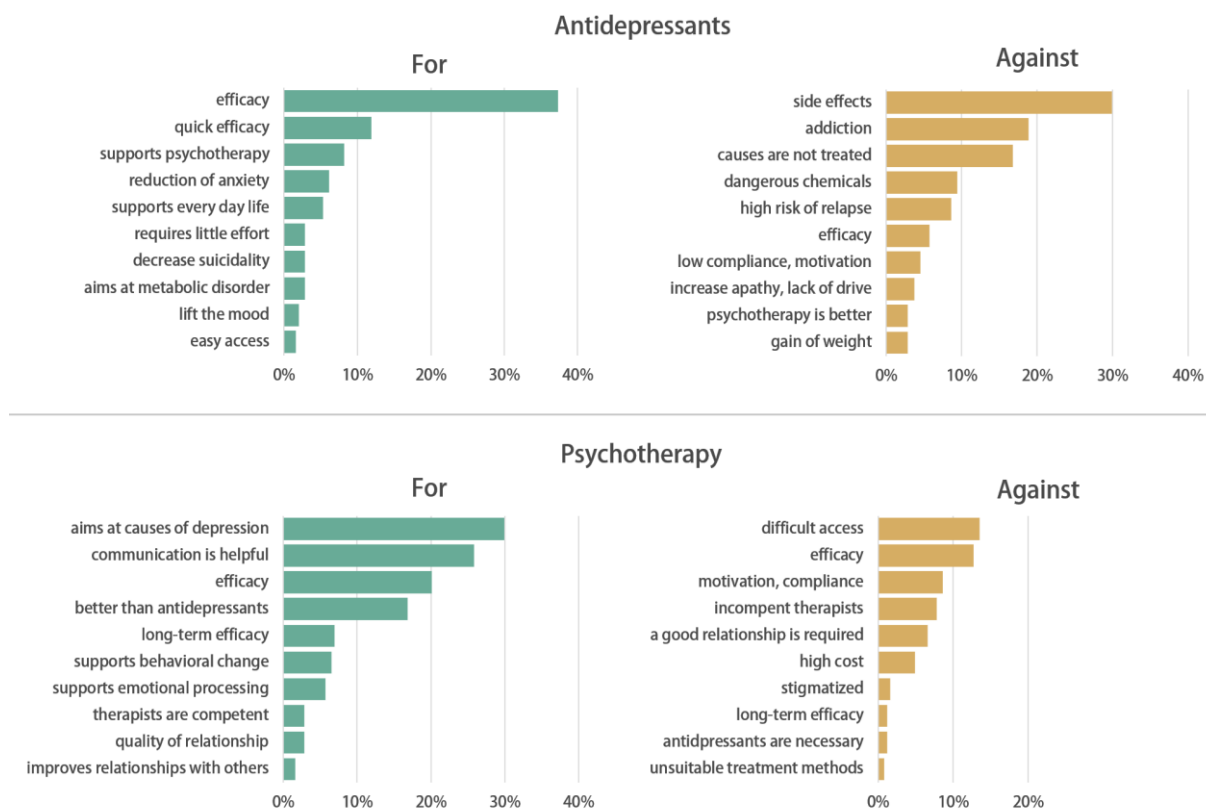

Figure 3.6 Arguments for and against the two treatments.

With H1b we expected that people would provide more favourable efficacy ratings for psychotherapy compared to antidepressants before navigation. Participants rated statements about the efficacy of both treatments on 8 items, on a scale from 1 to 7 (Figure 3.7). Since the internal consistency was high for both scales (antidepressants Cronbach alpha = .89, psychotherapy Cronbach alpha = .92), we pooled them. A paired t-Test showed a moderate effect on the preference for psychotherapy (mean = 5.24, SD = 1.10) over antidepressants (mean = 4.61, SD = 1.19; t(225) = 9.71, P < .001, d = .56; see items and response distribution in Figure 3.7). In sum, prior attitudes measured via pro and contra arguments, as well as via efficacy ratings, favoured psychotherapy over antidepressants. Both measures were moderately correlated (r = .41, P <.001).
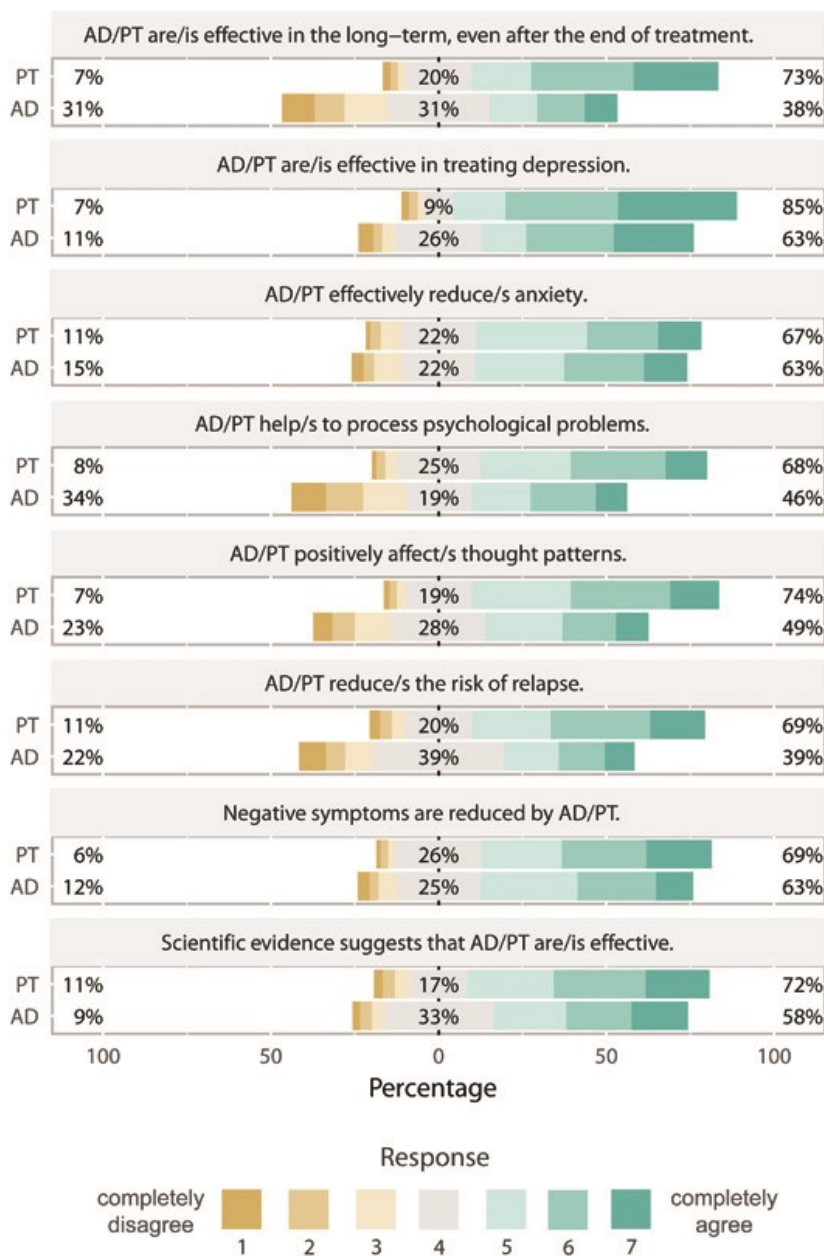
Figure 3.7 Prior attitudes about psychotherapy and antidepressants prior to information search.

