Brede, Max; Mazarakis, Athanasios; Peters, Isabella

Book Chapter — Published Version

What drives researchers to look up research publications they found in the news?

Suggested Citation: Brede, Max; Mazarakis, Athanasios; Peters, Isabella (2023) : What drives researchers to look up research publications they found in the news?, In: Broer, Irene et al. (Ed.): The Science-Media Interface. On the Relation Between Internal and External Science Communication, ISBN 9783110776546, De Gruyter, Berlin, pp. 79-115, https://doi.org/10.1515/9783110776546-004

This Version is available at:
http://hdl.handle.net/11108/594

Max Brede, Athanasios Mazarakis, and Isabella Peters

4. What drives researchers to look up research publications they found in the news?

Abstract: External science communication uses media and other means, such as news reports on scientific publications, to produce awareness and understanding of science and its results. Scientific publications that were featured in the news are linked to higher citations and altmetric-counts when compared to similar unfeatured articles. So far, the question about the relationship between attributes of scientific publications, their mentions in a news report, and their effect on researchers' decision to look up a scientific publication remained unanswered: a research gap this study attempts to fill. First, we conducted a three-phased variation of a Delphi survey to generate a selection of attributes that experts deem relevant for evaluating scientific publications. Then the attributes were discussed with a focus group and optimized for a large-scale online conjoint study with 642 respondents. Statistical analysis revealed that attributes which indicate expert opinion and methodological quality are the major drivers behind looking up scientific publications mentioned in news reports. This finding underscores that forms of external science communication and the highlighting of particular publication attributes positively affect the awareness of scientific publications that are also positively related with a publication's citation counts.

Keywords: conjoint study, Delphi study, news article, science communication

1 Introduction

Science communication is defined by Burns et al. (2003) as the use of media and other means to produce awareness, enjoyment, interest, opinions, or understanding of science and its aspects. In particular, being aware of new scientific findings is an essential part of every scientist's daily life, as their work must always be up to date. In a study that examined scientists' search and reading behavior, Tenopir et al. (2019) found that scientists report newspapers as an important resource of scholarly information. This highlights the important role of "external science communication" or the communication of scientific knowledge by individuals who are not necessarily part of a scientific communi-

https://doi.org/10.1515/9783110776546-004

ty (Dernbach et al., 2012). An example of this type of communication is a newspaper report in which a journalist reports on a research publication.

Scientific publications featured in the news have been shown to receive more citations (i.e., Anderson et al., 2020; Dumas-Mallet et al., 2020; Fanelli, 2013) and higher altmetric counts (i.e., Bowman & Hassan, 2019; Lemke, 2020). Most often, this effect is discussed to be attributable to one of two possible mechanisms or a mix of both (see also Chapter 5 in this book; Lemke, 2022). The first is the so-called earmark hypothesis (Kiernan, 2003) which attributes the observable advantage to a qualitative difference in publications mentioned in news reports because the publication responds similarly well to the selection strategies of researchers and journalists alike. The idea is that researchers search for and cite similar publications (and publication attributes) like journalists, resulting in higher citation counts, regardless of the increased reach and larger audience resulting from the non-academic news report. Here, one may assume that researchers and journalists have similar mental concepts about and selection strategies for the "newsworthiness" or "relevance" of scientific publications. This thesis of inherent qualities of research publications driving citations and other forms of attention is backed up by findings from a different scenario. Breuer and colleagues (2022) used retrieval test collections to compare relevance judgments for scientific publications with their citation rates, although they have not investigated in detail the role of publication qualities for relevance decisions. They showed, though, "that documents that receive a relevance rating are more likely also to be highly cited" (Breuer et al., 2022, p. 2470) and that they receive higher altmetric attention scores. This connection is unidirectional: highly cited documents are not necessarily (more) relevant for a search task.

In contrast to this pure attribution to the publication's attributes, the publicity hypothesis states that the advantage in citations can be linked to the additional reach gained by journalistic reporting. A strong indicator of this connection is the study by Phillips et al. (1991). The authors compared a sample of research publications featured in the New York Times to one that was meant to be featured in news reports but was not, due to a strike. The authors reported that the publications featured in the unpublished edition of the newspaper did not perform any better than comparable, not-featured articles. This citation advantage associated with a mention in the New York Times was replicated by Kiernan (2003), who was also able to show that the reported effect is not specific to this prestigious outlet.

Both hypotheses are not mutually exclusive. It could also be argued that the journalistic landscape and its interactions with academia have changed drasti-

cally since these explanatory models were formulated. Examples of these changes are the rise of social media and its usage by scholars (Lemke et al., 2019).

Furthermore, neither hypothesis goes into the specifics of which attributes of a scientific publication lead to the observed advantages. However, it is reasonable to assume that certain attributes might result in different effects in both models, and the way agents interact with scientific publications might depend on different attributes. One could, for example, expect that journalists choose publications depending on specific criteria, like the "newsworthiness" of their topic that is not necessarily equal to or related to newsworthiness in science (see also Chapter 2 in this book). In contrast, researchers may choose to look up the publication and cite it, depending on its relevance to their research, because of disciplinary norms, acquaintance with the authors, or because of many other reasons that are often not explicit (Cronin, 1981; Garfield, 1962; Tahamtan & Bornmann, 2019).

In fact, such selection and decision-making processes are driven by multiple criteria that are often intertwined and prioritized according to the actual topic or situation and the role of the person who selects. For example, it has been shown for an intermediary institution between science journalism and science, the Science Media Center Germany (Broer & Pröschel, 2021), that it selects publications by acknowledging four major sets of criteria: a) journalistic criteria, such as reach, relevance for the public, urgency; b) science-internal criteria, such as quality of the journal, sample size, used method, number of authors; c) strategic criteria, such as the impact on public discourse or on agenda-setting; and d) organizational criteria, such as availability of experts and editors in the institution. Those findings highlight that science-internal and science-external selection strategies from researchers and other actors can either reinforce each other (as in the earmark hypothesis) or shed light on what is valued by the different actors of the science communication system.

To study the specifics of these interactions of internal and external scholarly communication, citations are of particular interest since they should be sensitive to publicity and earmark effects. In most cases, bibliometric citation analysis is concerned with the characteristics of the publication, its authors, and the journal in which the publication appeared. In this regard, most of these studies are based on post hoc analyses of publications whose characteristics and citations were used for analysis (Tahamtan & Bornmann, 2018).

In contrast, Tenopir and colleagues (2011) conducted a "conjoint analysis" style experiment. Conjoint analysis, also known as "discrete choice experiment" (Louviere et al., 2010) or "choice-based conjoint analysis" (Backhaus et al.,

2015), attempts to describe an entity in terms of its attributes and to identify the attributes promising the most (partial) utility of that entity to a user.

This type of survey design is primarily used in marketing and consumer research studies of latent preferences (Backhaus et al., 2015; Louviere et al., 2010). The goal is always to determine which attributes of an entity influence a participant's preferences in which magnitude. The procedure is based on a part-worth model that defines an option's value or utility as the sum of its attributes' part-worth utilities (Louviere et al., 2010). Conjoint analysis has successfully been used in a variety of settings, i.e., to test biases in the choices of healthcare stakeholders (Crabtree et al., 2022), to evaluate information leak severities (Koguchi & Maeda, 2022), or to examine the perception of privacy issues in virtual reality technology of German consumers (Schuir et al., 2022).

Tenopir and colleagues (2011) used a conjoint analysis-based survey to identify the most important features of a publication that make a potential reader want to read it. To do this, the authors examined the three attributes "author reputation," "journal prominence," and "online accessibility of the publication." The authors concluded that the accessibility of the publication is the most important of the three attributes. This was followed by the reputation of the authors and, finally, the type of journal as the least important attribute. In an additional choice experiment, the authors found that the "topic of the article" played by far the most important role. Other than that, the results of the conjoint analysis were replicated. Since a realistic choice between two publications is likely to cover similar topics, Tenopir and colleagues (2011) conclude that the three attributes considered in the experiment are of most use.

Another conjoint approach to analyze attributes influencing citations was conducted by Lemke and colleagues (2021). Their goal was to determine which bibliometric indicators are most helpful for readers when deciding whether to cite the article. They concluded that citation counts and the journal impact factor are the attributes generating the highest utility.

The research described so far relates to the scholarly reading and use of scientific literature. The state of the art of research dealing with the interaction of internal and external science communication mainly refers to the mention of scientific publications in news media in terms of an altmetric perspective. These altmetrics-centered studies are mainly concerned with the meaning of individual altmetrics (Haustein et al., 2015) or the ways to collect them, e.g., Kousha and Thelwall (2019), i.e., more with their use than with their creation. Corresponding studies mainly deal with other, non-news media-related data, such as the use of social media by scientists (Van Noorden, 2014). However, a more detailed analysis of the influences on the effect of external science communication and,

more specifically, news reports on subsequent citations of original publications beyond the theoretical consideration described above is lacking.

