EMPIRICAL RESEARCH

Which Factors Affect the Scientific Impact of Review Papers in IS Research? A Scientometric Study

ARTICLE HISTORY
Compiled December 24, 2018

ABSTRACT
Review papers provide a foundation for knowledge development in information systems (IS) as well as in any other scientific discipline. While some of the prominent reviews in information systems are cited more than twice a day on average, others take years to accumulate single digit citations. The magnitude of these differences and the proliferation of review papers in recent years prompt us to empirically analyze what distinguishes those reviews that have proven to be integral to scientific progress from those that might not be considered impactful. Our results demonstrate that the attributes explaining scientific impact are unique for the different types of reviews: reviews for describing, understanding, explaining, and theory testing. Transparency of the applied methodology is important for reviews that target theory testing, understanding, or explaining; similarly, reviews for describing, understanding or explaining achieve a higher impact when they develop a research agenda. By providing nuanced insights into the attributes of review papers that are valued by subsequent research, our study contributes to the vibrant discourse on literature reviews in IS. We thereby inform the different stakeholders involved in the development and publication of review papers in the IS field.

KEYWORDS
Review papers; scientometric; scientific impact; citation analysis

1. Introduction

Review papers are a fundamental genre in every scientific discipline. In fact, "many of our greatest scientists have used, created, and contributed to the review literature" (Garfield, 1987, p.113). Especially in recent years, a vibrant discourse on paper genres in general and review papers in particular has started in IS research (cf. Boell & Cecez-Kecmanovic, 2015; Paré, Tate, Johnstone, & Kitsiou, 2016; Templier & Paré, 2018), rendering IS a pioneer discipline to contribute to this important conversation in social science research. As in most scientific disciplines, review papers in the information systems (IS) discipline provide a foundation for scientific progress (Webster & Watson, 2002), and the impact of this genre is manifest. With very few exceptions, every major IS journal accepts review papers, often as a separate genre. In addition, there are editorial initiatives to facilitate the publication of reviews in some of the field's most renowned journals, including *MIS Quarterly*, the *Journal of the Association for Information Systems*, the *European Journal of Information Systems*, and the *Journal of Information Technology*. Over the last 15 years, more than 200 review papers have been published in IS and, similar to other disciplines (Hoffman & Holbrook, 1993; Peters & van Raan, 1994), this genre is remarkably impactful in IS.

The variance that exists between reviews that achieve an outstanding impact (e.g., Alavi & Leidner, 2001) and those that receive limited attention poses a challenge not only for effective knowledge development in the IS discipline. It also puts prospective authors at risk of investing too much time, developing minimal impact and receiving almost no recognition. Furthermore, while prospective authors who aspire to make an impact with their reviews can only draw upon crude journal impact factors to assess the potential impact of their specific review paper, there is a lack of empirical insights into which attributes affect the scientific impact of reviews. We suggest that a systematic analysis contributes to our understanding of the factors that drive the scientific impact of IS review papers and therefore address the following question:

What are the attributes that affect the scientific impact of review papers?

To answer this question, we conducted a scientometric study of 220 reviews which have been published in the top-40 IS journals (Lowry et al., 2013) between 2000 and 2014. In developing our model, we focus on content-related attributes of a review, as opposed to more superficial information, or meta-data. Specifically, we explore the effect of two main variables, transparency and research agenda, which affect the scientific impact of review papers after controlling for effects related to the journal, the authors and the topic of the review. While methodologists and editors emphasize methodological rigor and the development of a research agenda as important qualities of review papers (e.g., Paré et al., 2016; Rivard, 2014; Webster & Watson, 2002), there is a lack of empirical insights into whether they actually lead to a higher scientific impact, i.e., whether subsequent research values these qualities.

Our findings contribute to the discourse and literature on review papers in IS and to the scientometric literature in general. By providing empirical evidence on which attributes influence scientific impact, as measured in terms of citations, our insights are useful for substantiating guidelines and methodological recommendations. More generally, we advance our scientometric understanding of the impact of standalone reviews as a fundamental research genre — an important task in itself (Judge, Cable, Colbert, & Rynes, 2007), which has been addressed in many top-journals in other business and management disciplines (e.g., Bergh, Perry, & Hanke, 2006; Colquitt & Zapata-Phelan, 2007; Stremersch, Verniers, & Verhoef, 2007), but neglected in IS. We expect that our study contributes to the discourse on literature reviews and informs the different stakeholders involved in the development and publication of review papers, including authors, methodologists, reviewers, and editors.

The remainder of this paper is organized as follows. Based on extant work from the methodological literature on reviews and the established research stream of scientometrics, in Section 2 we develop a model consisting of attributes relating to the paper, the author, and the journal that can be expected to affect scientific impact. In Section 3, we present the dataset. Next, in Section 4 we evaluate the attributes empirically and show that their effects on scientific impact are robust. We discuss the implications of our results for authors of review papers who strive to maximize their impact on subsequent research in Section 5 and conclude with an outlook on how the IS discipline could foster the impact of reviews in Section 6.

2. Model Development

There are different attributes on the paper, author, and journal level that can be expected to affect the scientific impact of different types of review papers. The development of the model, which is structured according to these levels, is informed by both scientometric literature, which analyses the impact and diffusion of knowledge contributions within the academic discourse (e.g., Grover, Raman, & Stubblefield, 2013; Hansen, Lyytinen, & Markus, 2006; Hyland, 1999; Jackson & Rushton, 1987), and the literature on reviews, which includes methodological and editorial papers discussing qualities of impactful review papers (e.g., Paré et al., 2016; Rivard, 2014; Rowe, 2014; Webster & Watson, 2002). While there are multiple possible attributes on each level (e.g., Mingers & Xu, 2010; Tahamtan, Afshar, & Ahamdzadeh, 2016), we deliberately develop a parsimonious model by selecting one control variable for each level, and two variables that are specific to review papers, i.e., transparency and the development of a research agenda. In selecting the attributes specific to review papers, we focus on those that are related to the content of the review and not to its meta-data, such as the length of the title and the number of keywords. At the same time, and to avoid confounding effects related to differences in visibility and reputation, control variables need to be included at each level. Furthermore, we consider attributes that are not limited to particular types of review papers. An overview of attributes examined in previous scientometric studies is provided in Appendix A. The most relevant alternative attributes will be revisited in the results section as part of the robustness checks. Figure 1 provides an overview of the research model, which covers different types of reviews and is structured according to the paper, the author, and the journal level, respectively.

2.1. Attributes at the Paper Level

At the paper level, methodological transparency and the development of a research agenda can be expected to lead to higher scientific impact. As the scientometric literature is remarkably silent with regard to these content-related attributes that predict the impact of review papers, they are primarily based on the discourse on literature reviews. As a further attribute, we include the popularity of the topic to control for general variations in potential readership.

We consider methodological rigor as one of the most important attributes affecting the scientific impact of review papers. Assessing the reliability of the knowledge contributions when citing a review, authors' decisions to cite are influenced by the degree to which the transparency of the review's reporting practices signals a systematic methodological approach. A systematic methodology, which depends on the specific review type (Paré et al., 2016; Templier & Paré, 2018)[1], is considered as the basis of reliability, validity and trustworthiness of a paper (Grover et al., 2013; Judge et al., 2007; Templier & Paré, 2018; vom Brocke et al., 2015). Although transparency and systematicity are twin concepts (Paré et al., 2016), systematicity of a review cannot be assessed directly. Instead, readers may perceive high transparency as signaling high systematicity. For example, they may implicitly or explicitly consider detailed methodological guidelines (cf. Templier & Paré, 2018) or more general recommendations on reviews (e.g., Rowe, 2014; Webster & Watson, 2002) and recognize familiar methodological items that have been reported or, similarly important, those that have been omitted. For reviews that are not transparent, readers may be unable to evaluate their systematicity. Transparent reviews, on the other hand, can only be unsystematic if reviewers and editors have failed to require corresponding changes. In extant scientometric research, transparency, which is perceived by the readers, has been found to be associated with the number of citations a paper attracts (Bergh et al., 2006; Grover et al., 2013; Judge et al., 2007; Montori, Wilczynski, Morgan, & Haynes, 2003).

Developing a research agenda has been suggested as an important contribution of review papers (e.g., Rowe, 2014; Webster & Watson, 2002). However, this attribute can rarely be found in scientometric analyses. In fact, the only scientometric study we are aware of is a survey that dates back more than two decades and its results on providing value for future research by developing a research agenda have largely been ignored in subsequent scientometric research (Sternberg & Gordeeva, 1996). In this paper, the authors have surveyed researchers to identify attributes that make papers influential, one of these attributes being value for future research. Except for theoretical significance, this attribute outranked all other factors in terms of importance, including substantive interest, methodological interest, practical significance and quality of presentation. Drawing on a comprehensive overview of extant research, literature reviews are in a position to make well-grounded recommendations on promising research gaps, thereby helping others to avoid reinventing the wheel (Zorn & Campbell, 2006). Authors who identify research gaps can either highlight white spots in the research landscape, challenge existing knowledge and its underlying assumptions (Alvesson & Sandberg, 2011; Sandberg & Alvesson, 2011), or outline which gaps are unlikely to be addressed successfully (Lacity, Solomon, Yan, & Willcocks, 2011). Authors of review papers can even go one step further and develop a structured research agenda that describes how research gaps should be closed. By including recommendations for further research as an attribute in our main model, our study provides empirical insights into whether efforts to pave the path for future researchers actually translates into scientific impact. Specifically, citations may indicate whether review papers that identify research gaps in the literature and develop research agendas, are successful in stimulating subsequent research that may follow these roadmaps and investigate unexplored avenues.

[1] Note that systematicity is a concept that applies to all types of reviews, not only the qualitative systematic review (Paré et al., 2016) that has been discussed critically (Boell & Cecez-Kecmanovic, 2015).

The topic of a review paper is a further scientometric attribute that is likely to correlate with its scientific impact. Taking into consideration that decisions to cite tend to be premised on an appropriate thematic fit with the manuscripts of citing authors, reviews addressing popular topics have more opportunities to be cited. This has been confirmed by scientometric research, which has not only uncovered that citation practices differ between disciplines (Braun, Glänzel, & Grupp, 1995; Hurt, 1987) but also that citation rates vary between different topics within the same discipline (Garfield, 2006).