Manipulation Checks

Confidence. Contrary to our expectations (H2a), we could not replicate the confidence manipulation (Petty et al., 2002). After recalling situations in which they had been confident (mean = 4.64, SD = 1.20), participants were not more confident about their arguments, compared to recalling situations in which they had been doubtful (mean = 4.68, SD = 1.09; t(224) < 1, P = .78).

Source Credibility. In contrast to our expectation (H4a), source credibility ratings in the high source credibility condition (mean = 4.87, SD = 1.26) did not significantly differ from source credibility ratings in the low source credibility condition (mean = 5.16, SD = 1.32; $t(224) = 1.67$, $P = .10$). Since the confidence and source credibility manipulations were ineffective, we used respective manipulation check scores in the following regression analyses as predictors.

Confidence in Prior Attitudes

Tag Selection

To analyse attitude consistent tag selection, we conducted logistic regressions with the dependent variable clicks on attitude (in)consistent treatment tags. The number of clicks on the respective treatment (0 = psychotherapy tag selected, 1 = antidepressant tag selected) was entered in a logistic regression (Table 3.2). As predictors we entered *prior attitudes* and *tag popularity* (0 = Psychotherapy tags popular, 1 = antidepressant tags popular), *confidence ratings* and *source credibility ratings* (see independent variables). We included 2-way interaction terms (Table 3.2, Step 2), and tested for the presence interactions with likelihood ratio tests (Menard, 1995; Peng, Lee, & Ingersoll, 2002).

We expected that high confidence should strengthen the relationship between prior attitudes and the proportion of clicks on attitude consistent tags (H2b). However, there was no significant interaction of the predictors confidence in prior attitude ratings and prior attitudes, predicting the selection of antidepressant tags (Table 3.2, Step 2). As likelihood ratio tests showed, including three-way interaction ($\chi^2(10) = 4.91$, $P = .90$), and four-way interaction ($\chi^2(11) = 4.98$, $P = .93$), did not improve model fit.

Table 3.2 Selection ratio of antidepressant tags.

	Step 1			Step 2		
	B	SE	P	B	SE	P
Intercept	-0.39	0.08	<.001	-0.41	0.08	<.001
Prior attitudes	-0.02	0.03	.37	-0.03	0.03	.32
Confidence score	0.002	0.05	.97	0.04	0.05	.82
Tag popularity	0.44	0.11	<.001	0.44	0.11	<.001
Source credibility score	-0.005	0.04	.92	-0.005	0.04	.92
PA x confidence score				-0.01	0.03	.65
	Model fit:			Model fit change (vs. Step 1):		
	$\chi^2(4) = 17.89$, P = .001			$\chi^2(1) = 0.20$, P = .65		

B, Beta coefficients, continuous predictors were centered; SE,

Standard Error of Beta coefficients; PA, Prior attitudes.

Blog Post Selection

A second logistic regression used the same predictors as in the regression predicting tag selection but with blog post selection as criterion variable (Table 3.3). We expected that high confidence should strengthen the impact of prior attitudes and consequently lead to increased proportion of clicks on attitude consistent blog posts (H2c). We observed an interaction between confidence and prior attitudes (beta = 0.11, SE = .02, P < .001). To disentangle the interaction, we compared slopes for high (+1 SD) and low (-1 SD) confidence ratings. This showed that when confidence ratings were low (-1 SD), participants selected a higher proportion of blog posts that were in line with their prior attitudes (beta = -0.26, SE = 0.05, P < .001; Figure 3.8). When confidence ratings were high (+ 1SD), there was no association with prior attitudes (beta = 0.02, SE = .03, P = .57; Figure 3.8). In contrast to our expectation, and in line with the defense motivation account, when confidence was low but not high, there was an association between prior attitudes and selection of attitude consistent blog posts.

Table 3.3 Selection ratio of antidepressant blog posts.