The empirical approach described below represents an attempt to gather initial indications about this problem. To the authors' knowledge, this study is the first attempt in this direction. So far, no detailed analysis of the interplay of publication attributes reported by news and their effect on later reads by researchers and citations of publications has been conducted. Therefore, the following reasoning underlies the study design: original research publications expose certain qualities (e.g., relevance, quality, rigor, innovation, topic), which can be translated to or operationalized via certain attributes that reflect those qualities (e.g., sample size, number of citations) and that can be linked to additional qualifiers allowing for nuance (e.g., large, groundbreaking). The attributes respond to researchers' mental selection strategies and drive decisions. The attributes can be mentioned in news media reports and other forms of external science communication that report on the original publication. We assume that 1) external science communication raises researchers' awareness of research publications (see also publicity hypothesis; Phillips et al. (1991) and Chapter 5 in this book) and 2) mentioning attributes in external science communication increases the likelihood of researchers' looking up (and then citing) the original publication.[1] The research question we attempt to answer in an exploratory examination of the aforementioned interplay focuses on the second part of our basic assumption: which are the key attributes of scientific publications whose inclusion in news reports might be beneficial to scientists in deciding to look up the publication that is being reported on?

If attributes of that kind could be found and their effect can be quantified, they could support an understanding of the extent to which news media influence scholarly citations. Since the research question aims at the implicit value-attribution of the reading scientists, a conjoint analysis was constructed based on the experiments by Tenopir et al. (2011) and Lemke et al. (2021).

The contribution of this article is threefold. First, we present a collection of expert-verified attributes of scientific publications that are deemed relevant (via a modified Delphi survey (Pollitt et al., 2016) when deciding whether a publication will be looked up (and subsequently cited) or not. The attributes served as

1 This is a simplification of the decision-making process, of course. We acknowledge that not all research publications that are looked up will be cited subsequently. However, we assume that awareness of an original publication is a necessary prerequisite for citation and that awareness can be induced via several mechanisms, e.g., formats of external science communication, search results lists, reference lists.

stimulus material in a conjoint experiment, which led to an empirically validated ranking of those attributes. This sheds light on those characteristics of research publications that drive look-up decisions the most. The third contribution comprises an evaluation of the congruence of implicit and explicit decision-making behavior of the subjects participating in the conjoint experiment: are they aware of their selection strategies and the publication's attributes they prefer most?

2 Methods

The conjoint study was constructed in multiple steps. The first step was the generation of relevant attributes and the stimulus material necessary for the online study. A Delphi-method (Gordon, 1994) oriented approach was carried out in three phases to generate an expert-judgment-based selection of attributes.

These attributes were then aggregated and specified in attribute levels to form stimulus material, which another group of experts again validated. A preliminary conjoint study was then conducted using this material to test it for the target audience. Finally, the results from this conjoint pre-study were used to optimize the design for a large-scale online conjoint study. This design process is depicted in Figure 1.

2.1 Delphi pre-study

A central step in conducting a conjoint analysis is the decision on attributes to be used for the choice sets. Since there are no prior studies examining the direct influence of the mention of specific publication characteristics in news media reports on the latter citation rates, attributes reported to influence the impact of a publication seem to be a good first selection on which to base the choice set (Tahamtan et al., 2016; Tahamtan & Bornmann, 2018; Tenopir et al., 2011, 2019). This decision was founded on our assumption that a greater impact moderated by news media distribution, as postulated in the publicity hypothesis, would be facilitated by mentioning attributes that are relevant to a scientific audience.

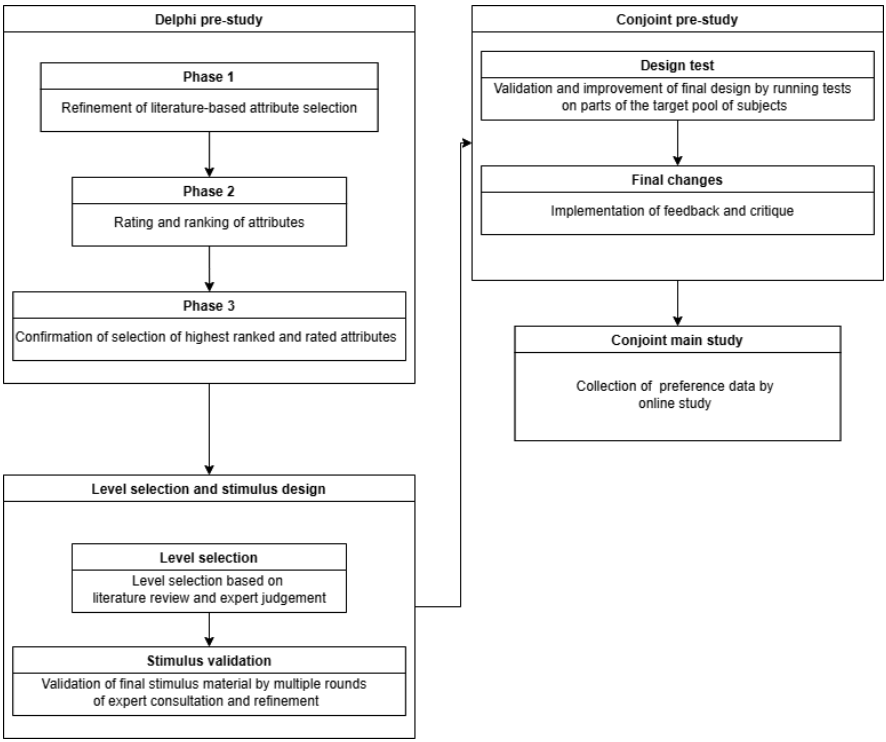

Figure 1: Overview of the study.

Considering the large variety of discussed factors influencing the impact of a publication, we first conducted a three-phased Delphi pre-study. The goal of this approach was to reduce the broad array of possible attributes to an easier-to-handle subset that could be presented in a conjoint setting. A focus group-based approach oriented on Pollitt et al. (2016) was chosen to do this. For the focus group's convenience, a LimeSurvey (LimeSurvey Project Team/ Carsten Schmitz, 2012) adaption of the Delphi method (Gordon, 1994) was implemented while trying to reach a consensus.

The main advantage of this Delphi-oriented method is its proven efficacy in reaching a productive solution based on small group discussions. Bloor et al. (2015) conclude that a sample of four experts can already produce useful results, given the balanced composition of this group. Similar group sizes of eight to 12 subjects (Robson & McCartan, 2016) or less (Krueger, 2014) have also been reported to work. We aimed to match these sizes while having a diverse group by recruiting seven scientists from different fields for our focus group (two from

information science and one each from business informatics, computer science, neuroscience, physics, and criminology). We used 64 attributes listed and classified by Tahamtan et al. (2016) and Tahamtan and Bornmann (2018) to ask our focus group whether the presented attributes might be relevant when looking up research publications mentioned in a news report. The complete list of the 64 attributes can be found in the Appendix.

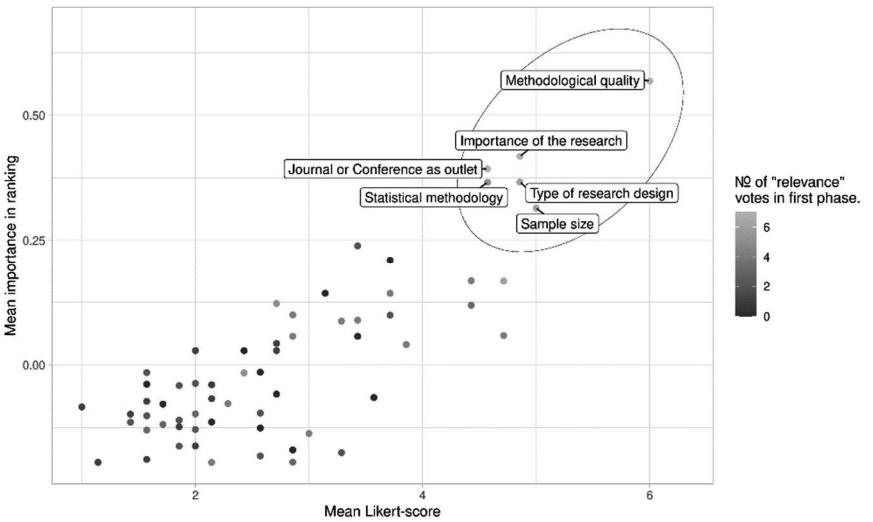

Figure 2: Results of the first and second phases of the Delphi pre-study. The importance values based on the ranking are calculated by setting the first rank to one and decreasing the value in steps of $1/n$, where n is the number of ranks assigned. Items that were not ranked were assigned a value of 0. The items were presented in German.

