A Framework for Guiding and Evaluating Literature Reviews

Mathieu Templier
Department of Management Information Systems, Université Laval, mathieu.templier@sio.ulaval.ca

Guy Paré
Research Chair in Information Technology in Health Care, HEC Montréal

Follow this and additional works at: https://aisel.aisnet.org/cais

A Framework for Guiding and Evaluating Literature Reviews

Mathieu Templier

Department of Management Information Systems,
Université Laval

mathieu.templier@sio.ulaval.ca

Guy Paré

Research Chair in Information Technology in Health
Care, HEC Montréal

Abstract:

Given that the synthesis of cumulated knowledge is an essential condition for any field to grow and develop, we believe that the enhanced role of IS reviews requires that this expository form be given careful scrutiny. Over the past decade, several senior scholars have made calls for more review papers in our field. While the number of IS review papers has substantially increased in recent years, no prior research has attempted to develop a general framework to conduct and evaluate the rigor of standalone reviews. In this paper, we fill this gap. More precisely, we present a set of guidelines for guiding and evaluating IS literature reviews and specify to which review types they apply. To do so, we first distinguish between four broad categories of review papers and then propose a set of guidelines that are grouped according to the generic phases and steps of the review process. We hope our work will serve as a valuable source for those conducting, evaluating, and/or interpreting reviews in our field.

Keywords: Literature Review, Standalone Review Paper, Research Synthesis, Review Methodology, Methodological Guidelines, Review Quality.

1   Introduction

Although the information systems (IS) field is relatively young compared to other social science fields (King & He, 2005), it has grown considerably since the early 1960s. Indeed, the domain's maturity is evidenced by the increasing volume of published IS research, the emergence of well-established journal outlets (Chen & Hirschheim, 2004), and the development of IS research perspectives and theories (Baskerville & Myers, 2002; Webster & Watson, 2002). As another sign of maturity, IS has recently become a reference field for studies in psychology, education, marketing, and other management domains (King & He, 2005).

However, to further its progress, the domain still needs to consolidate its research tradition. It has become increasingly difficult for researchers to remain knowledgeable of the many aspects of IS research. According to Card (2012), one obstacle to scientific progress is that researchers are limited in their ability to retain, organize, and synthesize earlier knowledge while staying abreast of new scientific contributions. Similarly, Cooper (2009) notices that an increased quantity of scientific inquiry has resulted in growing specialization in the social sciences. As a result, "time constraints make it impossible for most social scientists to keep up with primary research except within a few topic areas of special interest to them" (Cooper, 2009, p. 2). Since individual studies incrementally contribute to a larger understanding of a phenomenon of interest, building scientific knowledge requires cooperation and interdependent research work (Cooper, 2009). By uncovering prior knowledge, literature reviews offer foundations for further scientific research and are, thus, essential to any field's development. In the IS domain, the increasing amount of research and its rapid diffusion also call for reliable ways of integrating the findings of prior studies (Bandara, Miskon, & Fielt, 2011). In this era of information overload, it is essential that IS researchers engage in the process of research synthesis.

Hart (1999, p. 13) defines a literature review as:

> the selection of available documents (both published and unpublished) on the topic, which contain information, ideas, data and evidence written from a particular standpoint to fulfill certain aims or express certain views on the nature of the topic and how it is to be investigated, and the effective evaluation of these documents in relation to the research being proposed.

A literature review can either serve as the background for an empirical study or as an independent, standalone piece that provides a valuable contribution in its own right (Jesson, Matheson, & Lacey, 2011; Okoli & Schabram, 2010). The former, which can consist of a section in a journal paper or a chapter in a graduate thesis, is the most common type of review. The purpose of the review section or chapter is to help a researcher "acquire an understanding of (the) topic, of what has already been done on it, how it has been researched, and what the key issues are" (Hart, 1999, p. 1). In addition, the background section helps to contextualize the study's contributions and justify its approaches, research methods, tools, questions, and methods (Levy & Ellis, 2006a).

The second type, called here standalone literature review, is a "journal-length article whose sole purpose is to review the literature in a field, without any primary data…collected or analyzed" (Okoli & Schabram, 2010, p. 2). Such reviews are conducted for many different purposes (e.g., to make sense of existing knowledge on a particular topic, facilitate theory development, synthesize the extant literature on widely studied and mature areas, or identify research domains where further investigation is needed) (Webster & Watson, 2002; King & He, 2005; Okoli & Schabram, 2010). While providing a critical account of prior research might represent the sole objective of some reviews (Cooper, 1988), authorial critique can play a role (to different degrees) in all review types, which we illustrate in Section 2. In short, high-quality standalone reviews provide a valuable and trustworthy account of past research that other researchers might seek out for inspiration and use to position their own studies. In this line of thought, prior research has shown that review papers frequently become "core" or "milestone" papers in a field (Garfield, 1982; Paré, Trudel, Jaana, & Kitsiou, 2015; Rowe, 2014). Further, review papers play an important role in fostering fields as reference fields. They represent a key source of knowledge for new scholars and doctoral students entering a field and for those researchers outside the field.

Moreover, in light of the calls for increased use of evidence-based management (i.e., the systematic use of the best empirical evidence to improve management practice) (Pfeffer & Sutton, 2006; Rousseau, Manning, & Denyer, 2008; Reay, Berta, & Kohn, 2009), review papers become essential tools for

summarizing or synthesizing the extant literature in all applied fields such as management, marketing, finance, human resources, and information systems.

When rigorously conducted, reviews represent powerful information sources for researchers and practitioners looking for existing evidence to guide their decision making and managerial practices. The evidence-base paradigm in information systems aims to inform decisions about issues such as the adoption of new systems, software development methods and tools, and governance models because decision makers rely on and draw on comprehensive literature reviews (Oates, Edwards, & Wainwright, 2012). Like Atkins and Louw (2000) and Oates (2011), we call for the use of more high-quality review papers that synthesize available knowledge for professional practice.

The relevance of publishing standalone literature reviews in the IS field was first underscored by the creation of the *Management Information Systems Quarterly* (MISQ) review department in 2001. This department, which later became *MISQ Theory and Review*, is an outlet for review papers that synthesize prior research and aims to facilitate the dissemination of that knowledge (Watson, 2001). Its mission is to "accelerate the accumulation of IS knowledge" and "provide important input in setting directions for future research" (Webster & Watson, 2002, p. xiii). The lack of high-quality research syntheses is seen as a potential barrier to theoretical and conceptual progress in the IS field (Levy & Ellis, 2006a; Webster & Watson, 2002). For this reason, senior IS scholars have called for more review papers to foster the field's development and cumulative tradition (e.g., King & He, 2005; Rowe, 2014; Schwarz, Mehta, Johnson, & Chin, 2007; Watson, 2001; Webster & Watson, 2002;). To address the abovementioned problems and help information systems academics keep abreast with developments in our field, *Foundations and Trends in Information Systems*, a new journal fully dedicated to review papers, was launched in 2014.

According to Webster and Watson (2002), one of the challenges in integrating previous research and advancing knowledge is scholars' lack of familiarity with the methods used for structuring and presenting reviews. The issue of quality is a critical aspect that should be discussed and examined when producing or assessing standalone literature reviews (Paré et al., 2015). Indeed, calls for more IS review papers emphasize quality standards. Since literature reviews serve as "benchmarks" for other researchers in a field, they should cover the relevant literature to date and earn readers' confidence about the validity, reliability, and relevance of their findings. More specifically, we suggest that a research synthesis's quality involves three dimensions: rigor, relevance, and methodological coherence between the review's components and its objectives. Rigor refers to the soundness of the research process (Ogawa & Malen, 1991; Tobin & Begley, 2004). Without rigor, research has no scientific value and makes no contribution to knowledge (Morse, Barrett, Mayan, Olson, & Spiers, 2002). Rigor is a desired goal that is met by satisfying criteria such as internal validity, external validity, reliability, and objectivity (Morse et al., 2002; Ogawa & Malen, 1991). Relevance refers to the review's utility and usefulness. As we mention above, standalone literature reviews are conducted for many different purposes, such as to analyze the progress of a specific stream of research, to aggregate findings or reconcile equivocal results of prior studies, to review the application of a theoretical model or a methodological approach, to develop a new theory or research model, and to provide a critical account of prior research on a particular topic (Cooper, 1988). Lastly, methodological coherence refers to the "congruence between the research question and the components of the method" (Morse et al., 2002, p. 18). Methodological coherence links rigor and relevance, and validates the fit between a review's goals and the methodological guidelines chosen to attain them.

In this paper, we focus on methodological rigor because it is essential to research quality. Indeed, rigor is particularly important for reviews that are intended as milestones for future research. However, as Paré et al. (2015) stress, the discussion on methodological rigor in relation to review paper remains highly abstract and is thus prone to subjectivity and judgment on the part of researchers, assessors, and editors. Indeed, "in order for the reader—be it a researcher, a policy maker or a practitioner—to benefit from the evidence in the literature on a certain IS topic or subject, it is important to present an informative explanation about how the review was conducted" (Paré, Trudel, & Jaana, 2012, p. 19). Misconceptions about the methods used for conducting standalone literature reviews also raise issues related to rigor. In this regard, literature reviews are no different from other methodologies. A review's quality and coherence emerge from the application of a structured approach with specific guidelines. As Cooper (2009) states, "integrating separate research projects into a coherent whole involves inferences as central to the validity of knowledge as the inferences involved in drawing conclusions from primary data analysis" (p. 3). Therefore, in light of the growing interest in standalone reviews in the IS field, we posit that greater attention should be paid to the specific criteria used to evaluate their rigor.

Several fields, such as health sciences (e.g., Liberati et al., 2009; Higgins & Green, 2008), nursing (e.g., Cronin, Ryan, & Coughlan, 2008; Kirkevold, 1997; Whittemore, 2005), software engineering (e.g., Kitchenham & Charters, 2007), and social psychology (e.g., Cooper, 1982) have developed a strong tradition of review methods. However, what constitutes a rigorous and well-conducted review has received little attention so far. The few available guidelines either focus on a specific type of review (e.g., theoretical or systematic) or fail to provide comprehensive and specific recommendations. Consequently, in this paper, we fill this gap by helping researchers, reviewers, and editors to properly assess IS reviews' methodological rigor. While our framework covers all types of standalone reviews (see Section 2), it also highlights the guidelines that apply to particular forms of research syntheses. Ultimately, we provide a baseline from which researchers can use to evaluate and properly conduct standalone literature reviews in the IS field.