	Step 1			Step 2		
	B	SE	P	B	SE	P
Intercept	-0.75	0.07	<.001	-0.87	0.08	<.001
Prior attitudes	-0.05	0.03	.06	-0.12	0.03	<.001
Confidence score	-0.11	0.05	.02	-0.04	0.05	.46
Tag popularity	0.44	0.11	<.001	0.45	0.11	<.001
Source credibility score	0.02	0.04	.73	0.03	0.04	.52
PA x confidence score				0.11	0.02	<.001
	Model fit:			Model fit change (vs. Step 1):		
	$\chi^2(4) = 30.41$, P < .001			$\chi^2(1) = 25.56$, P < .001		

B, Beta coefficients, continuous predictors were centered; SE,

Standard Error of Beta coefficients; PA, Prior attitudes.
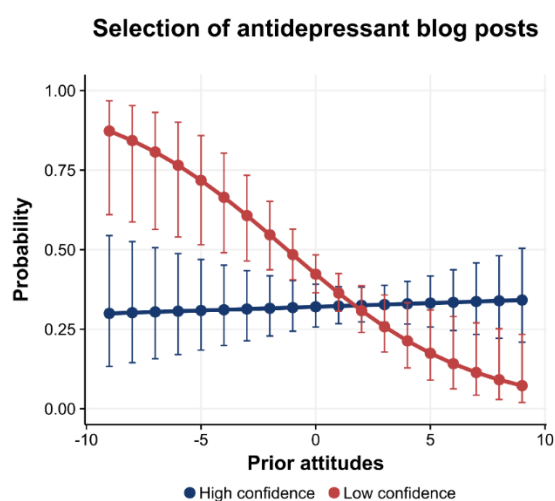

Figure 3.8 Predicted proportion of antidepressant blog posts selected, for high (+1 SD) and low (-1 SD) confidence (95% CI), with negative values indicating a preference for AD over PT.

Compared to the model including the 2-way interaction term (Table 3.3, Step 2), neither including three-way interaction term ($\chi^2(4) = 5.82$, P = .21), nor including the four-way interaction term ($\chi^2(5) = 5.99$, P = .31) yielded a better model fit (all respective lower order interaction terms were included as well).

Attitude Change

We conducted multiple linear regressions. First, with the predictor variables prior attitudes, confidence ratings, and source credibility ratings (all centered), and the dichotomous variable tag popularity (0 = psychotherapy popular, 1 = antidepressants popular). Additionally we included the predictor difference score of efficacy ratings (antidepressants subtracted from psychotherapy) prior to navigation, to analyze attitude change with a covariate approach (Van Breukelen, 2006). As criterion for attitudes after navigation, we included the difference score of efficacy ratings (Table 3.4).

We expected that high (vs. low) confidence would lead to higher confirmation bias and decreased attitude change, so for people who hold their attitudes with high (vs. low) confidence, prior attitudes should be more strongly associated with attitudes after navigation (H2d). The expected interaction between confidence and prior attitudes was not significant (Table 3.4, Step 2). However, the association between confidence and prior attitudes depended on source credibility (Table 3.4, Step 3). To disentangle this 3-way interaction, simple slopes were tested on low (-1 SD) and high (+1 SD) levels of source credibility ratings and confidence ratings. This revealed a strong association between prior attitudes and treatment efficacy ratings after navigation, for participants with lower ratings of confidence (-1 SD) and high source credibility ratings (+1 SD; beta = 0.34, SE = 0.13, P = .01), but no association for high confidence ratings (+1 SD) and low source credibility ratings (-1 SD; beta = 0.11, SE = 0.06, P = .053). There was also no association with low confidence (-1 SD) and low source credibility ratings (-1 SD; beta = 0.08, SE = 0.10, P = .42), and with high confidence (+1 SD) and high source credibility ratings (+1 SD; beta = -0.04, SE = .06, P = .50; Figure 3.9).

Table 3.4 Treatment Efficacy Ratings (AD-PT) after Navigation.

	Step 1			Step 2			Step 3		
	B	SE	P	B	SE	P	B	SE	P
Intercept	0.70	0.08	<.001	0.73	0.08	<.001	0.74	0.08	<.001
Efficacy ratings (PT-AD) prior to navigation	0.79	0.06	<.001	0.78	0.06	<.001	0.79	0.06	<.001
Tag popularity	-0.05	0.12	.64	-0.06	0.12	.60	-0.05	0.11	.65
PA	0.08	0.03	.01	0.10	0.06	.005	0.12	0.03	<.001
Confidence score	-0.08	0.05	.16	-.10	0.06	.08	-0.10	0.06	.07
Source credibility score	0.04	0.05	.34	0.05	0.05	.33	0.11	0.05	.03
PA x confidence score				-0.04	0.03	.16	-0.07	0.03	.02
PA x source credibility score							0.02	0.02	.40
Confidence score x source credibility score							-0.03	0.04	.53
PA x confidence score x source credibility score							-0.07	0.02	<.001
	Model fit:			Model fit (vs. Step 1):			Model fit (vs. Step 2):		
	adj. R^2 = .49,			Δ adj. R^2 < .01,			Δ adj. R^2 = .02,		
	$F(5,220) = 43.68$,			$F(1,219) = 1.94$,			$F(3,216) = 3.88$,		
	$P < .001$			$P = .16$			$P = .01$		

B, Beta coefficients, the continuous predictors were centered;