The focus group reported their decision on a binary scale consisting of "not relevant" and "relevant" labels. At the end of the first Delphi phase, the focus group had the chance to list additional attributes they found to be missing in the initial list. Of the initial 64, 54 attributes with at least one vote for relevance were available for the second Delphi phase. This attribute collection was supplemented by 12 additional attributes the participants mentioned to be missing. All initial attributes, the additional attributes, and the amount of "relevant" votes by the focus group are available in the Appendix. This collection marked the end of the first Delphi phase and led to the second Delphi phase.

The resulting list of 66 attributes was presented to the experts in the second Delphi phase, this time with the task to first rate the importance of each attrib-

ute on a seven-point Likert scale and then additionally to rank up to ten attributes as the most important. This approach resulted in a selection of six attributes that were ranked and rated as the most relevant, namely "methodological quality," "importance of the research," "journal or conference as an outlet," "type of research design," "statistical methodology" and "sample size." The distribution of rankings and ratings is shown in Figure 2. In addition, the results of all ratings and rankings are available in the Appendix.

In the third phase of the Delphi pre-study, the participants were asked to confirm the six selected attributes and explain their reasoning. The confirmation was collected online and within a focus group setting. This was done to avoid high scoring items being too broad in their possible interpretations. "Methodological quality," as an example, can be understood in different ways, depending on the background of the expert.

All participants of the focus group agreed to the six selected attributes. This result concluded the end of the Delphi pre-study and led to the specification of the attribute levels for all six confirmed attributes.

2.2 Specification of attribute levels

Because conjoint analyses are based on entities defined by combinations of levels of attributes, the selection of attributes had to be appended by appropriate levels in the next step. To do this, the explanations from the third phase of the Delphi pre-study and a literature review (Bhandari et al., 2007; M. Callaham et al., 2002; M. L. Callaham et al., 1998; Craig et al., 2007; Farshad et al., 2013; Figg et al., 2006; Kulkarni et al., 2007; Miettunen et al., 2002; Miettunen & Nieminen, 2003; Nieri et al., 2007; Patsopoulos et al., 2005; Willis et al., 2011) were used to generate up to four levels for each of the attributes, which were then combined into sentences to form mock-up-articles as the final choice set basis and thus the basis for the material of our conjoint study. The resulting levels for each attribute are displayed in Table 1. The number of attributes and levels thereof are based on the average conjoint designs, as reported by Marshall et al. (2010). More details about the attributes can also be found in the Appendix.

The complete mock-up design and the plausibility of the different values in Table 1 have been checked by representatives of the Science Media Center Germany who were partners in the MeWiKo research project (MeWiKo, n.d.) and who brought in journalistic expertise.

Table 1: The components of the choice sets resulting from the Delphi pre-study. The original choice sets were displayed in German.

Attribute	Value
Sentence 1	"DESIGN FORMAT was judged by scientists from the same area as IMPORTANCE."
DESIGN	The meta-analysis, which The experiment, which The study, which
FORMAT	was published in an English-language journal was published in an English-language conference
IMPORTANCE	very relevant relevant irrelevant extremely irrelevant
Sentence 2	"The study was conducted on a SAMPLESIZE sample for this research area. An STATISTICS was used to evaluate the overall methodologically QUALITY study."
SAMPLESIZE	small large
QUALITY	outstanding good bad extremely poor
STATISTICS	appropriate statistical procedure inappropriate statistical procedure

2.3 Conjoint-design

The conjoint analysis was then conducted using a lab.js-based (Henninger et al., 2019) online questionnaire, asking the participants to rank three mock-up news reports per trial. This was done by providing an instructional text at the top of each page that instructed the participants to arrange the mock-ups in the order they evoked curiosity to look up the original scientific publication. An example of such a mock-up news report can be seen in Figure 3.

The attributes' presentation was conducted with mock-up reports using a simple list of the attribute levels due to the results of the Delphi pre-study. Some of the experts in the Delphi pre-study reported issues in keeping the news media

focus of the study in mind. The presentation in the form of a news report was meant to keep the context present.

The whole study consisted of 12 trials, which were chosen based on Fedorov's exchange algorithm (Fedorov, 1972). This algorithm optimizes the information gained by an experimental design by choosing the design options that maximize the marginal effects.

Participants were invited to the study via e-mail in two steps, using a mailing list of economics researchers provided by the ZBW - Leibniz Information Centre for Economics. The first step intended to test the design on representatives of the target population to see whether they encountered any issues with the study material. Additionally, the first step was conducted to test the selected choice sets on their viability. The second step was then adapted to the lessons learned from the first step. Then the study was fully rolled out to gather and analyze the data regarding the research question, as depicted in Figure 1.

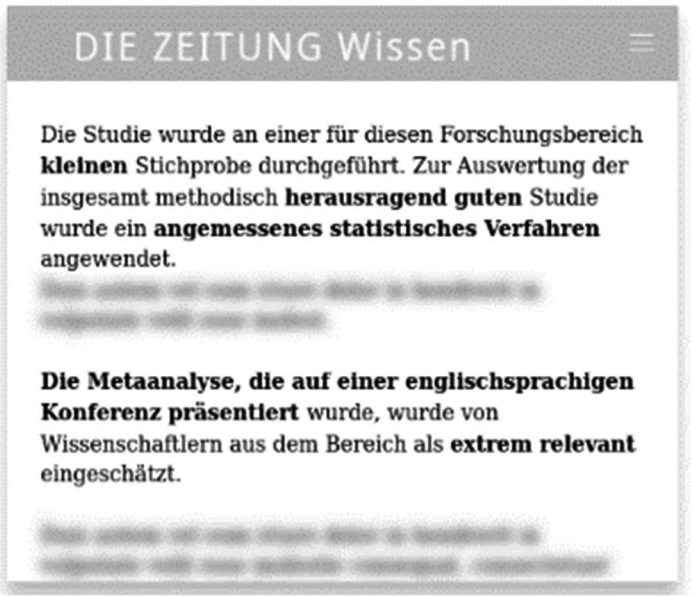

Figure 3: Mock-up news report as used in the conjoint study.

2.4 Conjoint step 1 – pre-study

The first step consisted of 2,000 e-mail invitations sent out between February 17 and March 15 2021. Of the 2,000 invited researchers, 156 accessed the study. Of those, two were excluded due to stating nonsensical employments, eight were excluded for taking more than ten minutes for one trial, and one subject was excluded for indicating not working scientifically.

Most of the 145 participants indicated working as a professor (42.86%), followed by Ph.D. students (20.41%) and research assistants (19.31%). The largest group of participants reported their research field as economics (40%), followed by business studies (34.48%). After finishing the 12 trials, the participants had the opportunity to give verbal feedback. Two independent raters categorized these verbal answers and achieved a satisfactory inter-rater reliability of $\alpha_{Krippendorff} = .742$ (cf. Krippendorff, 2004). These verbal answers were then used to adapt the study for conjoint step 2, more precisely, the actual data collection and leaving the pre-study phase. Some of these categories are now presented in detail.

The category "contradiction," which was by far the most frequently assigned by the raters (eight out of 25), was to be assigned in cases where the participants indicated a contradiction in the stimulus material, for example, as stated in the following comment: "I wonder how a study that uses an inappropriate statistical approach can nevertheless be rated as 'methodically well conducted.'"

The choice sets were redesigned to address this problem in the second conjoint step so that contradictions no longer occurred. This change was implemented because many participants described the contradictions as rendering the choice sets nonsensical. This change in design resulted in a high correlation between the attributes STATISTICS and QUALITY. Since high correlations between predictors render conditional logistic regressions unsolvable, one of these attributes must be ignored in the final statistical model.

Besides the remark about a contradiction, criticism of the length of the study and remarks about decreasing concentration were the most frequent type of feedback (five out of 25). A typical example of this type of criticism is the feedback "It is extremely difficult to stay concentrated with the very similar texts and the constant repetition." This criticism was not addressed in the second step because a limitation of statistical power was weighted more heavily than the possible effects of fatigue. However, to control whether selection decisions due to fatigue happened solely via heuristics, an item was added at the end of the questionnaire in which such heuristics were queried by asking the participants to state attribute-based decision-heuristics they were aware of.