2.2. Attributes at the Author Level

With regard to the author level, there are several attributes which have been suggested to influence the impact of a scientific paper. While some author attributes should be unrelated to the contribution of a paper (e.g., gender, nationality, and social status), other attributes may be considered indicators of the authors' impact, or a reputation for strong contributions to research. Citing a paper due to author-related attributes is commonly associated with particularistic citing behavior, while citing a paper for its scientific merits is associated with universalistic citing behavior (Baldi, 1998; Judge et al., 2007; Merton, 1973). A further dimension refers to teamwork between and beyond the authors, with author teams increasingly outperforming individual authors in terms of productivity and impact while sharing the credit for their work (Bikard, Murray, & Gans, 2015). Furthermore, authors' interactions with the research community, such as presenting early work and soliciting feedback, have been found to impact their research (Oettl, 2012).

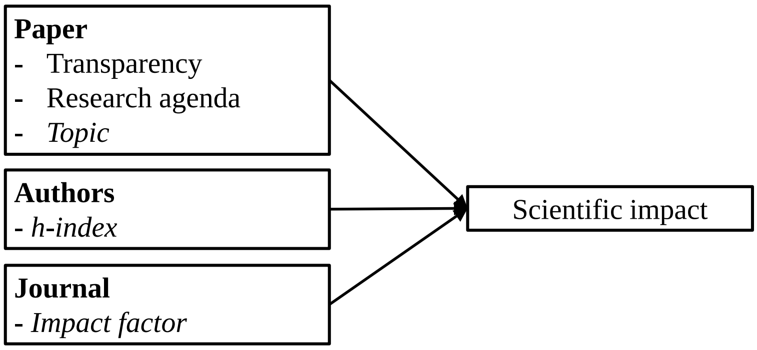
The category of author attributes most commonly used in the scientometric literature comprises indicators of an author's visibility (Judge et al., 2007; Peters & van Raan, 1994). Several indices have been developed to measure the impact of an author's publication record, such as the Hirsch, g, and h(2) indices (Grover et al., 2013). Other attributes, such as academic reputation and an author's affiliation, appear to be less trivial predictors of a papers' impact, but they tend to be correlated with an author's impact. This is due to the consideration of publication records and citation impact when academic reputation is evaluated and when tenure and promotion decisions for top-tier institutions are made.

2.3. Attributes at the Journal Level

Scientometric studies have found the publication outlet to be among the strongest predictors of the number of citations a paper receives (Judge et al., 2007; Leimu & Koricheva, 2005; Mingers & Xu, 2010; Peters & van Raan, 1994), regardless of its genre. Attributes at the journal level (e.g., visibility, access, reputation, and circulation) are significantly correlated with the scientific impact of their papers in IS (Grover et al., 2013), management (Judge et al., 2007), economics (Ellison, 2002), operations research (Mingers & Xu, 2010), psychology (Starbuck, 2005), and the health sciences (Patsopoulos, Analatos, & Ioannidis, 2005). As many of these attributes are interrelated, we follow a common practice in scientometric studies by adopting the journal impact factor as a proxy variable.

2.4. Model

Figure 1 illustrates our research model, which is structured according to three levels, namely the paper, the author, and the journal level. The same model applies to reviews pursuing different goals (Rowe, 2014), i.e., for reviews whose goal is to describe (narrative and descriptive reviews), understand (scoping and critical reviews), explain (theory development and realist reviews), or test (meta-analysis, qualitative systematic reviews and umbrella reviews). With this general model in mind, we explore empirical dependencies between the attributes and the different types of reviews in Section 4.



Note. The same model applies to reviews pursuing different goals (describing, understanding, explaining and theory testing). Control variables are in italics.

Figure 1. Model: Scientific Impact of Review Papers

3. Data

3.1. Sample of Review Papers

We collected an exhaustive set of review papers published between 2000 and 2014[2]. In agreement with extant definitions of literature reviews (e.g., Blumberg, Cooper, & Schindler, 2005; Hart, 1998; Levy & Ellis, 2006; Schwarz, Mehta, Johnson, & Chin, 2007), we include papers that *provide a synthesis of the body of knowledge of a specified domain.* As outlined in the following, we thereby exclude papers that collect primary data, or focus on questions of research methodology, as opposed to domain knowledge, for example. In contrast to other scientometric studies (e.g., Grover et al., 2013; Sidorova, Evangelopoulos, Valacich, & Ramakrishnan, 2008; Tams & Grover, 2010), we go beyond a few top-tier journals and focus on a broad set of 40 IS journals, which was identified by Lowry et al. (2013). These journals were identified based on expert judgment and impact factors. They include the AIS senior scholars' basket of journals; in addition the set primarily comprises journals from the IS discipline. To eliminate language-related effects, we exclusively focus on reviews published in English. We acknowledge that this is likely to increase homogeneity related to author nationalities and regions in our sample. Reviews were identified by scanning the tables of contents from each journal (approx. 17,500 papers). We then compiled a preliminary list of 470 candidates. From this list, we excluded papers which do not comply with our definition of review papers. Specifically, we excluded 70 candidates which do not provide a synthesis, 24 candidates which are short research commentaries, 28 candidates which collect primary data, 93 candidates which do not focus on domain knowledge, 8 candidates which do not focus on the academic literature, 9 candidates that focus on the history of a journal, 2 editorials, and 1 paper developing an artifact. In addition, three of the authors coded the type of review based on Rowe (2014), resulting in 74 reviews for describing, 48 reviews for understanding, 65 reviews for explaining, and 33 reviews for theory testing. After this coding process, we dropped 15 hybrid reviews that cannot be assigned to a unique goal, resulting in a final sample size of 220. Initial disagreements were discussed and reconciled during team meetings.

---

[2]By selecting the year 2000 as a starting point, our sample includes all reviews published in the prominent *Theory and Review* category of *MIS Quarterly*. For the most recent reviews published after 2014, three-year citation rates are not yet available.

3.2. Measures

Table 1 provides an overview of the measures used in our empirical study to operationalize the attributes of review papers. Descriptive statistics, including correlations, are provided in Appendix A. We measure the dependent variable using citation rates as commonly suggested in the literature (Grover et al., 2013; Judge et al., 2007; Tams & Grover, 2010). Citation data was extracted from Google Scholar[3] on February 20th, 2018. Self-citations were excluded because they do not represent real knowledge flow (Singh, 2005). We measure scientific impact in terms of three year citation rates and implement robustness checks to analyze the degree to which they correlate with long-term impact. To avoid possible measurement biases, we have to ensure that the dependent variable is measured after the same amount of time has elapsed since the publication of the reviews, i.e., we have to distinguish whether a review was published early or late in a certain year. Although databases such as Google Scholar and Web of Science only provide citation data on an annual, as opposed to a monthly or daily basis, we correct for the month of publication by adjusting the dependent variable proportionally. One alternative to the date of publication would be the date of (advanced) online publication. This date, however, is not available for more than 30% of the reviews, for 40% of the reviews for which the date of online publication is available, it is after the actual/official date of publication (in print, if applicable), and as our dataset contains some reviews published in the early 2000s, availability of a website for the journals would introduce confounding effects. We therefore consistently measured citations starting with the year of publication and we excluded citations in the years before the review was published in print.

Quality attributes of review papers include transparency and the development of a research agenda. Concerning transparency, we calculated a score that measures the percentage of items that were reported relative to the items required for each type of review. This approach is similar to previous studies (e.g., Aytug, Rothstein, Zhou, & Kern, 2012; Bergh et al., 2006). We coded whether items pertaining on six methodological steps as presented by Templier and Paré (2018) are reported. These methodological steps are: (1) developing the review plan, (2) searching the literature, (3) selecting studies, (4) assessing the quality of included studies, (5) extracting data, and (6) synthesizing. For each step, we coded a set of items that are required for the specific type of review. A detailed overview of the items for each type of review is provided in Appendix B. The first author extended the transparency coding of Templier and Paré (2018) to the whole set of reviews included in this study. To become familiar with the coding procedures, a random overlapping sample of 30 reviews was coded. A high inter-coder agreement was achieved and disagreements were resolved in a discussion between the authors. The remaining review papers were coded by the first author, and borderline cases were discussed by the authors until consensus was reached.

Concerning the development of a research agenda, we coded three possible levels. If the review briefly mentions topics that would benefit from future research, we coded *none*. If the review provides more specific starting points for subsequent studies by identifying research gaps, we coded *partial*. Exhaustive research agendas that are consistent with the recommendations of well-known editorials (e.g., Rowe, 2014; Webster & Watson, 2002) were coded as *complete*. Throughout the coding, higher levels were coded when the space for describing the research agenda was more substantial (e.g., number of pages, table summaries), when the relative importance in the review was more prominent (e.g., 1st vs. 2nd level sections, mentions in the abstract, mentions as a key contribution of the review), and when the guidance was more specific and actionable (e.g., description of methodological approaches, unambiguous recommendations on how to address the research gaps).

We specify control variables for the popularity of the topic, the visibility of the review's authors, and the journal. The topic refers to the average (citation) impact of other papers addressing the same topic as the review paper. Following Bergh et al. (2006), we measured

[3]Web of Science, as an alternative literature database, covers only 23 journals from our scope of 40 journals, e.g., it does not provide citation data for at least 39 review papers (not counting embargo years, such as for the Journal of the Association for Information Systems).

Table 1. Measurement: Attributes of Review Papers

Attribute	Measurement	Key references
Dependent Variable		
Scientific impact	Number of citations after three years, corrected proportionally for the month of publication (extracted from GoogleScholar on 20th of February, 2018)	Judge et al. (2007) Grover et al. (2013)
Paper Level		
Transparency	A score indicating the percentage of items that are reported transparently (relative to the items required to be reported by the review type). The items are structured according to six methodological steps: (1) developing the review plan, (2) searching the literature, (3) selecting studies, (4) assessing the quality of included studies, (5) extracting data, and (6) synthesizing (details are provided in Appendix B)	Templier and Paré (2018) Paré et al. (2016) Okoli (2015) Bergh et al. (2006)
Research agenda	Dummy variables: *None* if no guidance for future research is provided, *Partial* if the review identifies some research gaps, *Complete* if the review provides a detailed research agenda	Rowe (2014) Webster and Watson (2002) Te'eni, Rowe, Ågerfalk, and Lee (2015) Sternberg and Gordeeva (1996)
Topic	Average impact of papers addressing similar research topics (papers addressing the same topic were identified based on overlapping keywords)	Garfield (2006) Bergh et al. (2006)
Author Level		
h-index	Average of the h-indices of authors at the time when the review was published (calculated based on publication lists provided by Scopus)	Hirsch (2005)
Journal Level		
Journal impact factor	Journal impact factor provided by Thomson Reuters[a]	Judge et al. (2007) Mingers and Xu (2010)

Notes. Control variables are in italics. [a] Imputation of missing values based on the average impact factor of same-tier journals according to VHB-JOURQUAL3 (available at http://vhbonline.org/VHB4you/jourqual/vhb-jourqual-3/teilrating-wi/).

topic impact as the average citation rate of other IS papers which examine similar research topics. Research topics were considered similar between a focal review and other papers (using the same scope regarding journals and time, i.e., the Top-40 Journals between 2000 and 2015) if they share at least one keyword. The citation rates per research topic were then averaged using Scopus citation data (excluding the citations of the focal review).