SE, Standard Error of Beta coefficients; PA, Prior attitudes.
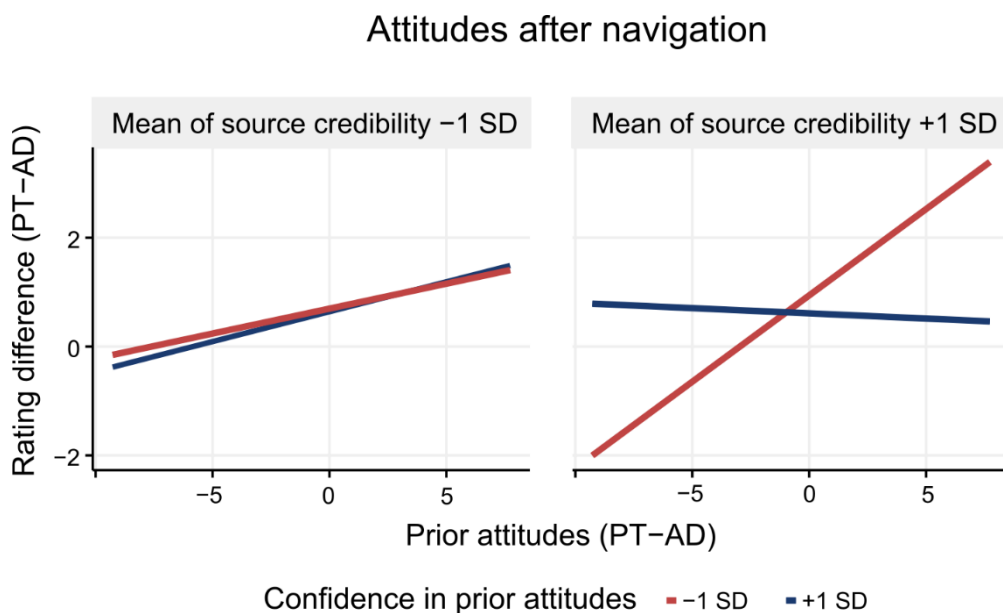
Attitudes after navigation

Figure 3.9 Prior attitudes, confidence and source credibility and treatment efficacy ratings after navigation, with negative values on all axes indicating a preference for AD over PT.

Tag Popularity of Treatments in the Social Tag Cloud

Tag Selection

In H3a we expected that popular treatment tags would be selected more often, independent of prior attitudes. To test this, we used a logistic regression model as described in the previous confidence section on tag selection (see Table 3.2, step 1). Tag popularity was the only significant predictor for the proportion of selected antidepressant tags (beta = 0.44, SE = .11, P < .001). This supports H3a, as participants selected a larger proportion of popular tags in the tag cloud. They did this independent of their prior attitudes as prior attitudes were not associated with tag selection.

Blog Post Selection

We also expected that participants would select more blog posts when related tags were more popular (H3b). We tested this with the logistic regression as described in the previous confidence section on blog post selection (see Table 3.3). This showed that participants selected a larger proportion of blog posts when related tags were popular in the tag cloud, supporting H3b (beta = 0.44, SE = .11, P < .001; Table 3.3, Step 1).

Attitude Change

We expected in H3c that the attitude change of would depend on tag popularity. More popular tags for a treatment should be associated with higher ratings of treatment efficacy. We conducted a linear regression analysis, as described in the previous confidence section (see also Table 3.4). We did not find an influence of tag popularity on efficacy ratings after navigation (beta = -0.05, SE = .12, P = .64; see Table 3.4, Step 1).

Source Credibility of the Tagging Community

Tag Selection

We expected that when the tagging community consists of experts (vs. novices), participants click on more tags (H4b). To test this, we conducted a negative binomial regression with the continuous, centered predictors source credibility score, prior attitudes, confidence score, the dichotomous predictor tag popularity (0 = psychotherapy tags popular, 1 = antidepressant tags popular), and the dependent variable number of selected tags (Table 3.5). We did not find support for H4b, as the total number of selected tags was not associated with source credibility ratings.

Blog Post Selection

We expected that when the tagging community consists of experts (vs. novices), participants click on more blog posts (H4c). We conducted another negative binomial regression with the continuous, centered predictors source credibility score, prior attitudes, confidence score, the dichotomous predictor tag popularity (0 = psychotherapy tags popular, 1 = antidepressant tags popular), and the dependent variable total number of selected blog posts (Table 3.5). As with the number of selected tags, perceived source credibility did not predict the total number of selected blog posts, not supporting H4c.

Table 3.5 Tags and blog posts selected.

	Number of Tags selected			*Number of Blog Posts selected*		
	B	SE	P	B	SE	P
Intercept	1.81	0.07	<.001	1.16	0.47	.01
PA	0.01	0.03	.63	0.03	0.04	.48
Confidence score	-0.02	0.05	.62	0.06	0.08	.43
Tag popularity	0.14	0.10	.18	-0.34	0.18	.06
Source credibility score	0.03	0.04	.51	0.02	0.07	.79
	Model fit:			Model fit:		
	$\chi^2(4) = 2.67$, P = .61			$\chi^2(6) = 5.10$, P = .28		

B, Beta coefficients, continuous predictors were centered;

SE, Standard Error of Beta coefficients; PA, Prior attitudes.

Attitude Change

When the tagging community consisted of experts (vs. novices), we expected that participants should change their prior attitudes to a greater degree (H4d). We conducted a linear regression analysis with efficacy ratings prior to navigation (AD-PT) as a covariate (Van Breukelen, 2006), and included the predictors tag popularity (0 = psychotherapy tags popular, 1= antidepressant tags popular), prior attitudes, confidence, and source credibility ratings. As criterion, we included the efficacy rating difference (AD-PT) after navigation. The covariate and the continuous predictor variables were centered (Aiken & West, 1991). High perceived source credibility was associated with higher treatment efficacy ratings after navigation (Table 3.6), supporting H4d.

Table 3.6 Treatment Efficacy Ratings (AD+PT) after Navigation.