These heuristics were tested by asking the respondents to put the attributes in the order in which they thought they influenced their decisions.

The third most frequent type of comment (four out of 25 each) was the praise of the design ("It is nice to see conjoint being used") and criticism of the stimulus material. There was especially repeated criticism about the vertical arrangement of the mock-up news reports, which resulted in annoying scrolling, as can be seen in this comment: "The display was unfortunately not optimal on my computer. Smaller images that could have been dragged down or up, if necessary, rather than to the right would have been better. Could unfortunately never see all three at once." To address this criticism, the design for the second conjoint step was changed to display mock-up news reports and categories horizontally on large screens.

We also received some comments from a few subjects who did not read the texts thoroughly but only focused on the highlighted passages. Others indicated that their decisions had been made based on a subset of these without attention to the overall context. Although these were only a few subjects, we took these indications seriously and used a manipulation check to control this aspect in the second conjoint step.

2.5 Conjoint step 2 – data collection

Data collection for the second step began on May 11, 2021, and ended on June 9, 2021. To recruit subjects for the second step, invitations were sent to 6,000 previously unused e-mail addresses from the list provided by the ZBW - Leibniz Information Centre for Economics on May 11, and reminders were sent on May 25. During the data collection period, 728 potential participants started the study, of which 36 were excluded for trials with completion times over ten minutes, ten for non-scientific employment, and 185 for not completing all 12 trials. The remaining participants took a mean of 43.9 seconds to complete a trial, with a standard deviation of 37.2 and a median of 35 seconds.

The 497 participants remaining were mostly professors (43.26%), followed by doctoral students (25.96%). As in the first step, most of these participants came from the fields of economics (34.81%) and business administration (29.18%), followed by macroeconomics (10.66%).

3 Results

In order to evaluate the data collected in the second conjoint step, two conditional logistic regression models were calculated, using the participant's arbitrary ID and the question number as stratum (this procedure is oriented on the method described in Aizaki and Nishimura (2008)). The first regression model analyzed the first rank priority, and the second regression model the second rank priority. However, as already noted, one of the two attributes, "methodological quality" (QUALITY) and (STATISTICS), must be excluded from the evaluation. Since the former was rated as much more critical in the Delphi-oriented phases held in advance, the decision for exclusion fell on "adequacy of the statistical procedure." Another argument for this exclusion are textual comments by participants, such as the following: "I come from a discipline that does virtually no quantitative or qualitative empirical work." The fact that "methodological quality" is a more flexible term and can be interpreted by many disciplines as relevant to themselves is another argument for preferring this attribute; "adequacy of statistical procedure," in contrast, is only relevant to disciplines that generate inference based on statistical models.

The results of the conditional logistic regressions are shown in Table 2. When looking at the estimated utilities, it is noticeable that "methodological quality" seems to be far more relevant for the second priority than for the first rank.

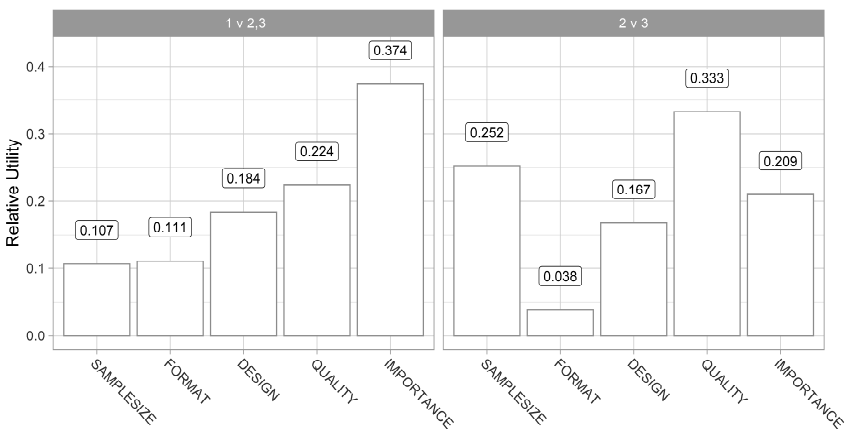

Figure 4: Relative utility of the presented attributes as the range between the highest and lowest level relativized by the sum of these ranges.

Table 2: Results of the two separate conditional logistic regression models. The first column indicates whether the decision for the first level over the others or the decision for the second level over the third was modeled.

Compared priority	Attribute	Comparative	Attribute level	95%-CI lower limit	b-Coefficient	95%-CI upper limit	t	p	
1 v 2,3	QUALITY	outstanding	good	1.079	1.459	1.840	7.514	0.000	***
1 v 2,3	QUALITY	outstanding	extremely poor	-1.352	-0.630	0.092	-1.711	0.261	
1 v 2,3	QUALITY	outstanding	bad	-0.136	0.827	1.790	1.683	0.261	
1 v 2,3	DESIGN	The Experiment, which	The study, which	-0.401	0.448	1.296	1.034	0.301	
1 v 2,3	DESIGN	The Experiment, which	The Metaanalysis, which	-1.750	-1.270	-0.789	-5.176	0.000	***
1 v 2,3	FORMAT	was published in an English-language journal	was published on an English-language conference	-1.215	-1.037	-0.860	-11.444	0.000	***
1 v 2,3	IMPORTANCE	relevant	extremely irrelevant	-3.967	-3.178	-2.389	-7.894	0.000	***
1 v 2,3	IMPORTANCE	relevant	extremely relevant	-1.972	-1.261	-0.549	-3.474	0.003	**
1 v 2,3	IMPORTANCE	relevant	irrelevant	-4.694	-3.489	-2.284	-5.674	0.000	***
1 v 2,3	SAMPLESIZE	large	small	0.695	0.997	1.298	6.483	0.000	***
2 v 3	QUALITY	outstanding	good	-1.619	-1.048	-0.477	-3.596	0.002	**
2 v 3	QUALITY	outstanding	extremely poor	-3.048	-2.426	-1.803	-7.632	0.000	***
2 v 3	QUALITY	outstanding	bad	-2.297	-1.521	-0.745	-3.841	0.001	***
2 v 3	DESIGN	The study, which	The Metaanalysis, which	-0.923	-0.556	-0.189	-2.972	0.015	*
2 v 3	DESIGN	The study, which	The Experiment, which	-2.053	-1.218	-0.384	-2.860	0.017	*
2 v 3	FORMAT	was published on an	was published in an	0.147	0.280	0.414	4.125	0.000	***

Compared priority	Attribute	Comparative	Attribute level	95%-CI lower limit	b-Coefficient	95%-CI upper limit	t	p	
		English-language conference	English-language journal						
2 v 3	IMPORTANCE	extremely irrelevant	relevant	0.921	1.526	2.130	4.949	0.000	***
2 v 3	IMPORTANCE	extremely irrelevant	extremely relevant	0.937	1.167	1.397	9.936	0.000	***
2 v 3	IMPORTANCE	extremely irrelevant	irrelevant	0.846	1.366	1.886	5.150	0.000	***
2 v 3	SAMPLESIZE	small	large	1.383	1.837	2.291	7.933	0.000	***

The attribute that evoked the most interest in looking up the original publication seems to be the judged "importance" for the research area. This impression becomes even clearer when the relative utilities are considered (Figure 4). An interesting observation is that the three most important attributes for the first level of priority were those that were linked to an expert judgment in the mock-up news report (IMPORTANCE, QUALITY, and DESIGN).

Regarding the second level of priority, it is noticeable that with "methodological quality" and "importance of the research area" extraneous judgments also weighed heavily in the decision. "Sample size," which had the least influence on the first level of priority, is also of great importance for the second level of priority, now achieving the second-highest relative utility. It should be noted that the phrase "large sample for the research area" can also be interpreted as an extraneous judgment, just like the other essential characteristics, but this was not our initial intention.

Another issue investigated in the second conjoint step of the study was whether the participants followed a conscious heuristic when conducting the trials. A question was included as a reaction to many participants of the first conjoint step, remarking that they did not read the texts attentively but only paid superficial attention to the highlighted text passages. They further indicated that their decisions had been made based on a subset of these without attention to the overall context. As already mentioned, this comment occurred particularly in combination with complaints about effort and fatigue resulting from conducting the study.