In Section 2, we describe the major phases and steps that are at the basis of standalone reviews and present four distinct categories of review papers with their objectives, characteristics, and challenges. In Section 3, we propose a set of guidelines for conducting rigorous IS literature reviews and specify to which review forms they apply. We believe that carefully considering these guidelines is likely to help the broader IS research community make informed judgments about review papers' contributions to a cumulative tradition. Finally, in Section 4, we conclude the paper.

2   Background

2.1   General Procedure for Conducting Literature Reviews

As a preliminary step, we sought out existing guidelines for standalone reviews. As expected, our search resulted in few sources from the IS field. Indeed, "the rigorous, standardized methodology that has developed from the health sciences and other fields is virtually unknown in information systems research" (Okoli & Schabram, 2010, p. 6). However, several relevant and useful sources emerged from other fields that have a stronger tradition in review methods and evidence-based practice. These include the health sciences and health information domains (e.g., Liberati et al., 2009; Higgins & Green, 2008), software engineering (e.g., Kitchenham & Charters, 2007), social psychology (e.g., Cooper, 1982; Cooper, 2009), and management and organization science (e.g., Rousseau et al., 2008; Wolfswinkel, Furtmueller, & Wilderom, 2013).

According to Kitchenham and Charters (2007), the process of conducting an effective literature review comprises several discrete activities. The authors argue that literature reviews involve stages that can be grouped into three main phases: (1) planning, (2) carrying out, and (3) reporting. The activities associated with planning include choosing a research question and developing a protocol. The second phase involves many activities, such as searching for and selecting pertinent and available documents on the topic, assessing the quality of the selected studies, and extracting, analyzing, and synthesizing the data. The last phase encompasses writing the review and developing strategies for disseminating the results to potentially interested parties. In this paper, we limit our analysis to the first two phases.

According to Kitchenham and Charters (2007), most medical and social science methodologists agree about the major steps involved in conducting a literature review. The six steps that comprise our framework reflect the main activities involved in developing and conducting a research synthesis: (1) formulating the problem, (2) searching the literature, (3) screening for inclusion, (4) assessing quality, (5) extracting data, and (6) analyzing and synthesizing data (see Figure 1). Note that the proposed sequence of steps need not be followed in a linear manner. Indeed, another important trait of the review process is its iterative nature: many activities are initiated during the planning phase and later refined during subsequent phases (Kitchenham & Charters, 2007). We describe the six steps in the following paragraphs.

Formulating the problem: authors must justify the need for a standalone literature review (Kitchenham & Charters, 2007; Webster & Watson, 2002), identify the review's purpose (Okoli & Schabram, 2010), and define the concepts or constructs at the heart of the synthesis (Cooper, 2009; Webster & Watson, 2002). They need to specify which research questions they are addressing (Kitchenham & Charters, 2007). As Jesson et al. (2011) stress, the research questions guide the entire study design because they underscore the type of information that is needed, inform the search for and selection of relevant literature, and guide the subsequent analysis.

Searching the literature: the data in a literature review includes the relevant information included in each primary study (Cooper, 2009). Before selecting and extracting data, the first step in collecting data involves searching through the literature. At this time, researchers identify a range of information sources and single out the studies that require further analysis. Primarily, researchers need to choose a population of studies (Cooper, 1982). According to Webster and Watson (2002), the literature search "ensure[s] that (reviewers) accumulate a relatively complete census of relevant literature" (p. xvi). When conducting a literature review, researchers want their results to pertain to all the extant studies on the problem (Cooper, 1982). Therefore, the specified search strategy and procedures, data sources, and search restrictions are critical if the researcher is to identify and retrieve all the relevant studies.

Screening for inclusion: the data collection process in a literature review involves both identifying primary studies and evaluating their applicability (Levy & Ellis, 2006a; vom Brocke et al., 2009). Indeed, once a group of potential primary studies has been identified, researchers must analyze them to determine their relevance (Kitchenham & Charters, 2007). A set of rules and selection criteria will provide a basis for including or excluding certain studies. This exercise requires a significant investment on the part of researchers, who must ensure enhanced objectivity and avoid mistakes (Liberati et al., 2009).
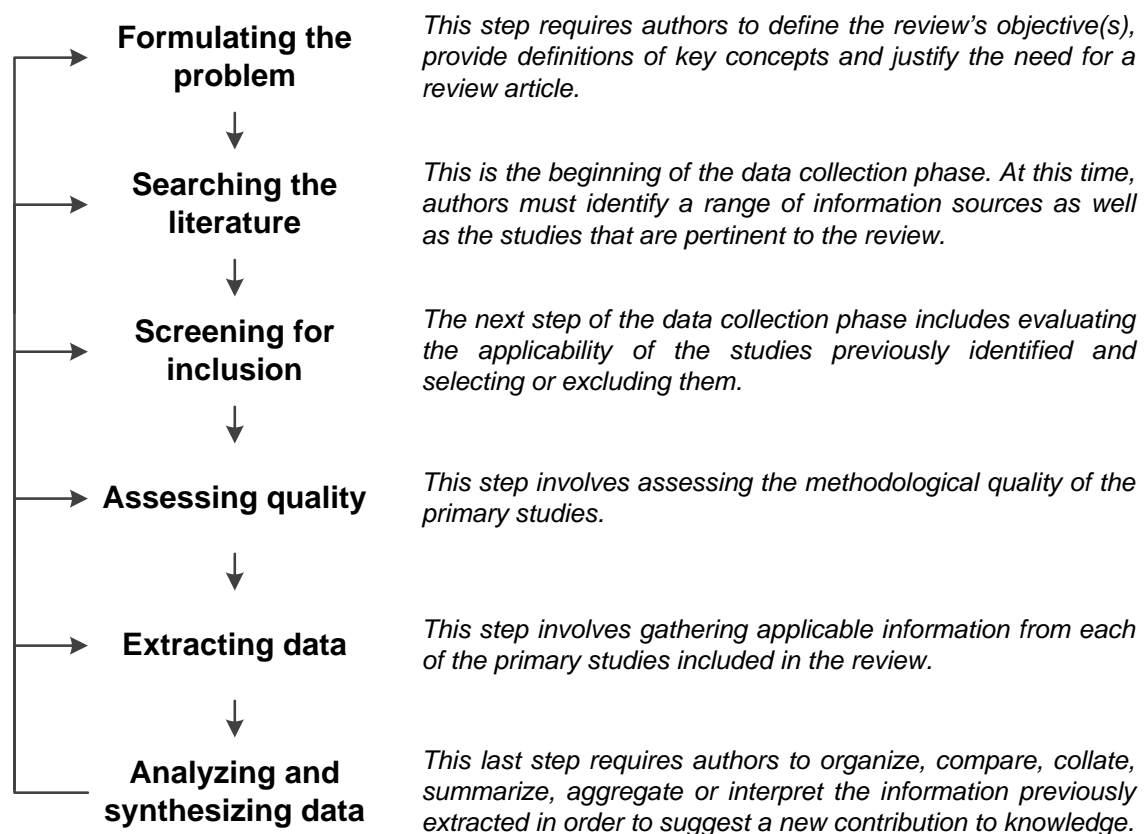
Formulating the problem	*This step requires authors to define the review's objective(s), provide definitions of key concepts and justify the need for a review article.*
↓	
Searching the literature	*This is the beginning of the data collection phase. At this time, authors must identify a range of information sources as well as the studies that are pertinent to the review.*
↓	
Screening for inclusion	*The next step of the data collection phase includes evaluating the applicability of the studies previously identified and selecting or excluding them.*
↓	
Assessing quality	*This step involves assessing the methodological quality of the primary studies.*
↓	
Extracting data	*This step involves gathering applicable information from each of the primary studies included in the review.*
↓	
Analyzing and synthesizing data	*This last step requires authors to organize, compare, collate, summarize, aggregate or interpret the information previously extracted in order to suggest a new contribution to knowledge.*

Figure 1. General Procedure for Conducting Literature Reviews

Assessing quality: in addition to screening studies for inclusion, researchers may need to formally assess the studies' quality to refine what ones they include or not, determine whether or not the differences in quality affect their review's results, or guide how they analyze the data and interpret the findings (Kitchenham & Charters, 2007). Assessing quality mostly pertains to appraising the research design and methods used in the primary studies. Indeed, researchers must assess these studies against recognized methodological standards (Jesson et al., 2011; Kitchenham & Charters, 2007). Ascribing quality scores to each study makes it possible to reflect on the extent to which the selected paper address possible biases and maximize validity (Kitchenham & Charters, 2007).

Extracting data: this step involves gathering applicable information from each primary study included in the review and deciding what is relevant to the problem of interest (Cooper, 1982). Indeed, the type of

data that should be recorded depends mainly on the research question (Okoli & Schabram, 2010). However, important information may also be gathered about how the primary study was conducted, the research design and methods, or statistical results (Cooper, 2009).

Analyzing and synthesizing data: researchers must collate, summarize, aggregate, organize, and compare the evidence extracted from the primary studies. The extracted information must be presented in a meaningful way that suggests a new contribution to knowledge (Jesson et al., 2011). Researchers are also expected to interpret the cumulative evidence and discuss the findings and conclusions derived from the data (Cooper, 1982). Webster and Watson (2002) warn researchers that literature reviews should be much more than lists of papers and should provide a coherent lens to make sense of extant knowledge on a given topic.