	Ratings T2 (AD+PT)		
	B	SE	P
Intercept	10.26	0.12	<.001
Efficacy ratings prior to navigation (AD + PT)	0.76	0.05	<.001
Tag popularity	0.02	0.18	.93
PA	0.08	0.04	.06
Confidence in PA score	0.14	0.09	.12
Source credibility score	0.24	0.07	<.001

Model fit:

adj. R^2 = .16, $F(5,220)$ = 9.47, P < .001

B, Beta coefficients; SE, Standard Error of Beta coefficients;

continuous predictors and critereon were centered; PA, Prior attitudes.

Discussion

With this randomized, controlled study, we aimed to investigate prior attitudes about antidepressants and psychotherapy, and the tendency to confirm prior attitudes when selecting and evaluating mental health-related information. We presented three factors to counter confirmation bias: popularity of treatment tags in a social tag cloud, confidence and the source credibility of the tagging community. We expected that people would select and favourably evaluate attitude inconsistent content, when confidence was low (vs. high). And we expected that *source credibility* and *tag popularity* should influence selection of tags independent of prior attitudes. We could not replicate the confidence manipulation (Petty et al., 2002), and participants did not distinguish source credibility as presented by banners, therefore we used manipulation check scores for correlational analyses.

As expected, people in the German population rated psychotherapy as more effective than antidepressants, and they reported according beliefs. Increasing tag popularity increased selection of tags, independent of prior attitudes and confidence. In contrast to our expectations, higher source credibility was not associated with increased tag or blog post selection. Participants with high confidence were more open to select attitude inconsistent blog posts, which is in line with the defense motivation account, but not with the accuracy motivation account we had expected (Hart et al., 2009). Moreover, we found people with low confidence rated treatment efficacy in accordance with their prior attitudes, but only when perceived source credibility was high.

Social Tags to Reduce Confirmation Bias

We expected that social tag clouds are a non-intrusive interface to circumvent prior attitudes, and popular tags would be selected more often in dependent of prior attitudes. We found that people selected popular tags and related blog posts more often. We think that these findings highlight the important role of popular content on the web, also in the context of mental health-related selection of information. When two treatment options are presented to a searcher, searchers will be guided by more popular information, even independent of their prior attitudes. This could help to design online platforms in which it is desirable to minimize the influence of prior attitudes and maximize the influence of a community.

A thorough discussion about nudges is beyond the scope of this paper, but we consider implications of implementing tag clouds as nudges. Whereas nudges are controversial in

general (Sunstein, 2015), educational nudges aiming to aid people in making better decisions are less controversial (Sunstein, 2015). Moreover, in the health context it is argued that it is impossible not to be influenced by policies of different stakeholders in general (Quigley, 2013). The way in which tools such as tag clouds influence behaviour might be considered more controversial since large tags automatically attract the searchers' attention (Bateman et al., 2008), thus influencing information selection (Cress et al., 2013), and therefore tags may restrict deliberate individual agency (Le Clainche & Tubeuf, 2015; Sunstein, 2015).

Defense Motivation in Mental Health-Related Information Search

We expected that people would be guided by accuracy motivation when searching mental health-related information. People would strive to select and evaluate information that is objectively correct, regardless of their prior attitudes. In contrast this, the pattern of results suggests that information searchers were defense motivated, they tended to confirm their prior attitudes, in order to avoid dissonant cognitions and to maintain a positive view of themselves (Hart et al., 2009; Kitchin, 2005; L. & Festinger, 1965).

This was reflected in blog post selection and resulting attitude change. We found that low confidence was associated with selecting attitude consistent blog posts, which suggests that participants may have felt increased threat under low confidence.

The findings on attitude change provide further support for the defense motivation account. People with high confidence were expected to change their attitudes in line with their prior attitudes. However, we found the opposite. When confidence was low, not high, people's attitudes after navigation were polarized in line with their prior attitudes. In contrast to blog post selection, however, this pattern was only found when source credibility was high but not when source credibility was low. This suggests that attitude inconsistent information could have posed a double threat when source credibility was high, in combination with low confidence. In all other instances, there was no association between prior attitudes and attitude change.

What follows from defensive processing? Not only when information presents a direct threat (e.g. anti-smoking images), but also when different treatment options are available, prior attitudes have an impact on online information search. When information acknowledges prior attitudes of the reader, the need to maintain a positive self-view can be reduced, and the reader becomes more open to attitude-inconsistent information (Harris & Napper, 2005; van Koningsbruggen, Das, & Roskos-Ewoldsen, 2009). So, content authors could anticipate the

attitudes of their readers when providing health information and acknowledge existing attitudes and views before providing potentially conflicting information.

Source credibility and Confirmation Bias

People do sometimes recognize source credibility on the web (Kammerer, Bråten, Gerjets, & Strømsø, 2012; Kammerer & Gerjets, 2014; Ma & Atkin, 2017), however, participants did not rate practitioners with years of experience as a more credible, compared to students of health-related subjects in their first semester. One possible explanation for this is that the banners on top of the page were too subtle.

Moreover, for student samples (vs. non-student samples) (Yang & Beatty, 2016), and content that is presented on common websites (vs. user generated content) (Ma & Atkin, 2017), searchers perceive experts as more credible. This might explain that for this representative sample on a specific tagging platform people did not distinguish high from low source credibility.

For content authors, this finding underlines the importance to consider the target audience as well as the impact of the type of platform that is being used to convey health-related messages. Whereas information searchers with high educational background, or searchers on general websites respond more to expertise, searchers on sites presenting user generated content (e.g. forums or blog posts) respond more to demographic similarity to the searcher (Ma & Atkin, 2017), and non-student searchers respond less to expertise when judging source credibility.

Attitude Confidence and Confirmation Bias

A recent study showed that individuals with depressive and anxiety symptoms exhibited lower confidence in a decision making task (Rouault et al., 2018). In the current study, people with lower confidence evaluated information content in line with their prior attitudes, when the source of information was highly credible. So particularly when searchers perceive information as highly credible, individuals with depressive or anxiety symptoms might be prone to select attitude consistent information. This should also be tested by future studies.