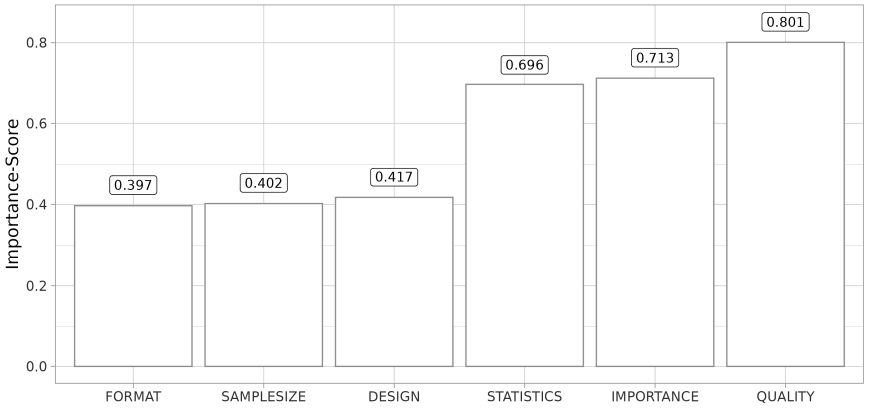

Figure 5: Mean scores of the attribute ranking asking for conscious judgment heuristics. Since the participants had the option to rank a subset of the attributes, the assigned ranks were transformed so that a score of 1 indicates a rank of 1, decreasing by 1/n for each of the n indicated ranks.

When examining the statements on conscious judgment heuristics obtained via a drag-and-drop ranking, it is noticeable that this does not seem to fully coincide with the findings obtained from the logistic regression (Figure 5). Here, the respondents were asked to rank the attributes: "Did you follow conscious decision rules while evaluating the items? If yes, please sort the presented aspects in descending order of importance for your decision."

"Presentation form" was one of the least important attributes in both response formats, and high relevance was attributed to "importance" of the research area. Nevertheless, overall, the structure of the responses is different; for example, "methodological quality" was ranked differently compared to its inclusion in the utility estimate. Still, it can be assumed that the results are due to the overall effects of the mock-up news reports and not to deliberate decision heuristics, leaving aside the possibility that they result from a lack of understanding about the task at hand. This statement leads to the conclusion and limitations in the following section.

4 Discussion and conclusion

We have conducted a conjoint analysis to determine which attributes of scientific publications influence the impact of news media reports on look-up decisions and subsequent citations most. For that, we followed the approach by Anderson et al. (2020) and studied which attributes drive the decision to look up an original research publication mentioned in a popular news report on science. The stimulus design was based on a three-phased Delphi study, resulting in six attributes as the most relevant characteristics for informing a decision to look up an original scientific publication, which were then presented in an online conjoint study. The study's participants – primarily senior scientists who conduct research in economics and business studies – were instructed to rank three mock-up news reports consisting of these six attributes each. This ranking task was repeated 12 times and analyzed using two separate conditional logistic regression models to estimate the utility of each attribute influencing the participants' decision-making.

We have shown clear differences in the utility of attributes used to describe scientific publications in a news report. Also, those attributes drive decision-making and future behavior (to look up the article or not) of researchers to a different extent, although the experts of the Delphi study deemed all attributes relevant for look-up decisions.

Overall, the subjects of the study have relied mainly on attributes based on expert opinion when looking up a publication underlying a news report. An indication of the importance of the publication, followed by statements about the methodological quality and the research design, positively influences decision-making towards looking up the original research publication. The selection decisions seem to be made on a case-by-case basis rather than based on general heuristics the subjects have followed, which is similar to selection processes in science journalism (Chapter 2 in this book; Broer & Pröschel, 2021). Since the population consisted primarily of economic researchers, these findings might be highly skewed. This is especially plausible since Lemke (2020) found press releases to mainly reference medical journals. Htoo and Na (2017) found significant differences in attention across disciplines in various altmetric indicators, including news coverage. An attempt to replicate the findings based on a different population would be desirable.

In addition, the conjoint analysis did not explicitly take into account the expert role of the participants. Although we did not observe one, a bias in the self-understanding of the researchers as experts in their respective fields could still be present. This could lead to answers that could be rational and consistent

or professionally-expected (habituated) explanations for their choices, ultimately leading to a social-desirability bias. However, observing participants' lack of consistent heuristics somewhat contradicts that idea.

Another issue is the possibility of complex interactions in the statistical utilities. Since the chosen models are not appropriate to estimate such interactions and the underlying conjoint model is one of independent attributes (Louviere et al., 2010), these possible complications in the interpretation of the results were not considered. Although contradictory combinations of attributes were omitted in the main study, one could still argue for amplification effects in certain combinations of attribute levels, such as quality and sample size.

Additionally, the presentation as a mock-up news report could further influence these effects. If the way the mock-up was presented primed a certain trust in the "experts" mentioned, for example, the "expert's opinion" could interact with the other attributes in other ways than a simple list of attributes would induce. However, this is also possible for real newspaper reports that are also impacted by the newspapers' or the journalists' perceived prestige, as they may serve as indications for the quality of the journalistic reporting. The decision to present the attributes in a mock-up format was made since the focus group consulted for the relevant attributes reported having rated the items as relevant for reading a publication in general, not based on news media. Therefore, to prevent this non-intended issue in the study, the mock-ups were used as the medium of presentation. To examine this possible caveat, one could reproduce the survey without the sentences or by presenting only one attribute at a time.

Another matter is the statistical model used to analyze the utilities. Including all decisions into one holistic ordinal logistic regression model, as described in Allison and Christakis (1994), would have been preferable. However, since the central assumption of ordinality of the criterion in every predictor was not met, the alternative approach of using two separate models, as described in Aizaki and Nichimura (2008), was taken. The resulting two conditional logit models came with the price of repeated testing and a lack of an estimate of the basic utility differences between the first and second rank priority levels. A different design of the attribute levels could alleviate this issue and make a more comprehensive model possible.

Our study is mainly theoretically rooted in the earmark hypothesis (Kiernan, 2003) and the publicity hypothesis (Phillips et al., 1991), which both try to explain higher citation counts for scientific publications covered in the media. However, to fully understand the intertwined relationship of external and internal science communication and how they affect each other, additional

theoretical considerations should be taken into account and used for further experimentation. For example, an explanation is needed to understand why not every research publication mentioned in the news will also receive more citations. Here, theoretical thoughts such as those stemming from the attention economy by Georg Franck may prove valuable, especially since it already focuses on the realm of science and academic reputation (van Krieken, 2019, p. 4). Attention is recognized as both a scarce resource and as a basic need. At the same time, however, attention generates more attention (van Krieken, 2019, p. 5): a phenomenon also described as "success breeds success" or the "Matthew Effect" (Klamer & van Dalen, 2002).

Furthermore, news values (initially called "news factors"), as proposed by Galtung and Ruge (1965), can have a significant influence on the flow of news. So it is not surprising that, among other factors, the factors "frequency" and "unexpectedness" of news can yield higher mentions in media (Galtung & Ruge, 1965). This can also be viewed closely with the relevance theory proposed by White (2011), who argued that authors usually cite research publications to strengthen their claims and that produce the least cognitive effort while retrieving and evaluating them (Breuer et al., 2022). Both arguments may also be applied to publishers of news reports and journalists when selecting original publications to be reported on. By no means is this selection of theories complete, which highlights the need for further quantitative and qualitative research on the overlapping processes and effects of internal and external science communication.

The attributes presented in our study are solely based on a literature review concerning influences on scientific impact. Furthermore, most of the attributes presented are not regularly reported in news reports. Therefore, although our study was supposed to have high external validity, this artificial limitation could present difficulties in interpreting the results. A large-scale natural language processing-based approach could help to test this caveat and to actualize the list of attributes with as realistic attributes as possible.

Overall, the results are promising, especially regarding the discussion on whether the observed effects of news reports and mentioning of certain publication attributes on look-up decisions (and, perhaps, later citations) are based on the research publications alone or on the additional visibility due to media coverage. Since the most valuable attributes were those with an external judgment implied, the additional, thematically, and methodologically classifying information that can be relayed by a news report could be one of the most significant driving factors concerning the impact advantage. This also highlights the effect

of intermediary institutions, such as Science Media Centers, that provide background and expertise to science journalists (Broer & Pröschel, 2021).

Our exploratory results are particularly critical since they bear room for discussion regarding the role of internal scholarly communication and science journalism and their relationship (Broer & Rotgeri, 2021). One may argue that news media coverage poses a threat to internal science communication and the reputation system of science, which – besides strong critique (Hicks et al., 2015) – still heavily relies on citations to research publications. Since we see tendencies confirming both the earmark and publicity hypotheses (see also Chapter 5 in this book), news media and external science communication can serve as gatekeepers or science influencers, who channel attention towards certain scientific topics, authors, etc. – and along with it, may raise or amplify awareness of scientific publications (Klamer & van Dalen, 2002) reflecting all positive and negative effects associated with the theory of attention economy (van Krieken, 2019). Amongst others, future work should apply a large-scale natural language processing-based approach to examine whether news reports that provide contextual information about ascientific publication come with higher citation rates of the original publication.