We measured author impact as the average h-index of the author team. This index has shaped the public perception of prominent authors in academia and it is based on both, productivity and impact. Specifically, an author with an index h has published h papers each of which has been cited at least h times (Hirsch, 2005). We therefore controlled for the average of the h-indices of the authors at the time when the review was published. The indices were determined based on the lists of publications of each author, as provided by Scopus.

As we cover a broad scope of journals, controlling for effects related to differences in visibility or circulation of journals was essential. Consistent with other scientometric studies covering many journals, we did not control for single journals but measures of journal impact (Judge et al., 2007; Mingers & Xu, 2010). This allowed us to control journal-related effects while at the same time avoiding model overfitting. We used the journal impact factors provided by Thomson Reuters as a measure to control journal-related effects.

4. Analyses and Empirical Results

4.1. Analyses

Consistent with previous scientometric research in IS (Loebbecke, Huyskens, & Berthod, 2007), we observe that it is only a small number of reviews that drive the aggregated impact while many reviews receive low single-digit or no citations with three year citation rates varying between 0 and 319 citations per year (median: 25, mean: 42, std. error: 51). We examined the attributes that explain the differences in citations of review papers by drawing on a generalized linear model (GLM) with a Poisson link function, which is appropriate for dependent variables that are skewedly distributed count data. To analyze the different effects of the variables, we initially ran regressions using the following equation as the control model:

$$Citations = \beta_0 + \beta_1 JournalImpactFactor + \beta_2 h\text{-}index + \beta_3 Topic + \epsilon.$$

Table 1 shows how the variables were measured and Appendix C provides descriptive statistics. By standardizing regression coefficients, we removed different units of measurement and determined the effects of standard deviation changes of the attributes on the dependent variable.

In the next step, we included indicator variables for the four types of review to capture differences in impact between reviews for describing, understanding, explaining, and theory testing. However, the new indicator variable for the type of review strongly correlates with the other variables of our model (cf. Appendix C), prohibiting us from pooling the different types of reviews and testing the effects of our main variables for the whole sample of reviews. For example, reviews for theory testing are naturally published on popular as opposed to emergent topics for which there is no established theory and a paucity of empirical research. Furthermore, these reviews tend to score highly with regard to transparency, but they rarely identify research gaps and develop research agendas. Concerning the same attributes, reviews for understanding are the complete opposite (cf. Appendix C). On the one hand, this indicates that the attributes explaining the impact of the four types of reviews are different. On the other hand, this issue needs to be addressed in our empirical strategy because these dependencies would potentially bias the confidence intervals of our estimates.

As these dependencies between variables prohibit us from estimating a single model for all review types, we split our sample according to the review type and conducted separate

Table 2. Results of a GLM Predicting Citations to Different Types of Reviews After 3 Years

Effect	Describing (I) (n=74)		Understanding (II) (n=48)		Explaining (III) (n=65)		Testing (IV) (n=33)	
	Control	Main	Control	Main	Control	Main	Control	Main
Journal Impact Factor	0.57**	0.57**	0.38**	0.35**	0.28**	0.19**	0.22**	0.01
h-index (average)	0.27**	0.27**	0.39**	0.45**	0.04	−0.05	0.15**	0.13**
Topic	0.01	−0.07	0.08*	0.00	0.28**	0.22**	0.14**	0.18**
Transparency (score)		0.10**		0.23**		0.26**		0.53**
Research Agenda (none)		−0.27**		[a]		−0.55**		[a]
Research Agenda (complete)		0.30**		0.51**		0.13*		[a]
R^2	0.29	0.32	0.46	0.48	0.30	0.41	0.37	0.47
ΔR^2		0.03		0.02		0.11		0.10

Notes. DV: citations. Model includes an intercept. Standardized beta coefficients are reported.
*significant at 0.01, **significant at 0.001. [a] Not enough observations available to include the variable.

analyses on each subset of review papers[4]. This step reduced our sample sizes considerably (cf. Table 2), resulting in lower test-power, in particular for theory testing reviews. As we analyzed all reviews in a broad scope of journals spanning 15 years, our options to extend the sample and increase test-power accordingly were limited. While low test-power poses the problem of a higher probability of missing effects that are actually significant, it can be acceptable when effects are found to be significant and other threats to validity are addressed (Cohen, 1988). Most importantly, low test power makes it necessary to check whether the observed effects arise by chance, as higher deviations from true effects are more likely in small samples. In the second part of the results section, we checked whether the results are robust regarding outliers and other variations in the sample. For the control model, the estimation results on each subset are provided in Table 2.

In the next step, we included the main variables and specified the following main model:

$$\begin{aligned} Citations = \beta_0 + \beta_1 JournalImpact + \beta_2 h\text{-}index + \beta_3 Topic + \beta_4 Transparency + \\ \beta_5 ResearchAgenda\_None + \beta_6 ResearchAgenda\_Complete + \epsilon. \end{aligned}$$

Transparency is the score of methodological items that are reported transparently. *Research Agenda_Complete* and *Research Agenda_None* are dummy variables indicating whether the review includes a complete research agenda or not. Reviews providing a *partial* research agenda serve as a natural reference group because they represent the common case of brief discussions of implications for future research, which is expected from most papers, including reviews. In two subsets, the research agenda variable did not have enough variance (not enough or too many observations) to be included in the analyses. For example, there are too few reviews for theory testing that propose a research agenda and there are too few reviews for understanding that omit a research agenda. These variables are not included in the corresponding result tables.

Table 2 displays the estimation results for the main model. Overall, there are different attributes that explain the variance in the scientific impact of review papers. In fact, the single best predictors vary throughout the subsets, reflecting substantial differences across the review types. This suggests that the way in which subsequent research perceives, evaluates, and cites reviews is contingent on the nature of the review. While some reviews are primarily valued for their rigorous application and reporting of methodologies, other reviews are valued for the usefulness of their research agendas. Our discussion of the results considers these differences in how other researchers evaluate the review types both directly and indirectly, i.e., by assessing other aspects of a review that may indicate qualities that are harder to assess. We discuss each variable in turn.

The journal impact factor is both a control variable and a possible indicator of the quality of papers. Although its measurement is similar to the topic variable, the observed correlations between these variables are only moderate and multicollinearity is not an issue with variance

[4]Another option would be to address this issue using interaction effects between the variables that are correlated. As Table 2 suggests, however, this would result in a complex model in which most variables interact with the type of review.

inflation factors (VIF) below a threshold of 2. The results show that the journal impact factor is the most important attribute for reviews aimed at describing. This suggests that citing authors consider the journals' reputation for selecting high quality reviews. The contributions of reviews for understanding, for which objective quality criteria are rare, are similarly difficult to evaluate, thus explaining the high effect of the journal impact factor. For reviews that aim to explain phenomena, the journal impact still has an effect, though not a predominant one. Concerning reviews that test a theory, there is no evidence that citing decisions are associated with a journals' reputation. For this type of review, it can be expected that citations are largely based on objective criteria that are applied to the reporting practices as suggested by methodologists. This low and insignificant effect of the journal impact factor contrasts with many other scientometric studies (e.g., Mingers & Xu, 2010). Although there are moderate correlations between transparent reporting practices and the journal impact, the change in the journal impact factor's coefficient from the control model to the main model suggests that the variance is explained by transparency and not by the journal impact.

Our model also controls for the authors' reputation. Taking into account the results of the robustness checks presented in the following subsections, author reputation has a positive impact on citation scores of reviews for describing, understanding, and theory testing. For reviews for explaining, we observe no or slightly negative effects. The question of "Who has published the review?" and a corresponding recognition of the ownership of ideas presented in the review is obviously important for reviews for describing and understanding, whose quality is difficult to assess objectively (cf. Hyland, 1999). For reviews aimed at theory testing, author impact correlates with the impact of the review. In this case, we are careful to speculate on underlying causalities associated with this coefficient, because the challenging methodologies associated with theory testing (e.g., meta-analysis) may be applied more often by experienced author teams (not necessarily captured by the h-indices of authors).

The topic variable controls for the popularity of different topics addressed by the reviews. This variable suggests that different types of reviews may have a higher impact when their timing with the popularity of the topic provides a good fit. While reviews for describing tend to exert a higher impact when published on emerging topics, theory testing reviews evidently require topics for which more empirical research is available. In this regard, our results contrast with Hwang (1996), who argues that meta-analyses may even be useful when research topics are still in exploratory phases. Reviews for understanding are not dependent on the popularity of the topic and may be published in its emergent or latter stages. Interestingly, reviews for explaining have the highest impact when published on established as opposed to emergent topics. This indicates that premature theorizing may not be valued by subsequent research.

The transparency variable measures the degree of transparent reporting practices with regard to methodological procedures; this is a quality that is considered critical by methodologists (Templier & Paré, 2018). Complementing methodological guidelines describing the levels of transparency associated with different types of reviews (Paré et al., 2016), our study shows to which extent subsequent research considers transparency when using and citing different types of reviews. Above all, we show that transparency is the most important attribute for theory testing reviews. Throughout our analyses, it is the single best predictor of impact with an effect size that is twice as high as the second best predictor. This underlines the importance of transparent and systematic reporting and suggests that these practices are valued by subsequent research. Interestingly, reviews aimed at understanding and explaining are cited more often when the methodology is reported in a transparent manner. This suggests that contributions to understanding and explaining phenomena, which often result in theoretical models, benefit from a rigorous methodology that supports the proposed explanation by grounding it in extant literature. In contrast, transparency is a less important attribute for reviews aimed at describing. This is consistent with the notion that this type of reviews aims at achieving useful problem shifts instead of aggregating empirical evidence in a reliable way. Subsequent research may perceive the goals of these reviews to be achieved by creative approaches to the topic, and not by the rigorous application of methodologies.