Public Attitudes towards Antidepressants and Psychotherapy

As for student (Schweiger et al., 2014) and representative samples in Germany (Angermeyer et al., 2013), we also expected prior attitudes to be more positive for psychotherapy than for antidepressants, and we found an according moderate effect. The results about the specific beliefs show that people are not satisfied with the current accessibility of mental health care

services, and the German population seems to have specific beliefs when it comes to side effects of antidepressants. However, side effects that can be found in the literature, such as nausea, insomnia, somnolence, fatigue, sexual dysfunction, and weight gain (Khawam, Laurencic, & Malone, 2006; Papakostas, 2008), were rarely associated with antidepressants.

Limitations

According to the Federal Office of Statistics, the current sample is representative for gender and age, but participants with lower education, such as people with a qualified job, are underrepresented, whereas participants with university degree are slightly overrepresented (Statistisches Bundesamt, 2018). Therefore, results of the current study should be interpreted with caution for people with lower level education. The recruitment process of the panel company uses online campaigns, search engine marketing and offline recruitment, where participants register at a portal where they can enrol for studies that match their demographics. So, it should be noted that this sample is restricted to online users of the German population.

The current study suggests that results for confidence and its interplay with source credibility are in line with predictions of defense motivation, however, due to the correlational design, potential correlated confounding influences could be at work and potentially could have been overlooked.

Moreover, all blog posts highlighted the efficacy aspect of prior attitudes, other important issues such as side effects or treatment of psychological causes were not mentioned in the blog posts. Thus, only one aspect related to prior attitudes, namely treatment efficacy was addressed in the blog posts. And all blog posts were formulated positively, such that information revealing limitations and boundary conditions of the treatments were addressed in the blog posts.

Since age could be an important covariate in this study, we exploratively checked the influence of age for each dependent variable, however, age was not a significant predictor in none of the analyses.

Conclusion

We presented correlational support for the defense motivation account in health-related search. That is, participants tended to confirm their prior attitudes when searching for information. We presented factors that influence this confirmation bias: First, social tags

reduced the influence of prior attitudes, and second, attitude confidence increased confirmation bias when source credibility was high. These findings have many implications for content creators, who should acknowledge existing attitudes in persuasive communication and consider demographics of their audience as well as the type of platform where content is published. Future studies should test whether this result extends to other health-related domains, beyond treatment of depression, and to other information platforms as well. Further it would be highly interesting to compare treatment attitudes toward internet-based psychotherapy including different delivery modes.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression. Multiple regression: Testing and interpreting interactions*. https://doi.org/10.1080/00031305.1972.10478934

Amick, H. R., Gartlehner, G., Gaynes, B. N., Forneris, C., Asher, G. N., Morgan, L. C., … Lohr, K. N. (2015). Comparative benefits and harms of second generation antidepressants and cognitive behavioral therapies in initial treatment of major depressive disorder: systematic review and meta-analysis. *Bmj*, *351*, h6019. https://doi.org/10.1136/bmj.h6019

Angermeyer, M. C., Breier, P., Dietrich, S., Kenzine, D., & Matschinger, H. (2005). Public attitudes toward psychiatric treatment: An international comparison. *Social Psychiatry and Psychiatric Epidemiology*, *40*(11), 855–864. https://doi.org/10.1007/s00127-005-0958-x

Angermeyer, M. C., Matschinger, H., & Schomerus, G. (2013). Attitudes towards psychiatric treatment and people with mental illness: Changes over two decades. *British Journal of Psychiatry*, *203*(2), 146–151. https://doi.org/10.1192/bjp.bp.112.122978

Angermeyer, M. C., van der Auwera, S., Carta, M. G., & Schomerus, G. (2017). Public attitudes towards psychiatry and psychiatric treatment at the beginning of the 21st century: a systematic review and meta-analysis of population surveys. *World Psychiatry*, *16*(1), 50–61. https://doi.org/10.1002/wps.20383

Bateman, S., Gutwin, C., & Nacenta, M. (2008). Seeing things in the clouds: the effect of visual features on tag cloud selections. *ACM Conference on Hypertext and Hypermedia*, *4250*, 193–202. https://doi.org/10.1145/1379092.1379130

Bohner, G., Dykema-Engblade, A., Tindale, R. S., & Meisenhelder, H. (2008). Framing of

majority and minority source information in persuasion: When and how "consensus implies correctness." *Social Psychology*, *39*(2), 108–116. https://doi.org/10.1027/1864-9335.39.2.108

Briñol, P., & Petty, R. E. (2009). *Persuasion. Insights from the Self-Validation Hypothesis. Advances in Experimental Social Psychology* (1st ed., Vol. 41). Elsevier Inc. https://doi.org/10.1016/S0065-2601(08)00402-4

Cress, U., & Held, C. (2013). Harnessing collective knowledge inherent in tag clouds. *Journal of Computer Assisted Learning*, *29*(3), 235–247. https://doi.org/10.1111/j.1365-2729.2012.00491.x

Cress, U., Held, C., & Kimmerle, J. (2013). The collective knowledge of social tags: Direct and indirect influences on navigation, learning, and information processing. *Computers and Education*, *60*(1), 59–73. https://doi.org/10.1016/j.compedu.2012.06.015

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry*, *12*(2), 137–148. https://doi.org/10.1002/wps.20038

Darke, P. R., Chaiken, S., Bohner, G., Einwiller, S., Erb, H. P., & Hazlewood, J. D. (1998). Accuracy motivation, consensus information, and the law of large numbers: Effects on attitude judgment in the absence of argumentation. *Personality and Social Psychology Bulletin*, *24*(11), 1205–1215. https://doi.org/10.1177/01461672982411007