5 Acknowledgments

This study is part of the research project MeWiKo funded by the German Federal Ministry of Education and Research (Grant numbers 16PU17018C and 16PU17018A). We also thank two anonymous reviewers for their valuable feedback on the article.

6 References

Aizaki, H., & Nishimura, K. (2008). Design and analysis of choice experiments using r: A brief introduction. *Agricultural Information Research, 17*(2), 86-94. https://doi.org/10.3173/air.17.86

Allison, P.D., & Christakis, N.A. (1994). Logit models for sets of ranked items. *Sociological Methodology, 24*, 199-228. https://doi.org/10.2307/270983

Anderson, P.S., Odom, A.R., Gray, H.M., Jones, J.B., Christensen, W.F., Hollingshead, T., Hadfield, J.G., Evans-Pickett, A., Frost, M., Wilson, C., Davidson, L.E., & Seeley, M.K. (2020). A case study exploring associations between popular media attention of scientific research and scientific citations. *PLOS ONE, 15*(7), e0234912. https://doi.org/10.1371/journal.pone.0234912

Backhaus, K., Erichson, B., & Weiber, R. (2015). Auswahlbasierte Conjoint-Analyse. In K. Backhaus, B. Erichson, & R. Weiber (Eds.), *Fortgeschrittene Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (pp. 175-292). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46087-0_5

Bhandari, M., Busse, J., Devereaux, P.J., Montori, V.M., Swiontkowski, M., Tornetta, P., Einhorn, T.A., Khera, V., & Schemitsch, E.H. (2007). Factors associated with citation rates in the orthopedic literature. *Canadian Journal of Surgery. Journal Canadien De Chirurgie, 50*(2), 119-123.

Bloor, M., Sampson, H., Baker, S., & Dahlgren, K. (2015). Useful but no oracle: Reflections on the use of a delphi group in a multi-methods policy research study. *Qualitative Research*, *15*(1), 57-70. https://doi.org/10.1177/1468794113504103

Bowman, T.D., & Hassan, S.-U. (2019). *Science news and altmetrics: Looking at EurekAlert!* Altmetric Conference 19, AM:6 workshop, Stirling. http://altmetrics.org/wp-content/uploads/2019/10/Bowman_altmetrics19_paper_6.pdf

Breuer, T., Schaer, P., & Tunger, D. (2022). Relevance assessments, bibliometrics, and altmetrics: A quantitative study on Pubmed and arXiv. *Scientometrics*, *127*(5), 2455-2478. https://doi.org/10.1007/s11192-022-04319-4

Broer, I., & Pröschel, L. (2021). *Das Science Media Center Germany: Ethnographische Einblicke in die Arbeitsweisen und Rollen eines Intermediärs zwischen Wissenschaft und Journalismus* (Vol. 57). Verlag Hans-Bredow-Institut. https://doi.org/10.21241/ssoar.73542

Broer, I., & Rotgeri, S. (2021). Unsere Verzerrte Sicht auf den Impact. *DUZ - Magazin Für Wissenschaft Und Gesellschaft*, 61-62.

Burns, T.W., O'Connor, D.J., & Stocklmayer, S.M. (2003). Science communication: A contemporary definition. *Public Understanding of Science, 12*(2), 183-202. https://doi.org/10.1177/09636625030122004

Callaham, M., Wears, R.L., & Weber, E. (2002). Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *JAMA - Journal of the American Medical Association, 287*(21), 2847-2850. https://doi.org/10.1001/jama.287.21.2847

Callaham, M.L., Wears, R.L., Weber, E.J., Barton, C., & Young, G. (1998). Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA - Journal of the American Medical Association, 280*(3), 254-257. https://doi.org/10.1001/jama.280.3.254

Crabtree, C., Holbein, J.B., & Monson, J.Q. (2022). Patient traits shape health-care stakeholders' choices on how to best allocate life-saving care. *Nature Human Behaviour, 6*(2), Article 2. https://doi.org/10.1038/s41562-021-01280-9

Craig, I.D., Plume, A.M., McVeigh, M.E., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact?: A critical review of the literature. *Journal of Informetrics, 1*(3), 239-248. https://doi.org/10.1016/j.joi.2007.04.001

Cronin, B. (1981). The need for a theory of citing. *Journal of Documentation, 37*(1), 16-24. https://doi.org/10.1108/eb026703

Dernbach, B., Kleinert, C., & Münder, H. (2012). *Handbuch Wissenschaftskommunikation*. Springer.

Dumas-Mallet, E., Garenne, A., Boraud, T., & Gonon, F. (2020). Does newspapers coverage influence the citations count of scientific publications? An analysis of biomedical studies. *Scientometrics*, 1-15.

Fanelli, D. (2013). Any publicity is better than none: Newspaper coverage increases citations, in the UK more than in Italy. *Scientometrics, 95*(3), 1167-1177. https://doi.org/10.1007/s11192-012-0925-0

Farshad, M., Sidler, C., & Gerber, C. (2013). Association of scientific and nonscientific factors to citation rates of articles of renowned orthopedic journals. *European Orthopaedics and Traumatology, 4*(3), 125-130. https://doi.org/10.1007/s12570-013-0174-6

Fedorov, V.V. (1972). *Theory of optimal experiments*.

Figg, W.D., Dunn, L., Liewehr, D.J., Steinberg, S.M., Thurman, P.W., Barrett, J.C., & Birkinshaw, J. (2006). Scientific collaboration results in higher citation rates of published articles. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 26*(6), 759-767. https://doi.org/10.1592/phco.26.6.759

Galtung, J., & Ruge, M.H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research, 2*(1), 64-90. https://doi.org/10.1177/002234336500200104

Garfield, E. (1962). Can citation indexing be automated? *Essays of an Information Scientist, 1*, 84-90.

Gordon, T.J. (1994). The delphi method. *Futures Research Methodology, 2*(3), 1-30.

Haustein, S., Bowman, T.D., & Costas, R. (2015). Interpreting "altmetrics": Viewing acts on social media through the lens of citation and social theories. *ArXiv:1502.05701 [Cs]*. http://arxiv.org/abs/1502.05701

Henninger, F., Shevchenko, Y., Mertens, U., Kieslich, P.J., & Hilbig, B.E. (2019). *lab.js: A free, open, online study builder*. PsyArXiv. https://doi.org/10.31234/osf.io/fqr49

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature News, 520*(7548), 429. https://doi.org/10.1038/520429a

Htoo, T.H.H., & Na, J.-C. (2017). Disciplinary differences in altmetrics for social sciences. *Online Information Review, 41*(2), 235-251. https://doi.org/10.1108/OIR-12-2015-0386

Kiernan, V. (2003). Diffusion of news about research. *Science Communication, 25*(1), 3-13. https://doi.org/10.1177/1075547003255297

Klamer, A., & van Dalen, H.P. (2002). Attention and the art of scientific publishing. *Journal of Economic Methodology, 9*(3), 289-315. https://doi.org/10.1080/1350178022000015104

Koguchi, T., & Maeda, S. (2022). The economic value of personal information: Analysis of information leakage incidents. In T. Jitsuzumi & H. Mitomo (Eds.), *Policies and challenges of the broadband ecosystem in Japan* (pp. 213-229). Springer. https://doi.org/10.1007/978-981-16-8004-5_10

Kousha, K., & Thelwall, M. (2019). *An automatic method to identify citations to journals in news stories: A case study of UK newspapers citing Web of Science journals*. https://doi.org/10.2478/jdis-2019-0016

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research, 30*(3), 411-433. https://doi.org/10.1111/j.1468-2958.2004.tb00738.x

Krueger, R.A. (2014). *Focus groups: A practical guide for applied research*. Sage publications.