Finally, as several editors have suggested (e.g., Webster & Watson, 2002), developing a research agenda is important. Reviews for theory testing are the only exception in our analysis as our sample does not contain enough observations to estimate the effect for this particular subset. Although this should not discourage authors of reviews for theory testing to speculate on fruitful paths for future research, it may not always be possible to provide a research agenda in this type of review. For reviews aimed at understanding, in contrast, the development of a comprehensive research agenda is the single best predictor of citation impact. These reviews seem to be more useful for subsequent research when the constructive problem shifts achieved by the understanding process are complemented by a research agenda specifying the implications for future research. For reviews aimed at describing and explaining, presenting the review without any implications for future research is associated with a significant decrease in terms of impact. In contrast, going beyond the presentation of a few open questions and developing a comprehensive agenda leads to higher impact.

In summary, our results provide a parsimonious and powerful explanation (in terms of ΔR^2) for the impact of four types of reviews. In contrast to previous scientometric research, in which a majority of the studies have identified journal impact as the single best predictor, we demonstrate the importance of including a transparent methodology as well as a research agenda. Although requiring considerable coding efforts, omitting these variables that have strong and significant effects on scientific impact poses a threat to the validity of scientometric models. The results also provide unprecedented evidence, showing that the most important attributes are determined by the type of review, and that the genre of literature reviews should not be considered as a monolithic block in this regard. As the robustness checks presented in the following subsection show, the model provides a robust and holistic explanation for the scientific impact of IS review papers. With its relatively high explanatory power, which exceeds many scientometric studies, it also provides a basis for tentative predictions of the impact of papers in the review genre.

4.2. Robustness Checks

There are several alternative attributes suggested in prior research that may bias our results (cf. Appendix A). Although our sample limits our ability to include further variables in the main model, we can check the robustness of the selected variables with regard to alternative effects. We therefore estimated six models that include various alternative variables and analyzed changes in effect size and significance of the main variables. Table 3 provides an overview of the robustness checks, their underlying rationale, and the corresponding models. We discuss the robustness checks relating to the journal, the authors, and the paper level in turn. On the paper level, transparency does not have a robust effect on citations of reviews for describing. For the sake of clarity, we do not repeat this result and discuss it at the end of this Subsection.

Table 3. Summary of Robustness Checks

No.	Robustness Check	Rationale	Model
1	Include special issue indicator.	Reviews published in special issues have a higher visibility (Bergh et al., 2006; Mingers & Xu, 2010).	(1)
2	Include accessibility.	Subscription access control (open access) could bias other coefficients (Björk & Solomon, 2012).	-
3	Include different measures of author visibility.	The average h-index of authors might not be a perfect measure for the visibility of the author team (Judge et al., 2007; Tams & Grover, 2010).	(2)
4	Include the number of authors.	Results might be biased due to the effects of teamwork (Fortunato et al., 2018; Wuchty, Jones, & Uzzi, 2007).	(3)
5	Include acknowledgment of conceptual feedback.	Results might be biased due to the effects of external feedback on paper impact (Oettl, 2012).	(4)
6	Include novelty of the review.	Results might be biased because novel reviews receive more attention (Uzzi, Mukherjee, Stringer, & Jones, 2013).	(5)
7	Check the effects of correcting citation data for month of publication.	Different results with uncorrected citation data provide evidence for the necessity of our corrections.	(6)
8	Compare shot-term and long-term impact.	Short-term impact might be weakly associated with long-term impact.	-

Regarding the journal level, reviews published in special issues may have a higher visibility and receive more attention in the field (Bergh et al., 2006; Mingers & Xu, 2010). To assess whether this affects the results of our main model, we include a variable indicating whether the review was published as part of a special issue or not (model 1). We checked all full-texts (i.e., PDFs) and identified 10 reviews that were part of a special issue (4 reviews for describing, 3 reviews for understanding, 3 reviews for explaining). The results are robust with minor changes in the significance of one level of the research agenda variable (reviews for explaining) and the topic control variable (reviews for understanding and explaining).

Accessibility is another variable that may explain differences in citations as papers published in open access journals are available to more researchers who could access, use, and cite them (Björk & Solomon, 2012). Since our sample does not include open access journals, this variable could not bias our results. Similarly, we checked for individual papers published openly and not under subscription access control, however, and found none.

We further checked robustness regarding different measures of author visibility. In particular, decisions to cite a review paper might be influenced by the reputation of the most prolific author. Considering the review of Xiao and Benbasat (2007), for instance, a value of 24.5 (as an average of the h-indices 1 and 48 at the time of publication) may not capture the visibility of Izak Benbasat, who authored the review paper with his former PhD student, Bo Xiao. To check how the h-index of the most prolific author affects the model, model 2 includes the highest h-index of the authors. While the coefficients for the highest h-index differ from the coefficients for the average h-index, the effect sizes and significance of the other variables remain robust. Citing decisions could therefore be influenced by both, the average visibility of the author team and the visibility of the most prolific author. Similarly, controlling for the h-index calculated for the whole author team (Glänzel, 2008) shows the robustness of the other coefficients.

Since the size of the author team has been identified as a significant predictor of high impact research output (Fortunato et al., 2018; Wuchty et al., 2007), we checked whether this variable biases our results. In model 3, we include the number of authors as a variable, which was log-transformed due to its skewed distribution. We also checked robustness regarding other functional transformations. The number of authors is a significant positive predictor of scientific impact in the case of reviews for describing and theory testing; however, it has a negative effect for reviews for understanding and explaining. This indicates that teamwork may have favorable effects if the review type is associated with a structured application of prescribed methodological procedures. For reviews that require creative thought and innovative problem shifts, such as reviews for understanding and explaining, teamwork seems to have a negative effect. This may be explained by the lack of a method for discovering novel insights (Bechtel, 1988) that could inform and guide an author team in developing reviews for understanding and explaining. Generally, the other coefficients do not change substantially, underlining the robustness of our main results.

Soliciting external feedback might be associated with a higher scientific impact and thereby complement the effects of collaboration within the author team. Helpful scientists who provide conceptual feedback have been shown to affect the performance of their collaborators (Fortunato et al., 2018; Oettl, 2012). Furthermore, the importance of soliciting feedback during the development of a review paper has been emphasized repeatedly (Bem, 1995; Daft & Why, 1995; Webster & Watson, 2002). In model 4, we therefore included a dummy variable indicating whether the authors acknowledge conceptual feedback. In accordance with Oettl (2012), this measure was coded manually from the acknowledgments section of the review papers, considering keywords such as "comments", "suggestion", "review", "discussion", and "criticism", provided by other scholars. For descriptive reviews, the results show that acknowledging feedback has negative effects. This is consistent with the nature of these reviews, which may be perceived as being more neutral and objective. Reviews for understanding, which can be more opinionated, benefit from the solicitation and acknowledgment of external feedback. While feedback has mixed effects on the impact of review papers, the results of the other variables do not change substantially. Complementing prior research on the effects of helpful researchers who provide feedback, our results suggest that effects might not only pertain to the productivity of their colleagues but also to the impact of their research.

Novelty of a paper has been shown to affect its impact (Fortunato et al., 2018; Tahamtan et al., 2016; Uzzi et al., 2013). The requirement to give credit to original works by appropriate citations directly contributes to the impact of papers that introduce new ideas in a certain domain. This "first-mover advantage" could explain the high impact of the first review on a particular topic, or reviews that introduce new ideas and refer to unexplored literatures (e.g., the resource-based view from the management disciplines), for example. Beyond the (qualitative) coding of novelty (e.g., Grover et al., 2013; Judge et al., 2007), Uzzi et al. (2013) implement a measure for novelty that does not require (subjective) judgment. The measure is based on novel, or a-typical combinations of references used in a given paper. We followed this approach and measured novelty of review papers as the percentage of cited works that have not previously been covered by a review paper. Reference data was extracted from the full-texts and matched with reference data from the other review papers. The suggested measure does not cover cases in which a review uses the same references that were included in a previous review but uses them in a distinct, novel way. It may also be questioned whether the percentage of novel references captures novelty in an equal way for reviews that cite few (potentially high-quality journal) papers and reviews that cite extensively (e.g., journal papers, conference papers, etc.). Nevertheless, the values of the novelty variable attest to the validity of this measure. While the early reviews have novelty-scores close to unity, the more recent reviews, in particular those on popular research topics that have been reviewed frequently, have lower novelty-scores. Evidently, descriptive reviews in particular benefit from including papers that have not been considered by previous reviews in IS. The results of model 5 suggest that our main model is robust to the effects of novelty, or "first-mover advantages".

As our procedure of correcting citation data for the month of publication (instead of using citation data that is aggregated on an annual level) is novel, we analyze its effect on the coefficients (model 6). While several coefficients changed slightly, the most substantial changes can be observed in the subset for descriptive reviews: the standardized coefficient for the topic and transparency variables changed twofold with a strong increase in significance for the topic variable. This suggests that our procedure is necessary and that short-term citation scores need to be crafted carefully when used as a dependent variable. This methodological detail is critical to avoiding biases in scientometric studies that intend to explain short-term impact.

Finally, review papers with a high short-term impact might not necessarily be on a trajectory to become high-impact in the long-term. Figure 2 shows the development of citation scores of the top-10%, the bottom-10% and a random selection of 10% of the review papers. As Figure 2 indicates, high-impact reviews can be distinguished after very few years. Furthermore, the three-year citation rates correlate significantly with citation rates after four (98.94%, $p < 0.01$), five (97.74%, $p < 0.01$) and six (96.74%, $p < 0.01$) years, respectively. The correlations suggest that short and long-term impact are strongly related. Explaining long-term impact would raise three empirical problems. First, self-reinforcing mechanisms, i.e., the Matthew effect (cf. Merton, 1968), may result in impactful review papers biasing the coefficients. Second, with an increasing time-lag the visibility of authors may increase due to an impactful review, thereby aggravating problems of reverse causality associated with correlational analyses. Third, an appropriate functional form for the development of citations over time would be necessary to analyze recent as well as dated papers.

In summary, the results of the robustness checks are qualitatively similar to the main results, as shown in Table 4 and 5. The only caveat is that the effect of transparency is not robust for reviews for describing. Due to our sample size, test power might not be sufficient to reliably detect lower effect sizes. In the case of reviews for describing, a low and non-robust effect indicates that citing decisions are influenced by other (proxy) variables such as the journal impact factor, rather than methodological transparency. Overall, the fact that our main results are not substantially affected by alternative explanations suggests that our (main) model provides a robust and parsimonious explanation for the scientific impact of review papers in our field.