2.2   Types of Literature Reviews

We concur with Cooper (1988) that there exist various forms of reviews that differ considerably in terms of fundamental objectives, motivations, and means by which they are conducted. In fact, considerable confusion and many ambiguities surround the use of the term "review" (Paré et al., 2015). Indeed, researchers have used numerous terms to depict their review methods and approaches, such as "narrative review" (e.g., Joseph, Ng, Koh, & Ang, 2007), "theoretical review" (e.g., Varey, Wood-Harper, & Wood, 2002), "critical review" (e.g., Fichman, 1992), "descriptive review" (e.g., Yang & Tate, 2009), "comprehensive review" (e.g., Liu, Min, & Ji, 2008), "systematic review" (e.g., Williams, Dwivedi, Lal, & Schwarz, 2009), and "meta-analysis" (e.g., King & He, 2006). For this reason, we distinguish the several types of standalone reviews before we develop a set of guidelines or criteria for assessing rigor.

We adopted an iterative approach in categorizing review types. We started out by searching and reviewing works by leading authors and methodologists. Backward and forward searches allowed us to uncover additional relevant material. Over the years, some scholars have discussed methodological issues related to research synthesis and proposed dimensions along which to distinguish between types of literature reviews (e.g., Cooper, 1988; Gough, Thomas, & Oliver, 2012; Grant & Booth, 2009; King & He, 2005; Paré et al., 2012; Rousseau et al., 2008). For instance, Cooper (1988) proposes a set of core characteristics along which the various types of review papers may differ. These characteristics include the review's main area of interest, its primary focus, the reviewers' perspective, the breadth of the literature search, the techniques used to organize the data, and the intended audience. By analyzing, comparing, and combining the descriptions and definitions from the abovementioned sources, we extracted four overarching categories of standalone reviews that we describe next. More precisely, we grouped and analyzed those dimensions that focused specifically on review aims and approaches. Indeed, because research objectives should suggest the components of the methods, they were the focus of our analysis. Hence, we developed four review categories that differ in terms of their suitability for synthesizing different forms of research findings (i.e., input), their appropriateness for answering specific research questions (i.e., process), and their intended outcomes (i.e., output). Table 1 describes the main characteristics of these four categories. Note that a particular review might pursue several objectives and, hence, share key characteristics that belong to different review types (Cooper, 1988). For instance, Joseph et al. (2007) conduct a narrative review to summarize the extant knowledge on IT turnover and propose a conceptual model of IT turnover intention that they validate through a meta-analytic review procedure. They then combine findings of the narrative review and the meta-analysis to develop a richer and more comprehensive model of IT turnover. Successfully conducting or evaluating hybrid reviews such as this one first requires a thorough understanding of the characteristics and methodological issues related to each review type involved and of their rationale. We believe the conceptual framework we introduce in this papers will help members of our research community better understand the basic principles of review methods and make more-informed choices about their designs.

Each review type should respect specific methodological requirements that are consistent with its general objective. Indeed, "clarity about the dimensions along which reviews vary provides a way to develop review methods further and to make critical judgments necessary for the commission, production, evaluation, and use of reviews" (Gough et al., 2012, p. 8). According to Oxman (1994), designing a research synthesis involves several decisions that must be aligned with the review's specific purpose and its main focus. Each review type, along with an adapted package of methodological guidelines and techniques, is indeed best suited to certain objectives, questions, or issues. Researchers widely accept that producing scientific knowledge depends very much on the tools and techniques that they apply (Pinsonneault & Kraemer, 1993). Since "better legitimization of every choice made during the review

process enhances the value of a review" (Wolfswinkel et al., 2013, p. 45), the methods and techniques used must be well suited to the review and, thus, carefully chosen. In the next several pages, we describe and illustrate each review type.

First, **narrative reviews** summarize previously published research on a topic of interest. They focus on concepts and theories, research methods, or research outcomes (Paré et al., 2012). They mainly assemble and synthesize extant literature and provide readers with a comprehensive report on the current state of knowledge in the area under investigation. Narrative reviews do not propose novel conceptualizations, criticize a body of literature, or validate a theory; instead, these review types "serve a scientific field by providing a much-needed bridge between the vast and scattered assortment of articles on a topic and the reader who does not have the time or resources to track them down" (Baumeister & Leary, 1997, p. 311). Such reviews are particularly useful for gathering a large and diverse volume of existing research on a subject (Dixon-Woods, Agarwal, Young, Jones, & Sutton, 2004). Narrative reviews often serve as an appropriate starting point for future inquiries and research developments and help researchers to determine and refine research questions or hypotheses (Cronin et al., 2008).

Very few sources offer guidelines for conducting rigorous narrative reviews (Baumeister & Leary, 1997). King and He (2005, p. 667) also note a lack of commonly accepted or standardized procedures: "researchers are relatively free to design their review strategy in terms of selecting relevant papers, categorizing research characteristics, and framing outcomes". In other words, researchers do not usually explain how they searched, selected, and synthesized their narrative reviews' primary studies (Paré et al., 2012). For this reason, narrative reviews are more prone to subjectivity than other review types. Efforts have been made to increase rigor and improve knowledge about conducting such reviews (e.g., Cronin et al., 2008; Green, Johnson, & Adams, 2006; Levy & Ellis, 2006b). Indeed, there is a growing number of well-structured narrative reviews in the IS field that present methodological details about the review process (Paré et al., 2012).

Table 1. Example of Format for Tables (Source)

		Narrative	Developmental	Cumulative	Aggregative
Input	**Primary study focus**	Allow researchers to gather studies that focus on thematically dissimilar concepts and findings		Require studies that focus on similar concepts	Require close conceptual and operational definitions
	Primary study design	Allow researchers to combine both conceptual and empirical studies with varying methods and designs			Require empirical studies that follow the same design
Process	**General objective**	Map the current state of knowledge and identify gaps in prior research	Assemble previous research to develop an innovative approach to the topic of interest	Compile cumulative evidence from earlier research in order to identify patterns and draw overall conclusions	Pool prior data and findings to test specific theories and hypotheses
	Literature coverage	Cover a representative set of the literature by including a sample that is illustrative of the larger population	Cover studies that are central or pivotal to a topic area and include a sample that considers all important aspects	Cover the literature in detail by identifying and including all pertinent data	
	Logic of synthesis	Follow a logic of configuration by drawing conclusions based on a coherent assembly of findings		Follow a logic of assimilation by confirming findings based on the repetition of evidence	
Output	**Product**	Narrative summary of knowledge in a topic area	Conceptual framework, theoretical model, methodological guidelines, etc.	Pooled summary of a specific research stream	Evidence-based validation of a theoretical model
	Domain of generalizability	Allow researchers to generalize inferences to a particular population of studies	Allow researchers to generalize theories, concepts, or new ideas to their domain of applicability	Allow researchers to generalize inferences to a particular population of studies	Allow researchers to generalize hypotheses to a pre-specified unit of analysis

Since narrative reviews summarize a research stream and, therefore, generalize their conclusions to a particular group of studies, researchers must adopt strategies for gathering a representative literature sample and follow structured procedures for analyzing and synthesizing the evidence (Levy & Ellis,

2006a). Cooper (1988) defines a representative coverage strategy as identifying and selecting earlier studies that typify a larger group. Authors of these reviews should also discuss the characteristics that make the sample illustrative of a larger population. As per Sandelowski, Voils, and Barroso's (2006) and Sandelowski, Voils, Leeman, and Crandell's (2012) definition, data analysis in narrative reviews relies on a logic of configuration. The authors distinguish between two broad categories of research syntheses based on different logics: configuration and assimilation. The configuration of findings refers to the arrangement of complementary evidence into a coherent argument (Sandelowski et al., 2006). Findings from previous studies usually address different aspects of the phenomenon of interest and are, therefore, linked and organized. Both narrative and developmental (see below) reviews mainly follow the configuration logic. The assimilation of findings reduces or averages evidence to make empirical statements (Gough et al., 2012; Sandelowski et al., 2012). Narrative reviews confirm findings through the repetition of thematically similar evidence.

Bélanger and Carter's (2012) paper represents an exemplar of a highly structured narrative review. In this paper, the authors explore the most important theories, findings, and approaches that researchers have used to study diverse electronic government platforms and services. To provide this historical synopsis, the authors examine two samples of e-government papers: the most highly cited e-government papers according to ISI Citations Index, and e-government research published in the Association for Information Systems (AIS) Senior Scholars' basket of journals. They clearly state their selection criteria and list the papers they include in distinct appendices. They also detail their coding procedures and inter-rater reliabilities for the classification of primary studies in an appendix. Their analysis of the extant literature reveals significant insights about the metamorphosis of e-government research over time. The authors' discussion of the gaps and opportunities for further work provides researchers with a starting point to investigate e-government phenomena.

Second, **developmental reviews** provide a research community with new conceptualizations, research models, theories, frameworks or methodological approaches. In general, they develop innovative ideas or approaches grounded in previous research from a particular body of knowledge. Developmental reviews are particularly valued for proposing new theoretical foundations, for developing new approaches to address existing research problems on a more mature topic, and for giving directions for further improvements (Webster & Watson, 2002). Developmental reviews usually adopt a theoretical (Webster & Watson, 2002) or a critical stance (Carnwell & Daly, 2001). As such, developmental reviews could propose new conceptualizations or theoretical approaches or could critically analyze previous knowledge and offer constructive information on problematic areas. Hence, the primary contribution of developmental reviews usually lies in the novelty of the proposed ideas and, therefore, goes beyond simply gathering and synthesizing prior studies.

As is the case with narrative reviews, authors of developmental reviews rarely explain the methods they used (Wolfswinkel et al., 2013). Guidelines for conducting rigorous developmental reviews have, however, been proposed in an effort to increase thoroughness and transparency (Webster & Watson, 2002). Examples of developmental approaches include realist synthesis (Pawson, Greenhalgh, Harvey, & Walshe, 2005), grounded theory review (Wolfswinkel et al., 2013), integrative review (Torraco, 2005), meta-synthesis (Hoon, 2013) and meta-ethnography (Noblit & Hare, 1988). Rigorous developmental reviews offer greater contributions and have a better chance of being published (Webster & Watson, 2002; Wolfswinkel et al., 2013). According to Torraco (2005), authors of developmental reviews are expected to identify a topic or issue that is appropriate for review, search and retrieve the most significant and relevant literature on that topic, analyze or critique the extant literature, and propose new concepts, theories, frameworks, or perspectives through one or more forms of synthesis.