Kulkarni, A.V., Busse, J.W., & Shams, I. (2007). Characteristics associated with citation rate of the medical literature. *PLOS ONE, 2*(5), e403. https://doi.org/10.1371/journal.pone.0000403

Lemke, S. (2020, December 18). *The effect of press releases on promoted articles' citations and altmetrics*. Metrics 2020: Workshop on Informetric and Scientometric Research (SIG/MET). Zenodo. https://doi.org/10.5281/zenodo.4351360

Lemke, S. (2022). An assessment of impact metrics' potential as research indicators based on their perception, usage, and dependencies from external science communication [Doctoral dissertation, CAU Kiel University]. https://macau.uni-kiel.de/receive/macau_mods_00003250

Lemke, S., Mazarakis, A., & Peters, I. (2021). Conjoint analysis of researchers' hidden preferences for bibliometrics, altmetrics, and usage metrics. *Journal of the Association for Information Science and Technology, 72*(6), 777-792. https://doi.org/10.1002/asi.24445

Lemke, S., Mehrazar, M., Mazarakis, A., & Peters, I. (2019). "When you use social media you are not working": Barriers for the use of metrics in social sciences. *Frontiers in Research Metrics and Analytics, 3*. https://doi.org/10.3389/frma.2018.00039

LimeSurvey Project Team / Carsten Schmitz. (2012). *LimeSurvey: An open source survey tool* [Manual]. http://www.limesurvey.org

Louviere, J.J., Flynn, T.N., & Carson, R.T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling, 3*(3), 57-72. https://doi.org/10.1016/S1755-5345(13)70014-9

Marshall, D., Bridges, J.F.P., Hauber, B., Cameron, R., Donnalley, L., Fyie, K., & Reed Johnson, F. (2010). Conjoint analysis applications in health—How are studies being designed and reported? *The Patient: Patient-Centered Outcomes Research, 3*(4), 249-256. https://doi.org/10.2165/11539650-000000000-00000

MeWiKo. (n.d.). Project website. https://mewiko.de/

Miettunen, J., & Nieminen, P. (2003). The effect of statistical methods and study reporting characteristics on the number of citations: A study of four general psychiatric journals. *Scientometrics, 57*(3), 377-388.

Miettunen, J., Nieminen, P., & Isohanni, M. (2002). Statistical methodology in general psychiatric journals. *Nordic Journal of Psychiatry, 56*(3), 223-228. https://doi.org/10.1080/080394802317607219

Nieri, M., Clauser, C., Franceschi, D., Pagliaro, U., Saletta, D., & Pini-Prato, G. (2007). Randomized clinical trials in implant therapy: Relationships among methodological, statistical, clinical, paratextual features and number of citations. *Clinical Oral Implants Research, 18*(4), 419-431. https://doi.org/10.1111/j.1600-0501.2007.01350.x

Patsopoulos, N.A., Analatos, A.A., & Ioannidis, J.P.A. (2005). Relative citation impact of various study designs in the health sciences. *JAMA - Journal of the American Medical Association, 293*(19), 2362-2366. https://doi.org/10.1001/jama.293.19.2362

Phillips, D.P., Kanter, E.J., Bednarczyk, B., & Tastad, P.L. (1991). Importance of the lay press in the transmission of medical knowledge to the scientific community. *New England Journal of Medicine, 325*(16), 1180-1183. https://doi.org/10.1056/NEJM199110173251620

Pollitt, A., Potoglou, D., Patil, S., Burge, P., Guthrie, S., King, S., Wooding, S., & Grant, J. (2016). Understanding the relative valuation of research impact: A best–worst scaling experiment of the general public and biomedical and health researchers. *BMJ Open, 6*(8), e010916. https://doi.org/10.1136/bmjopen-2015-010916

Robson, C., & McCartan, K. (2016). *Real world research: A resource for users of social research methods in applied settings* (Fourth Edition). Wiley.

Schuir, J., Pöhler, L., & Teuteberg, F. (2022). Zwischen Preisjägern, Datenschützern und Tech-Enthusiasten: Segmentierung des Virtual-Reality-Marktes am Beispiel Oculus. *HMD Praxis der Wirtschaftsinformatik, 59*(1), 261-279. https://doi.org/10.1365/s40702-021-00817-w

Tahamtan, I., & Bornmann, L. (2018). Core elements in the process of citing publications: Conceptual overview of the literature. *Journal of Informetrics, 12*(1), 203-216. https://doi.org/10.1016/j.joi.2018.01.002

Tahamtan, I., & Bornmann, L. (2019). What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018. *Scientometrics, 121*(3), 1635-1684. https://doi.org/10.1007/s11192-019-03243-4

Tahamtan, I., Safipour Afshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics, 107*(3), 1195-1225. https://doi.org/10.1007/s11192-016-1889-2

Tenopir, C., Allard, S., Bates, B.J., Levine, K.J., King, D.W., Birch, B., Mays, R., & Caldwell, C. (2011). Perceived value of scholarly articles. *Learned Publishing, 24*(2), 123-132. https://doi.org/10.1087/20110207

Tenopir, C., Christian, L., & Kaufman, J. (2019). Seeking, reading, and use of scholarly articles: An international study of perceptions and behavior of researchers. *Publications, 7*(1), Article 1. https://doi.org/10.3390/publications7010018

van Krieken, R. (2019). Georg Franck's 'the economy of attention': Mental capitalism and the struggle for attention. *Journal of Sociology, 55*(1), 3-7. https://doi.org/10.1177/1440783318812111

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature News, 512*(7513), 126. https://doi.org/10.1038/512126a

White, H.D. (2011). Relevance theory and citations. *Journal of Pragmatics, 43*(14), 3345-3361. https://doi.org/10.1016/j.pragma.2011.07.005

Willis, D.L., Bahler, C.D., Neuberger, M.M., & Dahm, P. (2011). Predictors of citations in the urological literature. *BJU International, 107*(12), 1876-1880. https://doi.org/10.1111/j.1464-410X.2010.10028.x

7 Appendix

Table: All attributes as extracted for the focus group-style interviews. The attributes removed after the first phase are missing rankings and ratings for the second phase. The papers used to determine the attribute levels after the Delphi pre-study are referenced by DOI or PMID.

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
methodological quality (i.e., RCTs vs. observational study)	Tahamtan et al., 2016, p. 1200	Qualität des Untersuchungsdesigns (z.B.: Metastudie oder Einzelfallbeschreibung)	10.1111/j.1600-0501.2007.01350.x (Nieri et al., 2007); 10.1001/jama.280.3.254 (M. L. Callaham et al., 1998); 10.1001/jama.287.21.2847 (M. Callaham et al., 2002); PMID: 17550715 (Bhandari et al., 2007)	6.00	0.57	7
importance of the research subject	Tahamtan et al., 2016, p. 1198	Wichtigkeit der Forschung nach Meinung/Urteil von Forschenden	10.1001/jama.280.3.254 (M. L. Callaham et al., 1998); 10.1001/jama.287.21.2847 (M. Callaham et al., 2002); 10.1016/j.joi.2007.04.001 (Craig et al., 2007)	4.86	0.42	7
presented on a conference or submitted to a journal	Tahamtan et al., 2016, p. 1207	Journal- oder Konferenzbeitrag	Levels in attribute	4.57	0.39	7

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
study design	Tahamtan et al., 2016, p. 1200	Angabe des Design-Typs unabhängig von der Qualität (Metaanalyse, randomisierte kontrollierte Studie, Beobachtung, Fallstudie,...)	10.1001/jama.280.3.254 (M. L. Callaham et al., 1998); 10.1001/jama.287.21.2847 (M. Callaham et al., 2002); 10.1111/j.1464-410X.2010.10028.x (Willis et al., 2011); 10.1001/jama.293.19.2362 (Patsopoulos et al., 2005); 10.1371/journal.pone.0000403 (Kulkarni et al., 2007); 10.1592/phco.26.6.759 (Figg et al., 2006)	4.86	0.37	6
type of statistical methods	Tahamtan et al., 2016, p. 1200	statistische Methodik (z.B.: verwendete Tests, Vorgehen bei der Datenbereinigung,...)	10.1023/a:1025056718587 (Miettunen & Nieminen, 2003); 10.1080/08039480231760721 9 (Miettunen et al., 2002); 10.1111/j.1600-0501.2007.01350.x (Nieri et al., 2007)	4.57	0.37	5
sample size	Tahamtan et al., 2016, p. 1200	Stichprobengröße	PMID: 17550715 (Bhandari et al., 2007); 10.1007/s12570-013-0174-6	5.00	0.31	6

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
			(Farshad et al., 2013)			
published in a journal with a local/international scope	Tahamtan et al., 2016, p. 1207	lokaler oder globaler Bezug		3.43	0.24	2
language of publication	focus group	Sprache der Publikation		3.71	0.21	0
type of document	Tahamtan et al., 2016, p. 1207	Angabe des Dokumententyps (Review, Paper, Letter to the Editor, ...)		4.43	0.17	4
age of the paper	Tahamtan et al., 2016, p. 1204	Alter der Publikation		4.71	0.17	6
whether others have already cited the paper	Tahamtan & Bornmann, 2018, p. 9	Anzahl der Zitationshäufigkeit der Publikation		3.71	0.14	4
open access status of a journal	focus group	Open-Access-Status des Journals		3.14	0.14	0
journal language	Tahamtan et al., 2016, p. 1207	Sprache des Journals		2.71	0.12	5
significance of results	Tahamtan et al., 2016, p. 1201	Statistische Signifikanz der Ergebnisse		4.43	0.12	3