De Maat, S., Dekker, J., Schoevers, R., & De Jonghe, F. (2006). Relative efficacy of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Psychotherapy Research*, *16*(5), 566–578. https://doi.org/10.1080/10503300600756402

Del Vicario, M., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2017). Modeling confirmation bias and polarization. *Scientific Reports*, *7*(December 2016), 1–9. https://doi.org/10.1038/srep40391

Gligorov, R., Hildebrand, M., van Ossenbruggen, J., Schreiber, G., & Aroyo, L. (2011). On the role of user-generated metadata in audio visual collections. *Proceedings of the 6th International Conference on Knowledge Capture (K-CAP '11)*, *0*, 145–151. https://doi.org/10.1145/1999676.1999702

Harris, P. R., & Napper, L. (2005). Self-affirmation and the biased processing of threatening

health-risk information. *Personality and Social Psychology Bulletin*, *31*(9), 1250–1263. https://doi.org/10.1177/0146167205274694

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological Bulletin*, *135*(4), 555–88. https://doi.org/10.1037/a0015701

Held, C., Kimmerle, J., & Cress, U. (2012). Learning by foraging: The impact of individual knowledge and social tags on web navigation processes. *Computers in Human Behavior*, *28*(1), 34–40. https://doi.org/10.1016/j.chb.2011.08.008

Johnson, D. R., Murphy, M. P., & Messer, R. M. (2016). Reflecting on explanatory ability: A mechanism for detecting gaps in causal knowledge. *Journal of Experimental Psychology: General*, *145*(5), 573–588. https://doi.org/10.1037/xge0000161

Jorm, A. F. (2000). Mental health literacy: Public knowledge and beliefs about mental disorders. *British Journal of Psychiatry*, *177*(NOV.), 396–401. https://doi.org/10.1192/bjp.177.5.396

Jorm, A. F., Christensen, H., & Griffiths, K. M. (2006). The public's ability to recognize mental disorders and their beliefs about treatment: Changes in Australia over 8 years. *Australian and New Zealand Journal of Psychiatry*, *40*(1), 36–41. https://doi.org/10.1111/j.1440-1614.2006.01738.x

Jorm, A. F., Nakane, Y., Christensen, H., Yoshioka, K., Griffiths, K. M., & Wata, Y. (2005). Public beliefs about treatment and outcome of mental disorders: a comparison of Australia and Japan. *BMC Medicine*, *3*. https://doi.org/10.1186/1741-7015-3-12

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2012). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2012.10.012

Kammerer, Y., Bråten, I., Gerjets, P., & Strømsø, H. I. (2013). The role of Internet-specific epistemic beliefs in laypersons' source evaluations and decisions during Web search on a medical issue. *Computers in Human Behavior*, *29*(3), 1193–1203. https://doi.org/10.1016/j.chb.2012.10.012

Kammerer, Y., & Gerjets, P. (2014). The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid Interface. *International Journal of Human–Computer Interaction*, *30*(3), 177–191.

https://doi.org/10.1080/10447318.2013.846790

Khawam, E. A., Laurencic, G., & Malone, D. A. (2006). Side effects of antidepressants: An overview. *Cleveland Clinic Journal of Medicine*, *73*(4), 351–361. https://doi.org/10.3949/ccjm.73.4.351

Kimmerle, J., Bientzle, M., Cress, U., Flemming, D., Greving, H., Grapendorf, J., … Sassenberg, K. (2017). Motivated Processing of Health-Related Information in Online Environments. *Informational Environments: Effects of Use, Effective Designs*.

Kitchin, P. L. (2005). Measuring the amount of statistical information in the EPT index. In *Environmetrics* (Vol. 16, pp. 51–59). https://doi.org/10.1002/env.670

Klayman, J. (1995). Varieties of Confirmation Bias. *Psychology of Learning and Motivation - Advances in Research and Theory*, *32*(C), 385–418. https://doi.org/10.1016/S0079-7421(08)60315-1

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 175–220. Retrieved from http://psycnet.apa.org/journals/xlm/6/2/107/

Kuo, B. Y.-L., Hentrich, T., Good, B. M. ., & Wilkinson, M. D. (2007). Tag clouds for summarizing web search results. *Proceedings of the 16th International Conference on World Wide Web - WWW '07*, 1203. https://doi.org/10.1145/1242572.1242766

L., J., & Festinger, L. (1965). Conflict, Decision and Dissonance. *Revue Française de Sociologie*. https://doi.org/10.2307/3318996

Le Clainche, C., & Tubeuf, S. (2015). Nudging, intervening or rewarding: A discussion on the constraints and the degree of control on health status. *Politics, Philosophy and Economics*, *15*(2), 170–189. https://doi.org/10.1177/1470594X15600832

Leary, M. R., Terry, M. L., Allen, A. B., & Tate, E. B. (2009). The concept of ego threat in social and personality psychology: Is ego threat a viable scientific construct? *Personality and Social Psychology Review*, *13*(3), 151–164. https://doi.org/10.1177/1088868309342595

Ley, T., & Seitlinger, P. (2015). Dynamics of human categorization in a collaborative tagging system: How social processes of semantic stabilization shape individual sensemaking. *Computers in Human Behavior*, *51*(PA), 140–151.

https://doi.org/10.1016/j.chb.2015.04.053

Liao, Q. V., & Fu, W.-T. (2013). Beyond the filter bubble. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (p. 2359). https://doi.org/10.1145/2470654.2481326

Lohmann, S., Ziegler, J., & Tetzlaff, L. (2009). Comparison of tag cloud layouts: Task-related performance and visual exploration. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5726 LNCS*(PART 1), 392–404. https://doi.org/10.1007/978-3-642-03655-2_43