Developmental reviews are usually highly iterative in nature. Authors start with a broad research topic that they will refine into a more nuanced question as evidence from studies informs the topic (Hoon, 2013). Thus, they can select and adapt the components of their methods as they develop the review (Gough et al., 2012). In the IS domain, Webster and Watson (2002) propose a method for structuring and presenting developmental reviews. The process is centered on two key methodological elements; namely, a structured search strategy for identifying relevant studies, and a concept-centric approach to analyzing, contextualizing, and synthesizing prior works. Developmental reviews usually adopt a central search strategy in to include both empirical and conceptual studies that consider all the important aspects of the phenomenon of interest. Cooper (1988) defines central or pivotal coverage as the identification and description of important efforts that provide direction to a field. From these seminal research works, reviewers might extend their search to accumulate a relatively complete census of the extant literature

(Webster & Watson, 2002). As with narrative reviews, developmental reviews follow a logic of configuration for analyzing diverse and complementary evidence. They may use various approaches and classifying techniques to coherently organize and make sense of the diverse streams that emerge from the primary studies (Hoon, 2013; Torraco, 2005; Webster & Watson, 2002). Developmental reviews emphasize the conceptual contribution of their outcomes, rather than formally assessing their findings' validity (Grant & Booth, 2009). Outcomes are usually evaluated and validated against criteria of logical reasoning and the value of the theoretical contribution.

DeLone and McLean's (1992) seminal paper on IS success provides a good example of a developmental review that adopts a theoretical stance. In this paper, the authors express the motivation that, if IS research is to make a contribution to the world of practice, it is essential to define a measure of IS success that will be used to evaluate IS policies, practice, and procedures. In recognition of this importance, they conduct a literature review of previously published empirical and conceptual studies that have attempted to measure various dimensions and factors pertaining to IS success. Taken together, these studies provide a representative sample of the work conducted in this particular domain from 1981 to 1988. Subsequently, the authors present a conceptual framework with six interrelated categories of IS success, which they use to organize the extant IS research in this area and discover patterns and commonalities. Based on this framework, they integrate the multiple dimensions of IS success that they discovered from the literature review and propose a comprehensive conceptual model of IS success to guide future research efforts. This developmental review represents one of the most highly cited papers in the IS field today.

For their part, Parker and Castleman (2009) propose a developmental review that adopts a critical stance. They evaluate the individual suitability of theories and frameworks in investigating e-business adoption in small and medium-sized enterprises (SMEs). To do so, the authors "take stock" of the available evidence in the extant literature and use a plethora of studies to illustrate that SMEs are idiosyncratic in nature when it comes to adopting e-business. Then, they critically analyze and evaluate existing theories and frameworks to examine whether this idiosyncratic nature was captured sufficiently when explaining e-business adoption decisions. Based on this critical appraisal, the authors conclude that all commonly used theories omit important aspects of small-firm idiosyncrasy and, thus, highlight the need for a new integrated framework that will more accurately predict SMEs' e-business adoption intentions. In addition, they offer preliminary ideas on this framework to direct future research efforts.

Third, **cumulative reviews** compile empirical evidence to map bodies of literature and draw overall conclusions regarding particular topics of interest. As with narrative reviews, they synthesize extant literature on a particular topic to provide readers with a comprehensive description of the current state of knowledge in the area. However, the diversity of the primary studies under scrutiny and their underlying logic of analysis are different. Specific methodological approaches for cumulative reviews have been proposed that differ in terms of the nature of the research area and the range of analysis methods they use. Such approaches include scoping review (Arksey & O'Malley, 2005; Levac, Colquhoun, & O'Brien, 2010), mapping review (Anderson, Allen, Peckham, & Goodwin, 2008; Petersen, Feldt, Mujtaba, & Mattsson, 2008), and descriptive review (King & He, 2005). When looking at emerging topics, the primary contribution of scoping or mapping reviews lies in their ability to evaluate the size and scope of available literature on a particular subject matter and inform researchers about a new area for future research (Arksey & O'Malley, 2005). When looking at more-mature areas, descriptive reviews determine, through frequency analyses, the coverage of the research field and the extent to which a body of empirical studies supports or reveals any interpretable patterns or trends with respect to pre-existing propositions, theories, methodologies, or findings (King & He, 2005; Petersen et al., 2008).

Cumulative reviews are much less iterative than the two previous types because reviewers must first identify a clear research question (Arksey & O'Malley, 2005). In fulfilling their objectives and ensuring the generalizability of results, researchers usually employ structured search methods to identify as many relevant studies as possible, and collect an exhaustive sample of published works on a topic (Arksey & O'Malley, 2005). According to Cooper (1988), researchers following an exhaustive coverage strategy should strive to include all relevant literature to base conclusions on all available materials. Further, researchers should accurately gather significant information from the chosen works. At this stage, applying a clear, replicable and consistent approach is recommended to avoid any potential outcome biases. Cumulative reviews tend to follow a logic based on assimilation; that is, they compile findings that are thematically similar. As a result, such reviews usually adopt an analytic framework: they classify techniques or thematic analyses in order to make sense of the data and comprehensively summarize

earlier evidence. By using evaluation criteria, authors of cumulative reviews extract characteristics of interest, such as publication year, research methods, data collection techniques, and direction or strength of final outcomes (e.g., positive, negative, or non-significant) from each study to produce quantitative results in the form of frequency analyses (Arksey & O'Malley, 2005; Bandara et al., 2011; King & He, 2005). In doing so, authors of cumulative reviews may claim that their findings represent the "state of the art" in a particular area or domain (King & He, 2005).

Dubé and Paré's (2003) paper on positivist IS case study research represents an exemplar of a cumulative review. These authors determine the extent to which the field had advanced in its operational use of case study methods. To do so, they performed a systematic search in seven major IS journals over a 10-year period to identify a representative number of empirical papers that employed the case study method. In all, they identified 183 case papers. Subsequently, they developed and used a coding scheme with pertinent evaluation criteria and attributes to code each paper and collectively gauge the extent to which positivist case study research in IS has or has not exploited the methodological guidelines that have been made available by leading methodologists. They formally validate the coding scheme and obtained a satisfactory inter-rater agreement rate from this process. They employed frequency analysis to inform the research community about the trends and patterns in the use of positivist case study research. As we mention earlier, while providing a critical account of prior research represents the sole objective of some developmental reviews, authorial critique can play a role in other forms of research synthesis as in the present cumulative review. Indeed, Dubé and Paré (2003) critically appraise the level of methodological rigor in prior case research and their overall assessment, which is somewhat equivocal, serves as an instrument to reflect on our progress and identify potential areas for improvement.

Lastly, **aggregative reviews** bring together prior findings and test specific research hypotheses or propositions. By rigorously collating and pooling prior empirical data, aggregative reviews are particularly valued for providing evidence-based validations of pre-specified theoretical models and propositions. As a result, such reviews support evidence-based practice, a movement that developed in the field of medicine in the early 1990s (Montori & Guyatt, 2008). Indeed, aggregative reviews have been advocated as essential to synthesizing the accumulated knowledge on a particular subject and supporting evidence-based management (Briner, Denyer, & Rousseau, 2009; Rousseau et al., 2008). These types of reviews can take three general forms; namely, systematic reviews, meta-analyses, and umbrella reviews (also known as overviews of reviews) (Higgins and Green, 2008), all of which use various qualitative or quantitative synthesis approaches and techniques (Paré et al., 2012). The distinction between systematic and meta-analytic reviews lies in the techniques used to analyze data, which are respectively qualitative or quantitative in nature. Further, whereas systematic and meta-analytic reviews collate data from previous empirical studies, overviews of reviews build on evidence available from previous systematic reviews or meta-analyses on the topic (Paré et al., 2012). Over the years, aggregative reviews have become increasingly popular across a broad spectrum of research domains including medicine, nursing, public health, medical informatics, education, and management.

Aggregative reviews follow "explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made" (Higgins and Green, 2008, p. 6). A variety of guidelines have been proposed for conducting aggregative reviews in various scientific domains including health sciences (e.g., Liberati et al., 2009; Higgins & Green, 2008; Shea et al., 2007), software engineering (e.g., Kitchenham & Charters, 2007), social psychology (e.g., Cooper, 1982), and information systems (e.g., Okoli & Schabram, 2010; Oates et al., 2012). A rigorous and consistent procedure for executing aggregative reviews involves the a priori definition of a clear set of research objectives and questions that will guide the development of a structured review protocol (Higgins & Green, 2008; Kitchenham & Charters, 2007). As with cumulative reviews, aggregative reviews require an exhaustive search and selection strategy in order to identify and gather all the relevant literature that is available (Higgins & Green, 2008; Liberati et al., 2009; Oxman, 1994). By gathering and combining close replications of prior evidence, aggregative reviews follow a logic of assimilation. They are also expected to consider potential errors, biases, or flaws in the available evidence (Higgins & Green, 2008; Kitchenham & Charters, 2007). Indeed, formal appraisal of the quality of primary studies is necessary for substantiating results. Overall, systematic reviews, meta-analyses, and umbrella reviews are recognized as instrumental in summarizing accurately and reliably vast amounts of research evidence. In short, compared to the previous types of research syntheses, aggregative reviews follow higher standards of rigor and quality because they provide an explicit, comprehensive, and reproducible account of an existing body of knowledge (Okoli & Schabram, 2010).

Turner, Kitchenham, Brereton, Charters, and Budgen (2010) offer a good example of a meta-analytic review that examines whether the technology acceptance model (TAM) is an accurate predictor of actual system use. In following a systematic review process, the authors present the narrow research questions the study addresses, the search strategy and criteria (inclusion and exclusion) they used to identify relevant empirical studies, the methods they used to assess the quality of the selected studies, and the strategy they used to extract their data. Originally, these authors planned to perform an effect-size based meta-analysis. However, due to the heterogeneity of reporting of the primary studies in terms of the TAM used or the statistical method used, they conducted a vote-counting meta-analysis instead of a full effect-size meta-analysis. From a methodological perspective, this study contains a wide array of elements (e.g., quality assessment of primary studies) that are required to ensure the clarity and transparent reporting of a quantitative systematic review. Furthermore, it highlights the importance of considering and investigating, before pooling data into a meta-analysis, the diversity of methods and measures across the included studies (i.e., heterogeneity).