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
the novelty of the paper	Tahamtan et al., 2016, p. 1199	Kreativität der Forschung nach Meinung/Urteil von Forschenden		2.86	0.10	4
amount of details shared in paper	Tahamtan et al., 2016, p. 1200	Anzahl der Details im Methodenteil		3.71	0.10	2
authors with or without Nobel Prize	Tahamtan et al., 2016, p. 1209	erhaltener Nobel-Preis		3.43	0.09	4
field and subfield of the paper	Tahamtan et al., 2016, p. 1199	Thematische Einordnung der Publikation in ein Untersuchungsfeld		3.29	0.09	4
audiences the document is intended for	Tahamtan & Bornmann, 2018, p. 9	Ausrichtung der Publikationen (populärwissenschaftliche oder wissenschaftliche Zielgruppe)		4.71	0.06	4
accessibility of data used	focus group	Zugänglichkeit der in der Publikation genutzten Daten		3.43	0.06	0
main study conclusion in the title	Tahamtan et al., 2016, p. 1203	Hauptbefund der Studie im Titel		2.86	0.06	4

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
extent the paper has been tweeted	Tahamtan et al., 2016, p. 1205	Anzahl Tweets mit Bezugnahme auf die wissenschaftliche Publikation		2.71	0.04	1
position of the paper in a preprint server	Tahamtan et al., 2016, p. 1205	Existenz eines Preprints (einer Vorabversion)		2.71	0.04	1
open access status	Tahamtan et al., 2016, p. 1205	Open-Access-Status		3.86	0.04	4
number of citations in the first year after publication	Tahamtan et al., 2016, p. 1204	Anzahl Zitationen im ersten Jahr		2.71	0.03	1
open/closed review-process	focus group	Angabe, ob es zu dem Artikel auch ein publiziertes/öffentlich zugängliches Review gibt		2.43	0.03	0
number of pages	Tahamtan et al., 2016, p. 1203	Seiten-/Wortzahl		2.00	0.03	1
authors' reputation	focus group	Reputation der Autor:innen		3.57	0.00	0
multidisciplinary or discipline-specific journal	Tahamtan et al., 2016, p. 1207	spezialisiertes oder multidisziplinäres Journal		3.29	0.00	2

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
reference age	Tahamtan et al., 2016, p. 1203	Alter der Literaturangaben		3.00	0.00	4
funding of the publication	focus group	Finanzierung der Publikation		2.86	0.00	0
Journal Impact Factor	Tahamtan et al., 2016, p. 1206	Journal Impact Factor		2.86	0.00	3
academic age of the author	focus group	Information darüber, wie lange Autor:innen publizieren		2.71	0.00	0
number of references	Tahamtan et al., 2016, p. 1203	Anzahl der Literaturangaben		2.57	0.00	2
number of cooperating organizations among authors	Tahamtan et al., 2016, p. 1210	Anzahl unterschiedlicher beteiligter Einrichtungen an Publikation		2.57	0.00	2
type of funding received	focus group	Art der erhaltenen Zuwendungen		2.57	0.00	0
information about reviewers	focus group	Informationen zu den Reviewer:innen (Disziplin, h-Index, seit wann wissenschaftlich aktiv etc.)		2.57	0.00	0

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
English or non-English journal	Tahamtan et al., 2016, p. 1207	Ist das Journal in englischer Sprache oder nicht		2.43	0.00	5
connections between clusters of co-citations	Tahamtan et al., 2016, p. 1199	Zitationen aus dem Themenbereich der Publikation		2.29	0.00	4
number of self-citations	Tahamtan et al., 2016, p. 1209	Anteil an Selbst-Zitationen		2.14	0.00	5
number of authors	Tahamtan et al., 2016, p. 1208	Anzahl der Autor:innen		2.14	0.00	1
number of previous citations of the author(s)	Tahamtan et al., 2016, p. 1209	bisherige Anzahl an Zitationen der Autor:innen		2.14	0.00	4
authors from (non-) English language institutions	Tahamtan et al., 2016, p. 1212	englischsprachige oder nicht englischsprachige Institution		2.14	0.00	1
existence of non-preprint versions of the publication	focus group	Vorliegen von Versionen der Publikation neben solchen in Preprint-Format		2.14	0.00	0

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
number of papers published on the project	Tahamtan et al., 2016, p. 1212	Anzahl an Publikationen in Bezug auf ein Projekt		2.00	0.00	2
articles published in high impact journals by department members	Tahamtan et al., 2016, p. 1212	durchschnittlicher JIF der Publikationen		2.00	0.00	1
h-Index	Tahamtan et al., 2016, p. 1209	h-Index		2.00	0.00	2
authors listed in ISI Highly Cited	Tahamtan et al., 2016, p. 1209	ISI Highly Cited (Datenbank aus dem Hause Clarivate mit den meistzitierten Wissenschaftlern eines Themenbereichs)		2.00	0.00	3
number of grants received	Tahamtan et al., 2016, p. 1213	Anzahl der Zuwendungen		1.86	0.00	2
number of databases the article is indexed in	Tahamtan et al., 2016, p. 1205	Listung in verschiedenen Datenbanken (WoS, Scopus,...)		1.86	0.00	2

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
report of study design in the title	Tahamtan & Bornmann, 2018, p. 8	Methodik im Titel		1.86	0.00	1
presence of certain trend words in abstract and keywords	Tahamtan et al., 2016, p. 1202	Trendthemen in Schlagworten und Abstract		1.86	0.00	2
journal age	focus group	Alter des Journals		1.71	0.00	0
productivity of department	Tahamtan et al., 2016, p. 1212	Anzahl der Publikationen der Organisation		1.71	0.00	3
academic rank of authors	Tahamtan et al., 2016, p. 1209	Akademischer Rang der Letzt-Autor*in		1.57	0.00	2
presence of certain trend words in abstract and keywords	Tahamtan et al., 2016, p. 1202	Anzahl von Begriffen in Bezug zu Trendthemen im Abstract		1.57	0.00	1
(non-)Asian origin of authors	focus group	Asiatische oder nicht asiatische Abstammung der Autor:Innen		1.57	0.00	0
diversity and number of keywords	Tahamtan et al., 2016, p. 1202	Diversität und Anzahl der Schlagworte		1.57	0.00	2

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
titles in question form or declarative titles	Tahamtan et al., 2016, p. 1202	Frage oder Aussage als Titel		1.57	0.00	1
prestige of references	Tahamtan et al., 2016, p. 1202	Prestige der Literaturangaben		1.57	0.00	3
academic rank of authors	Tahamtan et al., 2016, p. 1209	Akademischer Rang der Erst-Autor*in (z.B.: Professor*in, Assistenzprofessor*in, etc.)		1.43	0.00	2
amount of grants received	Tahamtan et al., 2016, p. 1213	Höhe der Zuwendung		1.43	0.00	1
race of authors	Tahamtan et al., 2016, p. 1211	Ethnie der Autor:innen		1.14	0.00	1
surname of authors	Tahamtan et al., 2016, p. 1213	Vornamen der Autor:innen		1.00	0.00	1
number of images	Tahamtan et al., 2016, p. 1202	Anzahl an Abbildungen				0

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
number of equations	Tahamtan et al., 2016, p. 1201	Anzahl an Formeln in der Publikation				0
number of words in abstract	Tahamtan et al., 2016, p. 1202	Länge des Abstracts				0
presence of appendices	Tahamtan et al., 2016, p. 1202	Länge des Anhangs				0
number of words in the title	Tahamtan et al., 2016, p. 1202	Länge des Titels				0
oral or poster presentation of a paper at a conference	Tahamtan et al., 2016, p. 1207	sind die Journal-Beiträge auf einer zugehörigen Konferenz mündlich oder via Poster präsentiert worden				0
department size	Tahamtan et al., 2016, p. 1212	Anzahl der Mitarbeiter:innen				0
income of the author's country	Tahamtan et al., 2016, p. 1211	Bruttoinlandsprodukt des Landes, in dem sich der Arbeitsort befindet				0

Attribute as reported in the original paper	Origin of attribute	German translation presented to focus group	Identifiers of publications used for level determination	Rating Delphi phase 2	Ranking Delphi phase 2	'Relevant'-votes Delphi phase 1
gender of authors	Tahamtan et al., 2016, p. 1211	Geschlecht der Autor:innen				0
university rank	Tahamtan et al., 2016, p. 1212	Platz in Universitätsranking				0