Ma, T. (Jennifer), & Atkin, D. (2017). User generated content and credibility evaluation of online health information: A meta analytic study. *Telematics and Informatics*, *34*(5), 472–486. https://doi.org/10.1016/j.tele.2016.09.009

Menard, S. (1995). *Applied Logistic Regression Analysis*. *Sage University Paper Series on Quantitative Applications in the Social Sciences* (Vol. 38). SAGE publications. https://doi.org/10.4135/9781412983433

Moore, D. A., & Healy, P. J. (2008). The Trouble With Overconfidence. *Psychological Review*, *115*(2), 502–517. https://doi.org/10.1037/0033-295X.115.2.502

Muchnik, L., Aral, S., & Taylor, S. J. (2013a). Social influence bias: A randomized experiment. *Science*, *341*(6146), 647–651. https://doi.org/10.1126/science.1240466

Muchnik, L., Aral, S., & Taylor, S. J. (2013b). Social influence bias: A randomized experiment. *Science*, *341*(6146), 647–651. https://doi.org/10.1126/science.1240466

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

O'Grady, L., Wathen, C. N., Charnaw-Burger, J., Betel, L., Shachak, A., Luke, R., … Jadad, A. R. (2012). The use of tags and tag clouds to discern credible content in online health message forums. *International Journal of Medical Informatics*, *81*(1), 36–44. https://doi.org/10.1016/j.ijmedinf.2011.10.001

Papakostas, G. I. (2008). Tolerability of modern antidepressants. *J Clin Psychiatry*, *69 Suppl E*, 8–13. https://doi.org/18494538

Park, J., Konana, P., & Gu, B. (2010). Confirmation Bias, Overconfidence, and Investment

Performance. *The Social Science Research Network Electronic Paper Collection*, (October), 59. https://doi.org/http://dx.doi.org/10.2139/ssrn.1639470

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *The Journal of Educational Research*, *96*(1), 3–14. https://doi.org/10.1080/00220670209598786

Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology*, *82*(5), 722–741. https://doi.org/10.1037//0022-3514.82.5.722

Quigley, M. (2013). Nudging for health: On public policy and designing choice architecture. *Medical Law Review*, *21*(4), 588–621. https://doi.org/10.1093/medlaw/fwt022

Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2017.12.017

Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., & Angermeyer, M. C. (2012). Evolution of public attitudes about mental illness: A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, *125*(6), 440–452. https://doi.org/10.1111/j.1600-0447.2012.01826.x

Schweiger, S., Oeberst, A., & Cress, U. (2014). Confirmation bias in web-based search: A randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of Medical Internet Research*, *16*(3), e94. https://doi.org/10.2196/jmir.3044

Seitlinger, P., Ley, T., & Albert, D. (2012). An Implicit-Semantic Tag Recommendation Mechanism for Socio-Semantic Learning Systems An Implicit-Semantic Tag Recommendation Mechanism. In L. Tobias, R. Mikko, L. Mart, & T. Arthur (Eds.), *International Conference, Open and Social Technologies for Networked Learning Open and Social Technologies for Networked Learning* (pp. 41–46). Tallinn, Estonia.

Seitlinger, P., Ley, T., & Albert, D. (2015). Verbatim and semantic imitation in indexing resources on the web: A fuzzy-trace account of social tagging. *Applied Cognitive Psychology*, *29*(1), 32–48. https://doi.org/10.1002/acp.3067

Statistisches Bundesamt. (2018). *Jahrbuch Deutschland*.

Sunstein, C. (2015). The Ethics of Nudging Recommended Citation. *Yale Journal on Regulation Article*, *32*(2). https://doi.org/10.2139/ssrn.2526341

Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, *59*(9), 920–925. https://doi.org/10.1016/j.jclinepi.2006.02.007

Van Der Auwera, S., Schomerus, G., Baumeister, S. E., Matschinger, H., & Angermeyer, M. (2017). Approval of psychotherapy and medication for the treatment of mental disorders over the lifespan. An age period cohort analysis. *Epidemiology and Psychiatric Sciences*, *26*(1), 61–69. https://doi.org/10.1017/S2045796015001134

van Koningsbruggen, G. M., Das, E., & Roskos-Ewoldsen, D. R. (2009). How Self-Affirmation Reduces Defensive Processing of Threatening Health Information: Evidence at the Implicit Level. *Health Psychology*, *28*(5), 563–568. https://doi.org/10.1037/a0015610

Van Schaik, D. J. F., Klijn, A. F. J., Van Hout, H. P. J., Van Marwijk, H. W. J., Beekman, A. T. F., De Haan, M., & Van Dyck, R. (2004). Patients' preferences in the treatment of depressive disorder in primary care. *General Hospital Psychiatry*, *26*(3), 184–189. https://doi.org/10.1016/j.genhosppsych.2003.12.001

White, R. W., & Horvitz, E. (2015). Belief Dynamics and Biases in Web Search. *ACM Transactions on Information Systems*, *33*(4), 1–46. https://doi.org/10.1145/2746229

Wood, W. (2000). Attitude Change: Persuasion and Social Influence. *Annual Review of Psychology*, *51*(1), 539–570. https://doi.org/10.1146/annurev.psych.51.1.539

Yang, Q., & Beatty, M. (2016). A meta-analytic review of health information credibility: Belief in physicians or belief in peers? *Health Information Management Journal*, *45*(2), 80–89. https://doi.org/10.1177/1833358316639432

Zell, E., & Krizan, Z. (2014). Do People Have Insight Into Their Abilities? A Metasynthesis. *Perspectives on Psychological Science*, *9*(2), 111–125. https://doi.org/10.1177/1745691613518075