3   Guidelines to Evaluate Standalone Literature Reviews

To develop a comprehensive set of methodological guidelines, we searched for recommendations on how to conduct rigorous standalone literature reviews. Our review is developmental in nature. Therefore, we followed a highly iterative search strategy in order to identify relevant studies that cover all phases of the review process and consider all types of reviews in our taxonomy. As a first step, we reviewed the reference lists of the abovementioned sources (e.g., Cooper, 1988; Gough et al., 2012; Grant & Booth, 2009; King & He, 2005; Paré et al., 2012; Rousseau et al., 2008). We selected those papers that offer practical or pragmatic guidelines on how to perform literature reviews. As expected, a rapid screening of the identified paper showed that they mainly propose guidelines for aggregative reviews. For this reason, we decided to perform a purposeful search on the other review methods. We consulted the ISI Web of Knowledge (Thomson Reuters) database and the Google Scholar search engine. We conducted the search using the terms "review", "research synthesis", and "research syntheses" in conjunction with any of the following keywords: "descriptive", "integrative", "quantitative", "critical", and "narrative". We added several additional references to our initial list. As a final step, we validated our list of papers using the backward and forward search techniques. Table 2 presents the list of key references that we used to help build our set of criteria. As far as was possible, we included trustworthy papers and papers that are considered pivotal in their own field: except for those papers published in the last three years, the papers in our list were highly cited (each had an average of 810 citations). As the table shows, we identified key references for each of the four review categories included in our framework, although we associated the majority with aggregative reviews.

As a second step, we carefully scrutinized each paper in Table 2. At first, we extracted for each step of the review process all available recommendations and pragmatic advice from the selected references. Next, from carefully examining each recommendation, we grouped the extracted guidelines under broader categories. For instance, we grouped the various recommendations regarding the search process (e.g., search databases and registers, scan important journals manually, contact experts for references, follow the backward and forward approaches, search the references from reports or key studies) under the common attribute "Use a combination of approaches and data sources". Similarly, the general guideline "Conduct parallel independent data extraction" comprises various specific recommendations (e.g., involve multiple researchers in the extraction process, assess inter-rater consistency, ensure that all researchers understand the extraction protocol, and use consensus and arbitration to resolve disagreements). In consequence, we developed a list of 19 guidelines that we divided into the different phases and steps depicted in our general framework (see Table 3).

Table 2. List of Included Papers for the Development of the Guidelines

	Narrative	Developmental	Cumulative	Aggregative
Arksey & O'Malley (2005)			X	
Bandara et al. (2011)			X	
Carnwell & Dally (2001)		X		
Cooper (2009)				X
Cronin et al. (2008)	X			
Higgins & Green (2008)				X

Table 2. List of Included Papers for the Development of the Guidelines

Kitchenham & Charters (2007)				X
Levac et al. (2010)			X	
Levy & Ellis (2006a)	X			
Liberati et al. (2009)				X
Okoli & Schabram (2010)				X
Oxman & Guyatt (1988)				X
Oxman (1994)				X
Pawson et al. (2005)		X		
Petersen et al. (2008)			X	
Shea et al. (2007)				X
Torraco (2005)		X		
Webster & Watson (2002)		X		
Whittemore (2005)				X
Wolfswinkel et al. (2013)		X		

As a third and final step, we reflected on the usefulness and necessity of each activity, or guideline, in the review process by questioning how it satisfied a specific purpose in terms of the study's methodological rigor. Cooper (1988) maintain not only that the process of reviewing the extant literature is in itself a scientific methodology, but that it should deserve the same attention to design and rigor as primary study methodologies. In fact, "each methodological decision at each stage of a synthesis may enhance or undermine the trustworthiness of its conclusion or, in common social science terms, can create a threat to the validity of its conclusion" (Cooper, 2009, p. 11). As we mention previously, the term rigor encompasses various principles and approaches that aim to minimize bias and error in the review process (Ogawa & Malen, 1991). For our own purposes, we focused on four criteria that have been proposed to reach the goal of rigor; namely, internal validity, objectivity, external validity, and reproducibility (Davies & Dodd, 2002; Morse et al., 2002; Tobin & Begley, 2004). First, we define internal validity as the extent to which the review represents accurately the phenomena it is intended to describe or explain (Hammersley, 1987). Internal validity should reflect the soundness in the choice of approach, including the decisions related to the sources searched, the keywords used, the period of time covered, the criteria used for selecting papers, or the type of data extracted. Second, we define objectivity as the extent to which a review's findings are determined by the objects of the inquiry and not by the researchers' biases and values. To claim objectivity, the review process should establish "that data and interpretations of the findings are not figments of the inquirer's imagination, but are clearly derived from the data" (Tobin & Begley, 2004, p. 392). Methods for overcoming such bias involve mainly a consistent observance of the review protocol and a parallel verification of sensitive elements of the process. Third, we define external validity as the extent to which the findings have applicability in other contexts (Beck, Keddy, & Cohen, 1994). External validity refers to the review conclusions' generalizability and involves two targets (Cooper, 1982). For narrative, cumulative, and aggregative reviews, findings should pertain to a larger population of studies comprising previous research on the topic of interest. Developmental and aggregative reviews should also generalize their findings across the unit of analysis that is of interest to the topic area; for instance, the domain of applicability of a particular theory. Whereas reviewers are constrained by the domains studied in primary research, they exert control over external validity through their choice of which sources to consider and how to search and select papers(Cooper, 1982). Lastly, reproducibility measures the extent to which the review is repeatable, traceable, and clearly documented. Reproducibility of a research synthesis is attained through thoroughly documenting the review method and process, including the search, selection, extraction, and analysis performed in the review. Reproducibility is essential to determine whether a review's findings would be the same if the process was replicated (Beck et al., 1994; Coryn, 2007).

Table 3. List of Included Papers for the Development of the Guidelines

Table 3. List of Included Papers for the Development of the Guidelines

	Narrative reviews	Developmental reviews	Cumulative reviews	Aggregative reviews
Step 1: Formulating the problem				
Specify the review's primary goal(s)	✓	✓	✓	✓
Clearly define the review's key concept(s) and establish its boundaries	✓	✓	✓	✓
Step 2: Searching the literature				
Specify the search procedure in sufficient detail			✓	✓
Use a combination of data sources and search approaches	✓	✓	✓	✓
Avoid restrictions that are not based on the research question(s)	✓	✓	✓	✓
Adopt strategies to minimize publication bias				✓
Step 3: Screening for inclusion				
Specify the screening and selection procedures in sufficient detail			✓	✓
Conduct parallel independent assessment of studies for inclusion			✓	✓
Use inclusion criteria that reflect the research question(s)	✓	✓	✓	✓
Identify and be explicit about duplicate studies			✓	✓
Include studies from reputable sources	✓	✓		✕1
Step 4: Assessing quality				
Use recognized quality assessment tools				✓
Consider the quality assessment in the selection of studies or the interpretation of the findings				✓
Step 5: Extracting data				
Specify the type of information to be extracted			✓	✓
Use a structured procedure for data extraction	✓	✓	✓	✓
Conduct parallel independent data extraction		✓	✓	✓
Step 6: Analyzing and synthesizing data				
Report the appropriate standards for the synthesis of the results				✓
Describe the logical reasoning and justifications behind the findings		✓		
Summarize the included studies in detail	✓	✓	✓	✓

3.1 Step 1: Formulating the Problem

Formulating the problem includes guidelines associated with identifying the review's purpose and its central topic's definition. This step is one of the most sensitive in the review process and goes beyond rigor. Indeed, this step also pertains to the other dimensions of review quality as indicated previously; namely, relevance and methodological coherence. Besides its methodological quality, a review's utility is evaluated against the relevance of the problem it aims to address. Even if it is rigorously conducted, a review addressing a problem that is not germane and relevant would have low value for its audience. In short, researchers must motivate the problem as being both timely and appropriate and, accordingly,

---

[1] Such a guideline is detrimental in the case of aggregative reviews and, hence, should be avoided.

justify why they conducted their review. Further, this step will guide the researchers in choosing an appropriate design and provide the structure for the next phases of the review process (Jesson et al., 2011; Levac et al., 2010). The alignment between methodological components of the review process and the research questions directly contributes to the goal of methodological coherence. Therefore, appropriately formulating the problem is required for all review categories. Nevertheless, for cumulative and aggregative reviews, formulating the problem serves the additional purpose of reproducibility because clear definitions of the problem and key concepts should help the reader understand the review process and outcomes and ensure that the review is repeatable.

3.1.1   Specify the Review's Primary Goal(s)

Defining the research objectives (or research questions) represents one of the most important steps to be taken in any study, be it an empirical study, a conceptual piece, or a review paper. Hence, researchers must satisfactorily motivate the need for their conducting a standalone review (Okoli & Schabram, 2010). As we mention earlier, standalone review papers can be undertaken to analyze the progress of a specific stream of research, to aggregate findings or reconcile equivocal results of prior studies, to review the application of one theoretical model or one methodological approach, to develop a new theory or research model, or to provide a critical account of prior research (Cooper, 1988). For instance, Jeyaraj, Rottman, and Lacity (2006) justify their aggregative review by stressing the richness and diversity of the IT-based innovation research stream and underlining the existence of conflicting and contradictory findings on the antecedents of adoption. In general, authors of a review must be explicit about the objectives they pursue since they are closely related to the form or category of review.

3.1.2   Clearly Define the Key Concept(s) and Establish the Review's Boundaries

All authors of reviews must necessarily exclude a multitude of work that lies near the boundary of their problem domain, even if they are works that other reviewers might choose to include. To solve this issue, researchers must define the key concept(s) at the heart of their review (Arksey & O'Malley, 2005; Cooper, 2009; Webster & Watson, 2002). Another characteristic is the research materials to which the review directs its main attention. Indeed, this guideline serves to distinguish relevant from irrelevant studies (Cooper, 2009). Narrative and developmental reviews usually start with a wide focus that is narrowed down throughout the process, whereas cumulative and aggregative reviews are often concerned with using predefined concepts and precise boundaries (Gough et al., 2012; Hoon, 2013). Drawing on Cooper (1988) and Whittemore (2005), literature reviews can potentially direct their focus in one or more of the following areas: research findings or outcomes, concepts, theories or research models, and research methods. Like primary goals, these areas of focus are not mutually exclusive, and some reviews might have more than one focus with varying degrees of attention.