Table 4. Robustness Checks (Models 1-6)

	Review for Describing (I)							Review for Understanding (II)						
Effect	Main	(1)	(2)	(3)	(4)	(5)	(6)[b]	Main	(1)	(2)	(3)	(4)	(5)	(6)[b]
Journal Impact Factor	0.57**	0.61**	0.58**	0.57**	0.57**	0.55**	0.64**	0.35**	0.35**	0.34**	0.33**	0.29**	0.35**	0.40**
Special issue		0.70**							−0.93**					
h-index (average)	0.27**	0.29**		0.27**	0.26**	0.31**	0.31**	0.45**	0.41**		0.53**	0.46**	0.46**	0.37**
h-index (highest)			0.16**							0.30**				
Team size (log)				0.22**							−0.43**			
Feedback					−0.16**							0.36**		
Topic	−0.07	−0.09*	−0.04	−0.04	−0.05	−0.01	−0.14**	0.00	0.10*	0.04	−0.03	0.03	−0.01	−0.09
Novelty						0.25**							0.07	
Transparency (score)	0.10**	0.09**	0.00	0.05	0.12**	0.18**	0.20**	0.23**	0.17**	0.13**	0.26**	0.19**	0.24**	0.18**
Research Agenda (none)	−0.27**	−0.34**	−0.34**	−0.34**	−0.24**	−0.26**	−0.16*	[a]	[a]	[a]	[a]	[a]	[a]	[a]
Research Agenda (complete)	0.30**	0.19**	0.34**	0.25**	0.33**	0.24**	0.32**	0.51**	0.55**	0.69**	0.49**	0.51**	0.49**	0.48**
R^2	0.32	0.34	0.28	0.34	0.32	0.35	0.30	0.48	0.49	0.46	0.51	0.48	0.48	0.44

Notes. DV: citations. Model includes an intercept. Standardized beta coefficients are reported.
[a] Not enough observations available to include the variable.
[b] Changes in this column indicate that the corrections of the independent variable were necessary.
*significant at 1%, **significant at 0.1%.

Table 5. Robustness Checks (Models 1-6)

	Review for Explaining (III)							Review for Theory Testing (IV)						
Effect	Main	(1)	(2)	(3)	(4)	(5)	(6)[b]	Main	(1)	(2)	(3)	(4)	(5)	(6)[b]
Journal Impact Factor	0.19**	0.21**	0.18**	0.18**	0.19**	0.19**	0.19**	0.01	0.01	0.02	−0.03	0.01	0.00	0.05
Special issue		0.45**							[a]					
h-index (average)	−0.05	−0.04		−0.06*	−0.05	−0.04	−0.06	0.13**	0.13**		0.14**	0.13**	0.13**	0.08
h-index (highest)			−0.14**							0.07				
Team size (log)				−0.16**							0.30**			
Feedback					0.04							−0.02		
Topic	0.22**	0.24**	0.24**	0.23**	0.22**	0.23**	0.22**	0.18**	0.18**	0.17**	0.20**	0.19**	0.16**	0.21**
Novelty						0.08**							−0.11*	
Transparency (score)	0.26**	0.25**	0.31**	0.28**	0.25**	0.28**	0.22**	0.53**	0.53**	0.54**	0.52**	0.53**	0.46**	0.42**
Research Agenda (none)	−0.55**	−0.56**	−0.53**	−0.52**	−0.54**	−0.48**	−0.55**	[a]	[a]	[a]	[a]	[a]	[a]	[a]
Research Agenda (complete)	0.13*	0.10	0.18**	0.14*	0.13*	0.12*	0.20**	[a]	[a]	[a]	[a]	[a]	[a]	[a]
R^2	0.41	0.42	0.40	0.41	0.41	0.41	0.40	0.47	0.47	0.46	0.47	0.48	0.47	0.44

Notes. DV: citations. Model includes an intercept. Standardized beta coefficients are reported.
[a] Not enough observations available to include the variable.
[b] Changes in this column indicate that the corrections of the independent variable were necessary.
*significant at 1%, **significant at 0.1%.

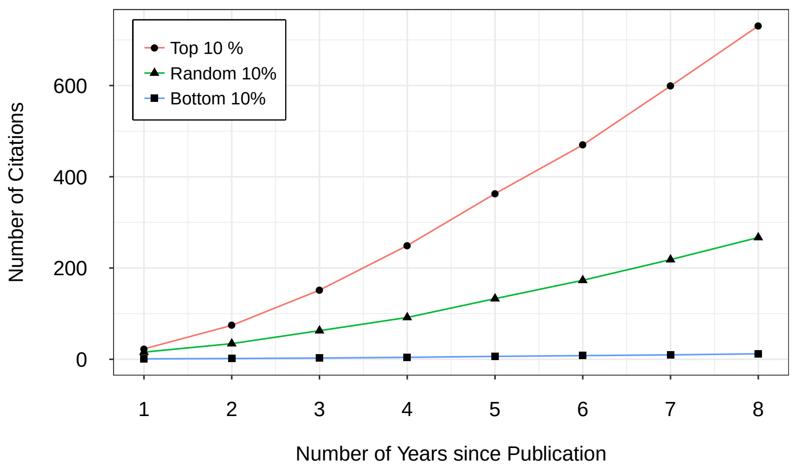


Figure 2. Cumulative Scientific Impact over Time

5. Discussion

Our study contributes to the vibrant discourse on literature reviews in IS, which is informed by its sister disciplines and offers many facets ranging from editorials, debates, methodological guidelines and opinion pieces to panels, tutorials and teaching material. This discourse, however, is largely based on anecdotal evidence and illustrative examples as opposed to reliable evidence that would support competing views on the attributes that distinguish impactful reviews. We think that our insights contribute to the discourse on literature reviews and inform the different stakeholders involved in the development and publication of review papers in the IS field, including authors (junior and senior), methodologists, reviewers, and editors.

5.1. Contributions

While our study offers many detailed results, we emphasize its broader contributions to both the literature on review papers and scientometric research in general. One contribution of our study is to provide substantial empirical evidence on attributes of different types of reviews that are successful in terms of scientific impact. The main results thereby suggest that it is necessary to consider nuances between types of reviews instead of conceiving the review genre as a monolithic block. Understanding the attributes that distinguish impactful reviews is crucial for various stakeholders involved in the development of review papers. Specific implications will be discussed in the following subsections.

The study also contributes to research by demonstrating that developing a research agenda is significantly associated with higher scientific impact. Developing a research agenda, as a scientometric variable, has received scant attention both within and beyond IS research. To the best of our knowledge, the only scientometric study analyzing this variable was conducted by Sternberg and Gordeeva (1996), who show that researchers expect papers providing value for future research, inter alia, to exert higher scientific impact. Despite its significance in the literature, especially in the literature on review papers (e.g., Rowe, 2014; Webster & Watson, 2002), the effect of this variable has not yet been analyzed in a scientometric impact model before. By including the development of a research agenda in our model and estimating its effect, we confirm its importance for IS review papers, and thus introduce a new variable to the arsenal of scientometric models (cf. Tahamtan et al., 2016) and show that it has a significant, high, and robust effect on scientific impact.

Our insights further contribute to recent debates on the role of transparency in review papers (cf. e.g., Boell & Cecez-Kecmanovic, 2015; Paré et al., 2016). They empirically show that the association between citations and transparency varies between different types of reviews. While this association is strong for reviews that systematically summarize evidence from prior research (theory testing reviews), it is weaker for traditional narrative reviews (reviews for describing). Surprisingly, reviews that require original, imaginative, or critical engagement (reviews for understanding and explaining) also achieve a higher impact when they are more transparent. This contrasts with the view that the original thought communicated through these types of reviews does not need to be complemented by a transparent methodology (Leidner, 2018).

Another contribution related to transparency is the scientometric insight that the effects of transparency and the journal impact factor are inversely related. This general tendency suggests that citing decisions rely either on the journal impact factor, which serves as a proxy variable for the quality of papers, or on the transparency of the paper itself. The result that the journal impact factor, one of the best predictors of scientific impact in many scientometric studies, can even become non-significant after including transparency suggests that future scientometric research should not avoid the efforts required for coding and including the transparency variable.

Overall, our model provides a powerful, parsimonious and robust explanation of impactful review papers and advances current scientometric analyses in IS in several regards. We assessed review papers published in an exceptionally large scope of journals and carefully developed a range of variables, which in turn are based on the content of the paper as opposed to meta-data.

We believe that much of the latent skepticism towards scientometric papers in IS is explained by their preoccupation with superficial meta-data and we are confident that a stronger focus on content-related aspects will raise the interest of a broad audience. This requires extensive efforts in content analysis, such as manually coding 21 distinct transparency items or categorizing the development of a research agenda. We deliberately focused on a small set of variables in our main model to avoid problems related to kitchen-sink models encountered in previous scientometric studies, including correlation between independent variables and model overfitting. To assess and control possible biases, we conducted a comprehensive set of robustness checks and implemented corrections of the dependent variable. We hope that our study thereby raises the standards and value of future scientometric studies both within and beyond IS research.

5.2. Implications

Prospective authors of review papers can draw on our results to inform their decisions on how to develop their specific type of review and which characteristics to focus on when confronted with page and time restriction. Our results enable them to focus their review on attributes, including methodological characteristics, that are coherent with the specific type of review and likely to stimulate scientific impact. One quality that is important across the different types of review papers is the transparent reporting of the methodological process. These new insights, representing substantial evidence of citing decisions of more than 50,000 papers, is useful to inform recent contradictory debates in which the role of methodology in reviews for explaining, and theoretical reviews in particular, has been contested (cf. Boell & Cecez-Kecmanovic, 2015; Leidner, 2018; Paré et al., 2016). Contributing to a solution of this confusion, we show that although transparency may not be the only aspect of quality, it is an important driver for stimulating follow-up research. Overall, we think that the preference for transparent reviews prevalent in subsequent research supports the argument that transparency is a necessary aspect for the trustworthiness of a review (Paré et al., 2016). If authors of a review do not report their methodological process, they deprive subsequent research of the ability to establish confidence in the reliability of the review's claims regarding what we know and do not yet know on a certain topic. Therefore, we encourage authors to refer to guidelines outlining transparency, systematicity and corresponding reporting standards (cf. Paré et al., 2016; Templier & Paré, 2018).

Concerning the development of a research agenda, our analyses show that it is not only a quality emphasized in editorials (e.g., Rowe, 2014; Webster & Watson, 2002), but one that also influences the future impact of review papers. Authors should therefore consider providing additional value by going beyond cursory gap spotting and instead provide more comprehensive guidance for future research efforts. Reviews for theory testing are an exception in this regard since developing a research agenda is not coherent with this particular type of review. They are also an exception regarding the importance of selecting a high-impact publication outlet. While authors may choose to target top-journals when submitting reviews for describing, understanding, and explaining, there is no evidence that this strategy would increase the impact of reviews for theory testing reviews. Extant scientometric research has heretofore failed to appropriately inform authors regarding these aspects.