3.2   Step 2: Searching the Literature

This step includes guidelines associated with identifying potentially relevant studies. Searching for relevant literature is critical for all review categories, though the choice about the population of studies will differ according to the review's objectives. Indeed, the literature search mainly refers to the goal of external validity. For narrative reviews, researchers should identify a sample of studies that is representative of the research stream of interest. For developmental reviews, researchers should accumulate a sample of studies that covers all important aspects of the topic of inquiry (Webster & Watson, 2002). In such reviews, researchers usually conclude the search and selection process when it reaches conceptual saturation. For their part, when conducting cumulative or aggregative literature reviews, researchers want their results to pertain to all the studies deemed relevant for the problem (Cooper, 1982; Petersen et al., 2008). Hence, searching the literature attempts to include all potential papers associated with the topic of interest. Researchers also need to thoroughly document their search process when performing cumulative and aggregative reviews to contribute to reproducibility.

3.2.1   Specify the Search Procedures in Sufficient Detail

Literature reviews draw exclusively on the information included in primary studies to provide clear answers to research questions. Therefore, searching for and identifying relevant studies is one of the most critical tasks. To ensure a high level of rigor during this process, researchers should determine and follow a rigorous search strategy (Kitchenham & Charters, 2007). The authors must clearly answer questions such as "where to search, which terms to use, which sources are to be searched, time span, and language"

(Levac et al., 2010, p. 3). Further, for cumulative and aggregative reviews, the authors must be explicit in describing their approach and justifying their decisions (Kitchenham & Charters, 2007; Okoli & Schabram, 2010). This is important to assure the reader that all the pertinent and important literature has been considered (Oxman, 1994; Okoli & Schabram, 2010). This information is also critical for replication purposes and for further updates (Liberati et al., 2009). Important information on the search procedure that should be reported includes the data sources (Kitchenham & Charters, 2007; Liberati et al., 2009), the search terms used (Kitchenham & Charters, 2007; Okoli & Schabram, 2010), and the number of hits for each source (Higgins & Green, 2008; Liberati et al., 2009). Wu and Lederer's (2009) meta-analysis on the technology acceptance model provides a good example of how to clearly specify the search and identification procedures. The authors specify the types of studies included in their review, the sources that were used to locate the studies and the detailed procedure followed to retrieve the studies. Such precisions allowed them "to maximize the number of studies, reduce the source bias, and thus increase the power of the meta-analysis" (p. 424).

3.2.2 Use a Combination of Data Sources and Search Approaches

A high-quality review should also cover all aspects of the relevant literature related to the initial research question(s) (Webster & Watson, 2002). Retrieving information from only one source is likely to not adequately answer the question(s) (Kitchenham & Charters, 2007; Liberati et al., 2009). As such, it's important to combine multiple sources and approaches to ensure the search's comprehensiveness. The commonly used sources in standalone reviews are electronic databases such as ABI/Inform, ISI Web of Knowledge, PubMed, JSTOR, and IEEE Xplore (Okoli & Schabram, 2010). Shea et al. (2007) suggest that at least two complementary electronic databases should be used to scan the extant literature on a given topic. Further, other authors recommend that electronic databases be used in combination with alternative search approaches, such as manually scanning journals and conference proceedings and contacting experts on the topic of interest (Kitchenham & Charters, 2007; Okoli and Schabram, 2010). Webster and Watson (2002) also propose that backward searches (identifying and examining references cited in an paper of interest) and forward searches (identifying and considering papers that cite a particular paper) be conducted. Turner et al. (2010) provide a good example of how multiple sources can be combined. In addition to the papers obtained through electronic searches, they include all the relevant references from a previous review on the same topic and performed a backward search.

3.2.3 Avoid Restrictions That Are Not Based on the Research Question(s)

When a review is concerned with the generalizability of its findings, then it should ideally capture all the studies of interest. Indeed, researchers "must select and justify a search strategy that is appropriate for (the) research question" (Kitchenham & Charters, 2007, p. 17). Ideally, restrictions pertaining to the publication language, statu,s and date should be avoided unless the limitations correspond to the boundaries of the review itself (Higgins & Green, 2008). When restrictions are applied, authors should report them (Liberati et al., 2009) and clearly justify them (Kitchenham & Charters, 2007).

3.2.4 Adopt Strategies to Minimize Publication Bias

Publication bias refers to the problem that significant (and supporting) results are more likely to be published than non-significant (and non-supporting) results. Publication bias is particularly harmful for the internal validity of aggregative reviews because such reviews test hypotheses and theories. Therefore, researchers should address this issue by adopting strategies such as scanning grey and unpublished literature or contacting experts on the topic of interest to locate unpublished material (Kitchenham & Charters, 2007; Liberati et al., 2009). For instance, in their meta-analysis, Wu and Lederer (2009) include doctoral dissertations and conference proceedings to minimize the risk of publication bias. They also sent an inquiry for working papers and conference proceedings to the IS community through the AISWorld mailing list.

3.3 Step 3: Screening for Inclusion

A broad and comprehensive search is likely to yield many papers that are not relevant to the research question(s) (Oxman & Guyatt, 1988). Therefore, researchers must select the appropriate papers among those retrieved during their initial search. To claim that their results have internal validity, all review types should include only those papers that are appropriate and relevant for conducting the synthesis. Researchers should ask themselves if each paper addresses the problem under inquiry and helps them

answer the research question (Pawson et al., 2005). As a proxy to appraise the quality of the included studies and thus increase their internal validity, narrative and developmental reviews might also restrict the selection of studies based on the expected quality and reputation of the sources. Further, for cumulative and aggregative reviews, researchers should make important efforts when selecting studies to enhance objectivity and avoid research bias (Liberati et al., 2009). Oxman and Guyatt (1988) recommend a systematic, explicit, and rigorous screening/selection procedure to protect against biased selection of studies. These authors show how two reviews that investigate the same question include two highly different sets of primary studies and, as a result, produce diametrically opposed conclusions. Lastly, as for the previous steps, researchers need to thoroughly document their selection process to contribute to their study's reproducibility.

3.3.1   Specify the Screening and Selection Procedures in Sufficient Detail

Researchers should be explicit about how they selected or chose studies to ensure greater transparency and allow replicability (Oxman, 1994; Okoli & Schabram, 2010; Higgins and Green, 2008). According to Kitchenham and Charters (2007), selecting studies is a multi-stage process. Researchers should first initially screen the titles and abstracts against their inclusion criteria to decide whether they are worth reading or not (Okoli & Schabram, 2010). Next, they should thoroughly examine the full papers to ensure compliance with their inclusion criteria (Higgins & Green, 2008). In addition to the screening procedure, researchers should report appropriate information on this stage, such as their inclusion criteria (Oxman & Guyatt, 1988; Liberati et al., 2009; Okoli & Schabram, 2010), the number of excluded studies at each stage with reasons for exclusion (Kitchenham & Charters, 2007; Liberati et al., 2009), and the final number of included studies (Liberati et al., 2009). Liberati et al. (2009) also advise using a diagram flow to summarize the study selection process. Hauge, Ayala, and Conradi's (2010) review on the adoption of open source software provides a good example of how to apply this guideline.

3.3.2   Conduct Parallel Independent Assessment of Studies for Inclusion

The procedure for selecting studies should minimize biases and the potential for errors of judgment (Oxman, 1994) to ensure and validate the objectivity and consistency of the inclusion process. According to Oxman and Guyatt (1988), "expert assessment of primary research studies generally results in a level of disagreement that is both extraordinary and distressing" (p. 700). As such, researchers should have their studies independently assessed to minimize their errors and subjectivity (Kitchenham & Charters, 2007; Okoli & Schabram, 2010). Disagreements should be discussed and resolved using techniques such as consensus or arbitration (Higgins and Green, 2008; Kitchenham and Charters, 2007). The level of agreement between researchers could also be measured using the Cohen Kappa statistic (Higgins & Green, 2008; Kitchenham & Charters, 2007).

3.3.3   Use Inclusion Criteria that Reflect the Research Question(s)

According to Okoli and Schabram (2010), research studies must explicitly state on what criteria they based their judgment. As Oxman (1994) mentions, "the criteria used to select studies for inclusion should be consistent with the (research) focus" (p. 649). Therefore, the criteria required for study inclusion are usually based on the study's content and refer to the research question and/or topic of interest (Cooper, 2009; Kitchenham & Charters, 2007; Liberati et al., 2009; Okoli & Schabram, 2010; Petersen et al., 2008), the theories and constructs included in primary studies (Levy & Ellis, 2006a), or the research design and methodology the primary studies follow (Okoli & Schabram, 2010). Other criteria might be based on practical considerations such as the publication status, language, years considered, and a lack of or missing information (Liberati et al., 2009; Okoli & Schabram, 2010). Using such pragmatic selection criteria reflects trade-offs between satisfactorily answering the research question and practically managing the review (Okoli & Schabram, 2010) and should be used with caution (Liberati et al., 2009). In conducting aggregative reviews, which validate theories by repeating primary evidence, reviewers should first contact the primary authors of those studies that lack important information to obtain the missing information before deciding to include them or not (Oxman & Guyatt, 1988; Kitchenham & Charters, 2007). A good example of how to clearly justify the choice of inclusion criteria is provided in Robey, Im, and Wareham's (2008) review on inter-organizational information systems. The authors indicate that "because [their] primary interest is assessing theoretical trends and future directions, [they] restrict [their] review to empirical articles because they best demonstrate the influence of theoretical choices on the production of research findings" (p. 499).

3.3.4   Identify and Be Explicit About Duplicate Studies

"Duplicate publication can take various forms, ranging from identical manuscripts to reports describing different numbers of participants and different outcomes" (Higgins & Green, 2008, p. 152). Duplicate studies might also involve multiple reports of similar or different results from the same sample. Including multiple publications associated with the same data set can be harmful, especially for cumulative and aggregative reviews, which base their conclusions on repeated evidence. Therefore, researchers should find ways to identify duplicates and make proper decisions regarding their inclusion or not (Kitchenham & Charters, 2007). For instance, Ma and Liu (2004) report that they "carefully checked the sample (of the included studies) to make sure they were not based on the same data" (p. 63). Similarly, Wu and Lederer (2009) "ensured the uniqueness of each study by carefully comparing its description and statistical data with those others" (p. 424). Authors usually remove duplicates from their sample, although some choose to average results from publications pertaining to the same study.

3.3.5   Include Studies from Reputable Sources

This guideline is specific to narrative and developmental reviews because they adopt coverage strategies that emphasize important contributions to the topic of interest. Contrarily, this guideline should be avoided in the case of aggregative reviews because it could introduce a form of publication bias. Authors of cumulative reviews could apply such restrictions depending on their research questions. Since narrative and developmental reviews usually do not formally assess the quality of the primary studies they include (see step 4), they must find other ways to ensure the quality of the evidence on which they base their conclusions (Levy & Ellis, 2006a). For instance, authors often recommend that researchers should limit their search to top-tier journals and conferences (Levy & Ellis, 2006a). In the same vein, Webster and Watson (2002) suggest starting the search process with leading journals, where the major contributions are most likely to be found.

3.4   Step 4: Assessing Quality

Whereas the previous step deals mainly with the relevance of primary studies, this step includes guidelines associated with evaluating the rigor of the included papers. In this paper, the quality assessment process refers to a formal evaluation of the methodological quality as it has been defined by leading methodologists (e.g., Higgins & Green, 2008; Kitchenham & Charters, 2007; Shea et al., 2007). Over the years, the process of quality appraisal has emerged as a formal and recommended guideline, particularly for certain types of aggregative reviews such as systematic reviews and meta-analyses (Whittemore, 2005). Indeed, in order to claim internal validity of their findings, aggregative reviews require high-quality primary evidence that they pool to validate theories and hypotheses. In these reviews, quality assessment usually refers to two main issues. First, internal validity is the extent to which the design, methods, execution, and analysis of the individual studies minimize or avoid potential sources of bias (Higgins & Altman, 2008). Second, external validity refers to "the extent to which the results of a study provide a correct basis for generalizability to other circumstances" (Jüni, Altman, & Egger, 2001, p. 42). If individual studies included in a review contain methodological deficiencies or invalid results due to systematic errors and flawed designs, then these studies might distort the findings of the review and introduce bias in its conclusions (Liberati et al., 2009; Oxman & Guyatt, 1988). For aggregative reviews, quality assessment usually leads to the exclusion of primary studies or guides researchers when interpreting their results. Other types of standalone reviews are often implicit regarding the evaluation of the methodological quality of their primary studies, Yet, assessing the quality of such reviews may, in fact, provide a valuable aspect of the narrative or descriptive component of a review or a lens for interpreting their findings.

3.4.1   Use Recognized Quality Assessment Tools

According to Oxman and Guyatt (1988), "important aspects of the design and conduct of each primary study should be critiqued and the standard used in these critiques made explicit" (p. 700). When conducting aggregative reviews, researchers should assess the quality of the primary studies by using recognized assessment tools and checklists such as those proposed by Pinsonneault and Kraemer (1993) for survey research, Dubé and Paré (2003) for positivist case research, and Paré, Cameron, Poba-Nzaou and Templier (2013) for ranking-type Delphi studies. Turner et al. (2010) and Hauge et al. (2010) provide good examples of how to perform such quality assessment. Both reviews rely on checklists that cover

various aspects related to the definition of the variables and their respective measures, the description of the research method, and the reporting of the results.

3.4.2   Consider the Quality Assessment in the Selection of Studies or the Interpretation of the Findings

Okoli and Schabram (2010) suggest that the quality-assessment process might serve two non-mutually exclusive purposes. First, the results might assist researchers when selecting primary studies. For instance, a researcher could apply a minimum quality threshold for including papers for further analysis. Second, researchers might consider the potential impact of methodological quality on their review's findings. In fact, researchers recommend investigating whether quality differences provide explanations for variations in results (Kitchenham & Charters, 2007). Quality scores might also serve to moderate the results of individual studies when aggregating those results and further guide the interpretation of the findings and the recommendations for future research (Kitchenham & Charters, 2007).

3.5   Step 5: Extracting Data

This step includes guidelines for the gathering and extracting applicable information from each study. According to Bandara et al. (2011), this step includes determining what to capture and how to capture things effectively. All categories of reviews are concerned with this step because the outcome of extracting data is the primary material for analysis, which needs to be accurate and meaningful (Kitchenham & Charters, 2007). Cooper (2009) suggests that errors or bias in extracting data can lead to a misrepresentation of the studies in the following analysis and, therefore, decrease the conclusions' internal validity. When extracting data, authors of developmental, cumulative, and aggregative reviews should also avoid bias and errors and ensure consistency. Extracting data for developmental reviews usually includes some thematic coding and conceptual classification and, therefore, involves making interpretations and judgments. Confidence in the results is achieved through an objective coding and extraction procedure. Further, for cumulative and aggregative reviews, a clear description of what and how the data was extracted serves the additional purpose of reproducibility.

3.5.1   Specify the Type of Data to Be Extracted

Authors of cumulative and aggregative reviews should plan and specify the type of data they plan to extract from their primary studies (Kitchenham & Charters, 2007; Higgins & Green, 2008) and propose a strategy to collect data (Higgins and Green, 2008). Most of the recorded data usually provide evidence about the a priori research questions (Okoli and Schabram, 2010). Nevertheless, important information about how the primary study was conducted, the research design and methods, or statistical results might also be gathered (Cooper, 2009). Dubé and Paré's (2003) review paper on rigor in IS case research clearly specifies the type of information gathered from the primary studies. The authors also provide a table that lists the coded variables that they included in their analysis.

3.5.2   Use a Structured Procedure for Data Extraction

In addition to the type of information extracted, authors of cumulative and aggregative reviews should document clearly how they extracted their data for reproducibility purposes (Higgins and Green, 2008). For all review categories, using a standardized "data extraction form" (Bandara et al., 2011; Higgins & Green, 2008; Kitchenham & Charters, 2007; Okoli & Schabram, 2010), a set of "bespoke forms" (Pawson et al., 2005), a "data charting form" (Arksey & O'Malley, 2005; Levac et al., 2010), or a "classification scheme" (Petersen et al., 2008) is a widely recommended approach for structuring this process. The data extraction form should allow researchers to collect all the information needed to address their research question(s) (Kitchenham & Charters, 2007). Several alternative methods have also been proposed, such as using an indexing or summary system (Cronin et al., 2008) or a concept-centric approach that involves using a concept matrix to extract key information from the selected studies (Webster & Watson, 2002). Jeyaraj et al.'s (2006) review on IT innovation adoption and diffusion provides a good example of how to structure the data extraction process. The authors first identified a list of dependent and independent variables from the primary studies. Then, they created a coding template "in order to uniformly code the findings between independent variables and dependent variables" (p. 5). They further developed a coding scheme that helped them assign values to the relationship between variables. The authors provide practical examples of how they used the coding scheme in their review.

3.5.3 Conduct Parallel Independent Data Extraction

The procedures used for extracting data should minimize biases and judgment errors. Therefore, researchers need to extract data in a reliable and consistent manner. Kitchenham and Charters (2007) recommend that two researchers independently perform the data extraction exercise in order to minimize errors and reduce potential bias. They suggest that at least a random sample of the primary studies be cross-checked by two or more members of the research team. Further, as for the screening and selection procedure, data that researchers have extracted should be compared and disagreements discussed and resolved (Kitchenham & Charters, 2007). Also, the Cohen Kappa statistics can be used to measure the level of agreement between researchers (Higgins & Green, 2008; Kitchenham & Charters, 2007). Higgins and Green (2008) further recommend using clear instructions and decision rules about coding the data in order to facilitate the consistency of the data-extraction process.

3.6 Step 6: Analyzing and Synthesizing Data

The last step refers to using appropriate techniques to make sense of the information gathered and to appropriately reporting the review results: "During data analysis, the separate data points collected by the researcher are summarized and integrated into a unified picture" (Cooper, 2009, p. 16). Analysis and synthesis can be done following different methods and approaches. For instance, narrative and developmental reviews usually present verbal descriptions of the data contained in primary studies, whereas cumulative and aggregative reviews might complement the narrative summary with quantified data, such as a frequency analysis or more-complex statistical methods (King & He, 2005). Consequently, even though all forms of reviews attempt to synthesize prior evidence and present findings that are both valid and reliable, objectives in terms of rigor and methodological guidelines will significantly vary among the different categories of reviews. As for the previous steps, cumulative and aggregative reviews also require a thorough documentation of the analysis and synthesis process for reproducibility.

3.6.1 Report the Appropriate Standards for the Synthesis of the Results

As Cooper (2009) mentions, "rules for summarizing and integrating data from the individual studies might be inappropriate and lead to incorrect cumulative results" (p. 249). Errors in analyzing and synthesizing data are particularly detrimental for aggregative reviews that validate theories and test hypotheses. A strategy to protect the analysis and synthesis process from potential internal validity threats is to be as explicit as possible about the approaches, procedures, and assumptions for analyzing data (Cooper, 1982; Kitchenham & Charters, 2007). This also makes the review process easier to replicate (Cooper, 2009). Likewise, researchers should report their review's results in sufficient detail to allow the reader to critically assess the foundations of the authors' conclusions (Oxman & Guyatt, 1988). Liberati et al. (2009) propose a list of appropriate standards that researchers should report when presenting their review's main results. For instance, in the case of a meta-analysis, the authors should present the sample size and the estimated effects with their confidence intervals and the pooled effect estimates across studies with a confidence interval for each relationship (Liberati et al., 2009). Such information is commonly shown in a table or a forest plot (Kitchenham & Charters, 2007; Liberati et al., 2009). Authors of meta-analyses should also assess the consistency of the data using statistics such as the I^2 index (Liberati et al., 2009). In the case of narrative systematic reviews, authors should explicitly present the qualitative inferences, make their conclusions as transparent as possible, and explain conflicting results (Cooper, 2009; Liberati et al., 2009; Oxman & Guyatt, 1988). They should also undertake a sensitivity analysis; for instance, by repeating the analysis and comparing the results for subgroups of studies (Kitchenham & Charters, 2007; Liberati et al., 2009). Sensitivity analyses are much less straightforward for qualitative syntheses, but researchers could consider the possibility of conducting sub-group analyses (Kitchenham & Charters, 2007).