Reviewers and editors should be aware of how citing behavior of subsequent research incentivizes authors of review papers. For example, reviewers should specifically encourage transparent methodology sections as well as well-grounded research agendas. In cases in which authors may be less attentive to these attributes (e.g., transparency for descriptive reviews, which is not a robust predictor of citation counts), requiring corresponding changes seems advisable. Although it may not immediately pay off in terms of citations, it is necessary for making reliable knowledge contributions. Reviews on less popular topics may also be cases in which trade-offs arise between scientific impact and scientific progress on the given topic.

Finally, publishing more review papers has been considered to be an effective means of pushing journal impact factors. Considering the increasing volume of low-impact review papers, this might not be an effective strategy. Furthermore, there is no evidence that theoretical review papers exert a higher scientific impact than reviews for describing, understanding, or theory testing, suggesting that reviewers and editors should be more open towards publishing all types of review papers. Instead, in particular for theoretical review papers, a lack of follow-up research may lead to the proliferation of uncontested knowledge, a tendency journal editors and reviewers should be aware of. In this regard, it is justified to consider the likelihood of stimulating subsequent research as a publication criterion.

5.3. Limitations and Avenues for Future Research

Although scientific impact is an important aspect of high quality reviews, we are cautious to present citations as the sole criterion guiding the discourse on review papers. For example, our results suggest that transparency is not a robust predictor of impact for reviews for describing. However, this should not lead to a neglect of transparency as transparent reporting is critical for the reliability and trustworthiness of these types of reviews (Templier & Paré, 2018). If transparency is not a high priority for authors who aim at increasing the impact of their reviews, reviewers and editors should require authors to adopt a systematic and transparent approach (i.e., to adhere to methodological reporting guidelines). Otherwise, striving for maximum scientific impact exclusively may have adverse effects on the reliability of knowledge development in our field. Similarly, the effects of authors' reputation require careful consideration. We do not consider them to suggest that junior scholars should reach out to senior scholars and to indiscriminately add any well-known author to the paper who does not immediately decline the request. Instead, we consider these effects to point to the role of experience and knowledge in the topic and the review methodology as an ingredient of high-quality reviews (Wong, Greenhalgh, Westhorp, Buckingham, & Pawson, 2013). This is also consistent with the results on teamwork and soliciting external feedback. We therefore encourage a more nuanced debate on how the field can draw on the experience of senior scholars. In this regard, we encourage further research on the process of developing high quality reviews. For example, surveys may offer insights into teamwork, solicitation of feedback, the use of methodological expertise and experience with the review methodology. Further tutorials and seminars on how to conduct various types of reviews should be integrated into PhD courses and mainstream IS conferences to raise awareness of and proficiency in applying appropriate methodologies.

Methodological limitations are related to our sample, which presents us with less than perfect conditions. The nature of our object of analysis prohibits us from implementing (experimental) research designs that are more appropriate for identifying causality. In a correlational, non-interventional design, we can only implement robustness checks to control for alternative explanations and further biases, such as those related to endogeneity. Dependencies between variables, particularly between those on the same level, necessitate a delicate selection of a few main variables. The low sample size limits the generalizability of our results although our scope covers as many as 40 journals and spans over 15 years. It also prohibits us from controlling for systematic trends over time. Furthermore, differences in focus on theory vs. practice and a general appreciation of methodological reporting standards suggest that there is no reason to expect our results to be representative beyond the IS discipline. Finally, citations are one possible measure of scientific impact that may not fully reflect scientific progress, knowledge development, or impact on research practice. These facets could be analyzed using other dependent variables (cf. Bollen, Van de Sompel, Hagberg, & Chute, 2009). For example, citation impact could be distinguished qualitatively, differentiating perfunctory impact, i.e., citations that do not engage with the content of the review, from ideational impact, i.e., citations expanding the knowledge developed in the review (Hassan & Loebbecke, 2017). Further, indicators for early diffusion into research practice such as reads, tweets, altmetrics or downloads (cf. Eysenbach, 2011; Fenner, 2014) could be analyzed.

Despite the volume of papers published on literature reviews, there are further avenues for future research and methodological advances. Concerning research agendas, procedural frameworks for their development and criteria for assessing the value of the output are a literal blind-spot in the literature. We therefore encourage a more comprehensive debate on the components that are required in a high-quality research agenda that can be considered a standalone contribution of a review paper. For example: Is it helpful to provide a long list of relatively unconnected research questions? Which aspects of research designs should be clarified to stimulate and enable subsequent author teams to follow up on the suggestions? When and in which way should the agenda be exposed to feedback from domain experts from the industry and academia? We hope that methodologists finally acknowledge the repeated calls for agenda development that can be found in editorials and, combined with our evidence, act upon these calls by developing corresponding methodological guidelines.

Furthermore, while scientific impact primarily reflects relevance to an academic audience, reviews should also be positioned for practical relevance and impact. Disciplines like the health sciences have been successful in positioning review papers as a channel for communicating knowledge and informing practice based on evidence. With less than 10 % of the reviews in our sample outlining implications for practice, review papers in IS do not yet fulfill their potential. By providing a methodology to identify topics that are relevant to practitioners but lack attention from researchers, Marrone and Hammerle (2016) take a valuable first step in this direction. We think that reviews are an appropriate genre to provide an overview of the current state of research from which practitioners can pull their topics of interest and inform their decisions. Building on the work of Oates (2015), the IS field has much to learn from the evidence-based practice (EBP) methodologies that have proven to be useful in other disciplines (e.g., Denyer & Tranfield, 2006).

6. Conclusion

At the outset of this study we noted that some of the prominent reviews in IS are cited more than twice a day on average, while others take years to accumulate single digit citations. Considering the magnitude of these differences and the proliferation of review papers in recent years, we conducted empirical analyses to understand what distinguishes those reviews that drive scientific progress from those that might not be considered impactful. Overall, our scientometric analyses of four types of review papers offer nuanced empirical insights into the content-related attributes affecting the scientific impact of review papers. Based on a parsimonious and powerful model, we show that on the paper level, the degree of methodological transparency and the development of a research agenda distinguish high-impact reviews in IS. These attributes have a significant effect on the citation impact of review papers after controlling for the journal impact factor, the average h-index of the author team and the topic of the review. We demonstrate the robustness of these effects by contrasting them with several alternative explanations. In short, our results are an important contribution to an informed debate on how we can leverage the power of review papers to drive scientific impact and knowledge development in IS. The IS research community in turn should continue its efforts to effectively build on the foundation provided by review papers.

References

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, *25*(1), 107–136.

Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, *36*(2), 247–271.

Aytug, Z. G., Rothstein, H. R., Zhou, W., & Kern, M. C. (2012). Revealed or concealed? transparency of procedures, decisions, and judgment calls in meta-analyses. *Organizational Research Methods*, *15*(1), 103–133.

Baldi, S. (1998). Normative versus social constructivist processes in the allocation of citations: A network-analytic model. *American Sociological Review*, *63*(6), 829–846.

Bechtel, W. (1988). *Philosophy of science: An overview for cognitive science.* Hillsdale, NJ: Lawrence Erlbaum.

Bem, D. J. (1995). Writing a review article for Psychological Bulletin. *Psychological Bulletin*, *118*(2), 172–177.

Bergh, D. D., Perry, J., & Hanke, R. (2006). Some predictors of SMJ article impact. *Strategic Management Journal*, *27*(1), 81–100.

Bikard, M., Murray, F., & Gans, J. S. (2015). Exploring trade-offs in the organization of scientific work: Collaboration and scientific reward. *Management Science*, *61*(7), 1473–1495.

Björk, B.-C., & Solomon, D. (2012). Open access versus subscription journals: a comparison of scientific impact. *BMC Medicine*, *10*(1), 73.

Blumberg, B., Cooper, D. R., & Schindler, P. S. (2005). *Business research methods* (2nd ed.). New York: McGraw-Hill Education.

Boell, S. K., & Cecez-Kecmanovic, D. (2015). On being "systematic" in literature reviews in IS. *Journal of Information Technology*, *30*(2), 161–173.

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLOS ONE*, *4*(6), 1–11.

Braun, T., Glänzel, W., & Grupp, H. (1995). The scientometric weight of 50 nations in 27 science areas, 1989–1993. Part I. All fields combined, Mathematics, Engineering, Chemistry and Physics. *Scientometrics*, *33*(3), 263–293.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum.

Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, *50*(6), 1281–1303.

Daft, R. L., & Why, I. (1995). Why I recommended that your manuscript be rejected and what you can do about it. *Publishing in the Organizational Sciences*, 164–183.

Denyer, D., & Tranfield, D. (2006). Using qualitative research synthesis to build an actionable knowledge base. *Management Decision*, *44*(2), 213–227.

Ellison, G. (2002). The slowdown of the economics publishing process. *Journal of Political Economy*, *110*(5), 947–993.

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, *13*(4).

Fenner, M. (2014). Altmetrics and other novel measures for scientific impact. In S. Bartling & S. Friesike (Eds.), *Opening science* (pp. 179–189). Springer.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., ... Barabási, A.-L. (2018). Science of Science. *Science*, *359*(6379), 1–7.

Garfield, E. (1987). Reviewing review literature. Part 1. Definitions and uses of reviews. *Essays of an Information Scientist*, *10*, 113–116.

Garfield, E. (2006). The history and meaning of the Journal Impact Factor. *Journal of the American Medical Association*, *295*(1), 90–93.

Glänzel, W. (2008). H-index concatenation. *Scientometrics*, *77*(2), 369–372.

Grover, V., Raman, R., & Stubblefield, A. (2013). What affects citation counts in MIS research articles? An empirical investigation. *Communications of the Association for Information Systems*, *34*, 1435–1456.

Hansen, S., Lyytinen, K., & Markus, M. L. (2006). The legacy of "Power and Politics" in disciplinary discourse: A citation analysis. In *Proceedings of the 27th International Conference on Information Systems.* Milwaukee, WI, USA.

Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination.* London: Sage.

Hassan, N. R., & Loebbecke, C. (2017). Engaging scientometrics in information systems. *Journal of Information Technology*, *32*(1), 85–109.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences of the United States of America*, *102*(46), 16569–16572.

Hoffman, D. L., & Holbrook, M. B. (1993). The intellectual structure of consumer research: A bibliometric study of author cocitations in the first 15 years of the Journal of Consumer Research. *Journal of Consumer Research*, *19*(4), 505–517.

Hurt, C. D. (1987). Conceptual citation differences in science, technology, and social sciences literature. *Information Processing & Management*, *23*(1), 1–6.

Hwang, M. I. (1996). The use of meta-analysis in MIS research: Promises and problems. *ACM SIGMIS Database*, *27*(3), 35–48.

Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, *20*(3), 341–367.

Jackson, D. N., & Rushton, J. (1987). *Scientific excellence: Origins and assessment.* New York, NY, USA: SAGE Publications.

Judge, T. A., Cable, D. M., Colbert, A. E., & Rynes, S. L. (2007). What causes a management article to be cited – Article, author, or journal? *Academy of Management Journal*, *50*(3), 491–506.

Lacity, M. C., Solomon, S., Yan, A., & Willcocks, L. P. (2011). Business process outsourcing studies: A critical review and research directions. *Journal of Information Technology*, *26*(4), 221–258.

Leidner, D. E. (2018). Review and theory symbiosis: An introspective retrospective. *Journal of the Association for Information Systems*, *19*(6), 552–567.

Leimu, R., & Koricheva, J. (2005). What determines the citation frequency of ecological papers? *Trends in Ecology & Evolution*, *20*(1), 28–32.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of Information Systems Research. *Informing Science*, *9*, 181–212.

Loebbecke, C., Huyskens, C., & Berthod, O. (2007). Research importance in the Information Systems field: A citations analysis. In *Proceedings of the 28th International Conference on Information Systems.* Montreal, QC, Canada.

Lowry, P. B., Moody, G. D., Gaskin, J., Galletta, D. F., Humpherys, S. L., Barlow, J. B., & Wilson, D. W. (2013). Evaluating journal quality and the Association for Information Systems Senior Scholars' Journal Basket via bibliometric measures: Do expert journal assessments add value? *MIS Quarterly*, *37*(4), 993–1012.

Marrone, M., & Hammerle, M. (2016). An integrated literature review: Establishing relevance for practitioners. In *Proceedings of the 37th International Conference on Information Systems.* Dublin, Ireland.

Merton, R. K. (1968). The Matthew Effect in science. *Science*, *159*(3810), 56–63.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations.* University Of Chicago Press.

Mingers, J., & Xu, F. (2010). The drivers of citations in Management Science journals. *European Journal of Operational Research*, *205*(2), 422–430.

Montori, V. M., Wilczynski, N. L., Morgan, D., & Haynes, R. B. (2003). Systematic reviews: A cross-sectional study of location and citation counts. *BMC Medicine*, *1*(1), 1–7.

Oates, B. J. (2015). On systematic reviews for evidence-based practice. *Journal of Information Technology*, *30*(2), 177–179.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science*, *58*(6), 1122–1140.

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, *37*(1), 879–910.

Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in Information Systems literature reviews. *European Journal of Information Systems*, *25*(6), 493–508.

Patsopoulos, N. A., Analatos, A. A., & Ioannidis, J. P. A. (2005). Relative citation impact of various study designs in the Health Sciences. *The Journal of the American Medical Association*, *293*(19), 2362–2366.

Peters, H. P. F., & van Raan, A. F. J. (1994). On determinants of citation scores: A case study in Chemical Engineering. *Journal of the American Society for Information Science*, *45*(1), 39–49.

Rivard, S. (2014). Editor's comments: The ions of theory construction. *MIS Quarterly*, *38*(2), iii–xiv.

Rowe, F. (2014). What literature review is not: Diversity, boundaries and recommendations. *European Journal of Information Systems*, *23*(3), 241–255.

Sandberg, J., & Alvesson, M. (2011). Ways of constructing research questions: Gap-spotting or problematization? *Organization*, *18*(1), 23–44.

Schwarz, A., Mehta, M., Johnson, N., & Chin, W. W. (2007). Understanding frameworks and reviews: A commentary to assist us in moving our field forward by analyzing our past. *ACM SIGMIS Database*, *38*(3), 29–50.

Sidorova, A., Evangelopoulos, N., Valacich, J. S., & Ramakrishnan, T. (2008). Uncovering the intellectual core of the Information Systems discipline. *MIS Quarterly*, *32*(3), 467–482.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management*

Science, *51*(5), 756–770.

Starbuck, W. H. (2005). How much better are the most-prestigious journals? The statistics of academic publication. *Organization Science*, *16*(2), 180–200.

Sternberg, R. J., & Gordeeva, T. (1996). The anatomy of impact: What makes an article influential? *Psychological Science*, *7*(2), 69–75.

Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. *ACM SIGMIS Database*, *71*(3), 171–193.

Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, *107*(3), 1195–1225.

Tams, S., & Grover, V. (2010). The effect of an IS article's structure on its impact. *Communications of the Association for Information Systems*, *27*, 149–172.

Te'eni, D., Rowe, F., Ågerfalk, P. J., & Lee, J. S. (2015). Publishing and getting published in EJIS: Marshaling contributions for a diversity of genres. *European Journal of Information Systems*, *24*(6), 559–568.

Templier, M., & Paré, G. (2018). Transparency in literature reviews: An assessment of reporting practices across review types and genres in top IS journals. *European Journal of Information Systems*, *27*(5), 503–550.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472.

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in Information Systems Research. *Communications of the Association for Information Systems*, *37*(9), 205–224.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), xiii–xxiii.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: Realist syntheses. *BMC Medicine*, *11*(1), 1–14.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.

Xiao, B., & Benbasat, I. (2007). E-Commerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly*, *31*(1), 137–209.

Zorn, T. E., & Campbell, N. (2006). Improving the writing of literature reviews through a literature integration exercise. *Business Communication Quarterly*, *69*(2), 172–183.

1. Attributes Included in Previous Scientometric Studies

To provide an overview of the various attributes considered in scientometric studies, we searched literature reviews on scientometrics (e.g., Tahamtan, Afshar, & Ahamdzadeh, 2016), background sections (e.g., Mingers & Xu, 2010) and major studies of scientometric impact (e.g., Bergh, Perry, & Hanke, 2006; Colquitt & Zapata-Phelan, 2007; Stremersch, Verniers, & Verhoef, 2007).

The summary in Tables 1 to 3 provides a representative overview of these attributes rather than an exhaustive list. To provide a condensed table, attributes are summarized using the same label if they are measured similarly (e.g., article length and number of pages) or equivalent after transformations (e.g., age of the paper and year of publication). The last column indicates how the attribute is considered in our analyses and the underlying rationale.

Table 1. Paper Level Attributes Included in Previous Scientometric Studies

Attribute	Exemplary references	Inclusion/Rationale
Methodology	Bergh et al. (2006); Judge, Cable, Colbert, and Rynes (2007)	Main model
Research agenda	Sternberg and Gordeeva (1996)	Main model
Paper type	Bergh et al. (2006); Judge et al. (2007); Mingers and Xu (2010); Stremersch et al. (2007)	Used for the subsets
Goal	Judge et al. (2007); Mingers and Xu (2010)	Used for the subsets
Topic/Field size	Bergh et al. (2006); Stremersch et al. (2007)	Control model
Novelty	Stremersch et al. (2007); Uzzi, Mukherjee, Stringer, and Jones (2013)	Robustness checks
Strength of theoretical contribution	Grover, Raman, and Stubblefield (2013); Judge et al. (2007); Merton (1957)	Not considered (difficult to assess reliably; Does not apply to all types of reviews (e.g., meta-analyses))
Attention grabbers (e.g., length of title, number of keywords)	Ayres and Vars (2000); Stremersch et al. (2007)	Not considered (superficial attribute)
Number of references	Mingers and Xu (2010); Stremersch et al. (2007)	Not considered (superficial attribute)
Presentation	Judge et al. (2007); Stremersch et al. (2007)	Not considered (superficial attribute)
Article length	Bergh et al. (2006); Judge et al. (2007); Leimu and Koricheva (2005); Mingers and Xu (2010); Stremersch et al. (2007)	Not considered (superficial attribute)
Awards	Stremersch et al. (2007)	Not considered (potential of reverse causality)
Year of publication	Bergh et al. (2006); Judge et al. (2007)	Dependent variable measured after 3 years for every review paper
Language	Leimu and Koricheva (2005)	Limited analysis to reviews published in English

Table 2. Author Level Attributes Included in Previous Scientometric Studies

Attribute	Exemplary references	Inclusion/Rationale
Publication record (e.g., h-index, top-tier publications)	Bergh et al. (2006); Judge et al. (2007); Leimu and Koricheva (2005); Stremersch et al. (2007)	Control model
Number of authors	Bergh et al. (2006); Fortunato et al. (2018); Mingers and Xu (2010); Stremersch et al. (2007); Wuchty, Jones, and Uzzi (2007)	Robustness checks
Soliciting feedback	Oettl (2012)	Robustness checks
Affiliation	Bergh et al. (2006); Judge et al. (2007); Leimu and Koricheva (2005); Stremersch et al. (2007)	Not considered (no IS-specific ranking for the time-frame available)
Nationality	Leimu and Koricheva (2005)	Not considered (requires many dummy variables)
Gender	Judge et al. (2007); Leimu and Koricheva (2005)	Not considered
Age	Ayres and Vars (2000); Bergh et al. (2006)	Not considered (related to publication record)
Editorial board membership	Stremersch et al. (2007)	Not considered (differences in visibility between boards of different journals)
Self-citations	Stremersch et al. (2007)	Excluded from dependent variable

Table 3. Journal Level Attributes Included in Previous Scientometric Studies

Attribute	Exemplary references	Inclusion/Rationale
Journal impact	Judge et al. (2007)	Control model
Special issue	Bergh et al. (2006); Mingers and Xu (2010)	Robustness checks
Position in the journal	Judge et al. (2007); Stremersch et al. (2007)	Not considered
Accessibility	Björk and Solomon (2012); Mingers and Xu (2010)	No open access papers included in our sample

2. Measures

Measure for methodological transparency of review papers

Methodological transparency is measured as the percentage of items relevant to the specific review type that were reported transparently:

Transparency score $= \sum_{i \in I_R} \mathrm{i}/|I_R|$,

with $i = 1$ if the item is reported ($i = 0$ otherwise) and I_R:= set of items required for the type of review $R = \{$Narrative review, descriptive review, scoping review, critical review, theory development review, qualitative systematic review, meta analysis$\}$ (cf. Paré, Trudel, Jaana, & Kitsiou, 2015). These types of reviews are aligned with the categories of reviews for describing, understanding, explaining, and theory testing. In the following, we briefly describe the items, which are based on Templier and Paré (2018). A mapping of the items required for each review type is provided in Table 4.

Problem formulation (Step 1): We coded whether the review transparently states the problem it addresses by specifying the primary goals/research questions. We further noted whether the key concepts or theories that are investigated are clearly defined.

Literature search (Step 2): We read the methodology sections to understand the literature search process. Specifically, we noted whether the authors describe how the literature search was performed, whether the application of multiple search strategies is outlined, and whether multiple publication types (such as journal papers, conference papers and books) are considered. Further, we coded whether the authors make the comprehensiveness of the search and its restrictions transparent. Furthermore, we analyzed whether the authors describe how the reputation of the sources is considered and whether strategies for minimizing publication bias are applied (if applicable).

Screening for inclusion (Step 3): To analyze the reporting of paper inclusion, we extracted data on the screening and selection of primary studies, and results of parallel independent study selection (such as inter-coder reliability coefficients). In addition, we noted whether it is made transparent how studies using the same dataset are treated and whether the screening process is illustrated by providing a corresponding flow diagram or description.

Quality assessment (Step 4): For those review types that are required to assess the quality of the primary studies, we coded whether the quality assessment procedures are described and whether results on parallel independent codings are provided.

Data extraction (Step 5): We captured reporting of data extraction by coding whether a data extraction plan is provided, by looking for descriptions of tools or methods used to extract the data and by searching for parallel independent coding processes.

Data analysis and interpretation (Step 6): Finally, we coded items on the data analysis and interpretation phase. These include a description of how the data analysis is performed, how study quality is considered in the interpretation of the findings and whether a profile of the studies is included (providing a distribution of the included papers over journals and time, for example). In addition, we analyzed whether the data analysis methods or techniques are justified and whether methodological limitations are made transparent.

Table 4. Mapping of Review Types and Required Items (adapted from Templier & Paré, 2018)

	Describing			Understanding	Explaining	Testing	
	Narrative	Descriptive	Scoping	Critical	Theory development	Qualitative systematic	Meta-analysis
Step 1: Problem formulation							
Primary goals or research questions	required	required	required	required	required	required	required
Key concepts or theories being investigated	required	required	required	required	required	required	required
Step 2: Literature search							
How the literature search is performed	required	required	required	required	required	required	required
Multiple search strategies			required	required	required	required	required
Multiple publication types			required	required	required	required	required
Comprehensiveness of search & restrictions if applicable		required	required	required	required	required	required
How reputation of the sources is considered	required			required	required		
Strategies used to minimise publication bias						required	required
Step 3: Screening for inclusion							
How primary studies are screened or selected	required	required	required		required	required	required
Results of parallel independent study selection		required	required			required	required
How studies using the same data-set are treated						required	required
Flow diagram or description of screening process			required			required	required
Step 4: Quality assessment							
How quality assessment is performed						required	required
Results of parallel independent assessment						required	required
Step 5: Data extraction							
Data extraction plan		required	required			required	required
Tools or methods used to extract data		required	required		required	required	required
Results of parallel independent coding process		required	required			required	required
Step 6: Data analysis and interpretation							
How data analysis is performed			required		required	required	required
How study quality is considered in interpretation of findings						required	required
Profile of the included studies		required	required			required	required
Justification of data analysis methods or techniques			required			required	required
Methodological limitations	required	required	required	required	required	required	required

3. Descriptive Statistics and Correlations

Table 5. All Reviews: Descriptive Statistics

	Factor	Mean	(Standard deviation)	Min/Max
1	Impact	42.17	51.29	(0/318.6)
2	Journal impact factor	2.09	1.51	(0.18/5.31)
3	h-index (average)	9.77	6.83	(0/38)
4	Topic	4.86	3.37	(0/24.07)
5	Transparency (score)	0.43	0.25	(0/1)

Table 6. All Reviews: Correlations

	Effect	1	2	3	4	5	6
1	Impact	1					
2	Journal impact factor	0.44**	1				
3	h-index (average)	0.25**	0.21**	1			
4	Topic	0.28**	0.43**	0.07**	1		
5	Transparency (score)	0.2 **	0.14**	−0.03**	0.2 **	1	
6	Research agenda	−0.21**	−0.22**	−0.15**	−0.02**	−0.08**	1
7	Type: Describing	−0.11**	−0.19**	−0.05**	−0.17**	0.3 **	0.1 **
8	Type: Understanding	−0.07**	−0.06**	−0.1 **	−0.1 **	−0.15**	−0.38**
9	Type: Explaining	0.12**	0.23**	0.04**	−0.02**	−0.34**	0.15**
10	Type: Testing	0.06**	−0.01**	0.13**	0.32**	0.24**	0.13**

Notes. *significant at 1%, **significant at 0.1%. N = 220. No correlations reported for the levels of the review type variable.

Table 7. Reviews for Describing: Descriptive Statistics

	Factor	Mean	(Standard deviation)	Min/Max
1	Impact	36.17	48.67	(0.7/318.6)
2	Journal impact factor	1.80	1.21	(0.18/5.31)
3	h-index (average)	9.41	7.86	(0/34.5)
4	Topic	4.31	2.80	(0.34/11.38)
5	Transparency (score)	0.51	0.23	(0.12/1)

Table 8. Reviews for Describing: Correlations

	Effect	1	2	3	4	5	6
1	Impact	1					
2	Journal impact factor	0.49**	1				
3	h-index (average)	0.35**	0.23**	1			
4	Topic	0.22**	0.5 **	0.04**	1		
5	Transparency (score)	0.02**	0.02**	−0.23**	0.24**	1	
6	Research agenda	−0.26**	−0.26**	−0.19**	−0.05**	0.15**	1

Notes. *significant at 1%, **significant at 0.1%. N = 74.

Table 9. Reviews for Understanding: Descriptive Statistics

	Factor	Mean	(Standard deviation)	Min/Max
1	Impact	37.86	48.09	(0/216.2)
2	Journal impact factor	1.97	1.43	(0.18/5.31)
3	h-index (average)	8.94	5.39	(0.5/26)
4	Topic	4.43	3.06	(0/15.29)
5	Transparency (score)	0.38	0.25	(0.06/1)

Table 10. Reviews for Understanding: Correlations

	Effect	1	2	3	4	5	6
1	Impact	1					
2	Journal impact factor	0.42**	1				
3	h-index (average)	0.33*	0.16	1			
4	Topic	0.23*	0.42	0.08	1		
5	Transparency (score)	0.2	0.21	−0.26	0.28	1	
6	Research agenda	−0.41*	−0.1	−0.25	−0.14	−0.17	1

Notes. *significant at 1%, **significant at 0.1%. N = 48.

Table 11. Reviews for Explaining: Descriptive Statistics

	Factor	Mean	(Standard deviation)	Min/Max
1	Impact	49.58	58.68	(2/288.6)
2	Journal impact factor	2.53	1.77	(0.18/5.31)
3	h-index (average)	10.07	6.77	(0/38)
4	Topic	4.79	2.95	(0/12.47)
5	Transparency (score)	0.32	0.26	(0/1)

Table 12. Reviews for Explaining: Correlations

	Effect	1	2	3	4	5	6
1	Impact	1					
2	Journal impact factor	0.4 **	1				
3	h-index (average)	0.13**	0.17**	1			
4	Topic	0.35*	0.49	0.29	1		
5	Transparency (score)	0.38**	0.24**	0.27**	0.14**	1	
6	Research agenda	−0.25**	−0.36**	−0.19**	−0.15**	−0.29**	1

Notes. *significant at 1%, **significant at 0.1%. N = 65.

Table 13. Reviews for Testing: Descriptive Statistics

	Factor	Mean	(Standard deviation)	Min/Max
1	Impact	47.29	45.47	(0/202.6)
2	Journal impact factor	2.05	1.54	(0.18/5.31)
3	h-index (average)	11.19	6.37	(1.5/25)
4	Topic	6.85	4.87	(0.65/24.07)
5	Transparency (score)	0.52	0.20	(0.05/0.9)

Table 14. Reviews for Testing: Descriptive Statistics and Correlations

	Effect	1	2	3	4	5
1	Impact	1				
2	Journal impact factor	0.42	1			
3	h-index (average)	0.16	0.32	1		
4	Topic	0.34*	0.38	−0.28	1	
5	Transparency (score)	0.45	0.56	0.14	0.2	1

Notes. *significant at 1%, **significant at 0.1%. N = 33.

References

Ayres, I., & Vars, F. E. (2000). Determinants of citations to articles in elite law reviews. *The Journal of Legal Studies*, *29*(S1), 427–450.

Bergh, D. D., Perry, J., & Hanke, R. (2006). Some predictors of SMJ article impact. *Strategic Management Journal*, *27*(1), 81–100.

Björk, B.-C., & Solomon, D. (2012). Open access versus subscription journals: a comparison of scientific impact. *BMC Medicine*, *10*(1), 73.

Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, *50*(6), 1281–1303.

Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., ... Barabási, A.-L. (2018). Science of Science. *Science*, *359*(6379), 1–7.

Grover, V., Raman, R., & Stubblefield, A. (2013). What affects citation counts in MIS research articles? An empirical investigation. *Communications of the Association for Information Systems*, *34*, 1435–1456.

Judge, T. A., Cable, D. M., Colbert, A. E., & Rynes, S. L. (2007). What causes a management article to be cited – Article, author, or journal? *Academy of Management Journal*, *50*(3), 491–506.

Leimu, R., & Koricheva, J. (2005). What determines the citation frequency of ecological papers? *Trends in Ecology & Evolution*, *20*(1), 28–32.

Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, *22*(6), 635–659.

Mingers, J., & Xu, F. (2010). The drivers of citations in Management Science journals. *European Journal of Operational Research*, *205*(2), 422–430.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science*, *58*(6), 1122–1140.

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing Information Systems knowledge: A typology of literature reviews. *Information & Management*, *52*(2), 183–199.

Sternberg, R. J., & Gordeeva, T. (1996). The anatomy of impact: What makes an article influential? *Psychological Science*, *7*(2), 69–75.

Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: Drivers of article impact. *ACM SIGMIS Database*, *71*(3), 171–193.

Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, *107*(3), 1195–1225.

Templier, M., & Paré, G. (2018). Transparency in literature reviews: An assessment of reporting practices across review types and genres in top IS journals. *European Journal of Information Systems*, *27*(5), 503–550.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*(6157), 468–472.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039.