3.6.2 Describe the Logical Reasoning and Justifications Behind the Findings

Because developmental reviews focus mainly on creating new research concepts, models, theories, or frameworks, the authors should clearly demonstrate how key outcomes emerged from the analysis and synthesis of the extant literature. Moreover, a developmental review's outcomes are usually validated against criteria of analytical logic, creativity, and added value. As Webster and Watson (2002) mention, the reasoning and justifications behind the findings stand for a crucial part of the data analysis and synthesis process. Presenting the underlying logic and conceptual reasoning also helps the reader to follow the connections between the research purpose, the analysis and synthesis of the evidence, and the

outcome of the review (Torraco, 2005). Researchers can use various forms and methods of logical reasoning to make sense of the evidence from previous studies. For instance, Webster and Watson (2002) suggest that researchers should justify their findings and propositions by using a combination of theoretical explanations, past empirical findings, and practical examples. Torraco (2005) suggests other approaches such as using a guiding theory or competing models in order to provide a coherent structuring of the evidence. For their part, Wolfswinkel et al. (2013) recommend using grounded theory to analyze the content of the selected studies because it provides "disciplined ways of analyzing and integrating findings and insights" (p. 47).

3.6.3   Provide a Detailed Summary of the Included Studies

All forms of reviews should contain figures or tables that descriptively summarize the included studies' characteristics and findings (Higgins & Green, 2008; Levac et al., 2010; Pawson et al., 2005). Narrative, developmental, and cumulative reviews usually demonstrate both the reliability and accuracy of their findings through a clear and structured reporting strategy (Arksey & O'Malley, 2005; Levy & Ellis, 2006a; Pawson et al., 2005). Therefore, researchers should trace the usage and non-usage of primary materials (Pawson et al., 2005) and, thus, describe the chain of evidence and how they constructed their findings from the data they extracted (Levy and Ellis, 2006a). In addition, for cumulative reviews, researchers should explicitly describe all characteristics of the included studies, such as the overall number of studies, years of publication, research methods, context of the studies, or characteristics relevant to the population (Arksey & O'Malley, 2005; Levac et al., 2010). In the case of aggregative reviews, Liberati et al. (2009) recommend reporting study-level information on the main characteristics of the primary studies included in the review. Indeed, "publication of summary data from individual studies allows the analyses to be reproduced, and other analyses…to be investigated" (Liberati et al., 2009, p. 17). In addition, descriptive summaries may help in demonstrating appropriate relevance, representativeness. and generalizability pertinent to the population of studies (Cooper, 2009). As an example, DeLone and McLean's (1992) review clearly indicates which primary studies supported each dimension of their IS success framework.

4   Discussion and Concluding Remarks

In this paper, we address the growing issue of review papers' methodological rigor. After identifying four broad categories of reviews, we propose a list of 19 methodological guidelines that cover all steps of the review process and grouped them under the six following general headings: (1) formulating the problem, (2) searching the literature, (3) screening for inclusion, (4) assessing quality, (5) extracting data, and (6) analyzing and synthesizing data. Our framework is also comprehensive inasmuch as it covers all categories of reviews that are published in the IS field and also highlights those guidelines that apply to particular forms of reviews. Indeed, we contend that the set of guidelines we propose here covers all the main aspects of the review process and can be used appropriately for all forms of reviews. Note that we found only two sources that include guidelines for conducting narrative reviews. This confirms the lack of commonly accepted methods and procedures associated with this particular review type (Baumeister & Leary, 1997; King and He, 2005). At the same time, Paré et al. (2015) observes that narrative reviews are the second most important type of reviews published in leading IS journals. As such, more effort is needed to improve our collective knowledge about how to conduct and evaluate narrative reviews.

To the best of our knowledge, no prior research has attempted to develop such a comprehensive set of guidelines to evaluate standalone reviews. As a first attempt to demystify the rigor associated with review papers, our framework is broadly integrative, and we realize that particular review types (e.g., systematic review, scoping review, meta-analysis) deserve more attention. For instance, the aggregative category includes systematic reviews, vote-counting reviews, meta-analyses, and overviews of reviews. Hence, there is a need for further discussion about the various forms and nature of literature reviews in our field. Paré et al.'s (2015) recent work on the types of review papers is a valuable contribution at this point and one of the first attempts to provide a descriptive account of IS researchers' current review behaviors and practices. Next, while a comprehensive and general framework is relevant, it is also important to caution researchers that there are no sets of specific rules and recommendations that apply to all types of reviews. Therefore, it is imperative to rely on guidelines that take into account the singularity of each review type. We acknowledge that we propose a series of general guidelines and realize that not all review papers have to implement all facets of these guidelines. Future work could develop in more detail the appropriate rules and recommendations that need to be followed for each type of review.

Further, we believe our framework's value rests on two principal elements since it applies to both the process of conducting and the output of a review paper (see Table 1). As we mention previously, a method is a tool available to researchers to produce new knowledge. The Oxford Dictionary's[2] definition of a tool involves two meanings: both "a device…used to carry out a particular function" and "a thing used to help perform a job". While methodologists faithfully consider the impact of the method on the output's quality, they frequently neglect practical considerations and ease of use. However, a strong and valuable methodology should improve the rigor of research work and help ease researchers in conducting their own studies. Therefore, for each guideline, we propose specific strategies and describe available rules of thumb that can help researchers in conducting their reviews. By decomposing the whole process of reviewing into less complex and more manageable tasks, our guidelines also provides a framework that should be useful to researchers and journal reviewers and editors. We believe our work also serves an important educational purpose were it introduced to doctoral students in research methods seminars. Ultimately, by providing the IS community with a structured approach to reviewing and a better understanding of review methods, we hope to reduce the effort required in the future to produce rigorous and effective standalone literature reviews. In that respect, future work could decompose each guideline into more detail, help operationalize the guiding principles into practical strategies, and discuss the range of methodological practices available at each step. For instance, vom Brocke et al. (2009) investigate the methodological rigor specifically associated with the search process. Other contributions could focus on the remaining steps of the review process. Notably, the last step (analyzing and synthesizing) is one of the less documented. Indeed, data analysis techniques and processes are hard to evaluate because they usually involve mental activities and reasoning mechanisms that are difficult to state formally. Only recently have authors, such as Wolfswinkel et al. (2013) and Hoon (2013), proposed structured methods and approaches to rigorously analyze data from prior literature. Further developments are definitely needed in this particular area.

Regarding reviews' output, we believe a careful consideration of the guidelines listed in Table 3 is likely to enhance the overall rigor of review papers. To validate the contribution of our framework to the methodological rigor of the output, we reviewed the various guidelines available at each step and compared their appropriateness in answering specific issues of internal validity, external validity, objectivity, and reproducibility. Therefore, we hope our work will provide a baseline from which to proceed with conducting and evaluating standalone reviews in the IS field. However, while we regard the list of attributes in Table 3 to be a positive sign, we caution IS researchers not to treat this list as a cookbook recipe. While these guidelines might contribute to rigor, they do not guarantee it. Indeed, simply reporting a procedure does not necessarily guarantee that it was performed appropriately or effectively. Nevertheless, to further mature the IS field, we hope our framework will help the research community make an informed judgment about their respective contributions to a "cumulative tradition" in our field. Further research could test the applicability of our framework by exploring the current practices of IS researchers and assessing the extent to which they adopt and implement the recommended approaches and guidelines in conducting standalone reviews. Using our list of guidelines, future work could also develop and validate an instrument and apply it to the IS research in order to provide an in-depth evaluation of the methodological rigor of standalone reviews. In addition, future research could be directed toward the relevance dimension of quality. As we mention previously, a review rigorously conducted but addressing a problem that is not relevant would have low value for its audience. Standalone literature reviews play an important role in evaluating a research domain by inspiring and providing directions for further research (Webster & Watson, 2002). In this way, review papers frequently become "core" or "milestone" papers in a field (Garfield, 1982; Paré et al., 2012). Relevance is quite possibly a substantial explanation for this kind of influence of review papers on research streams. Therefore, the issue of relevance deserves further scrutiny and more careful consideration.

To conclude, we expect this research essay to contribute to the growing interest in stand-alone literature reviews and the recent trends of evidence-based management. Like Paré et al. (2015) and Rowe (2014), we believe that the enhanced role of IS review papers requires that this expository form be given careful scrutiny. We hope the framework and series of detailed guidelines proposed here will serve as a valuable framework for those interested in evaluating or properly conducting literature reviews both within and outside our field.

---

[2] http://www.oxforddictionaries.com/

Acknowledgments

References

Anderson, S., Allen, P., Peckham, S., & Goodwin, N. (2008). Asking the right questions: Scoping studies in the commissioning of research on the organisation and delivery of health services. Health Research Policy and Systems, 6(7), 1-12.

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19-32.

Atkins, C., & Louw, G. (2000). Reclaiming knowledge: A case for evidence-based information systems. Proceedings of the 8th European Conference on Information Systems, 39–45.

Bandara, W., Miskon, S., & Fielt, E. (2011). A systematic, tool-supported method for conducting literature reviews in information systems. Paper presented at the Proceedings of the 19th European Conference on Information Systems, Helsinki, Finland.

Baskerville, R. L., & Myers, M. D. (2002). Information systems as a reference discipline. MIS Quarterly, 26(1), 1-14.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. Review of General Psychology, 1(3), 311-320.

Beck, C. T., Keddy, B. A., & Cohen, M. Z. (1994). Reliability and validity issues in phenomenological research. Western Journal of Nursing Research, 16(3), 254-267.

Bélanger, F., & Carter, L. (2012). Digitizing government interactions with constituents: An historical review of e-government research in information systems. Journal of the Association for Information Systems, 13, 363-394.

Briner, R. B., Denyer, D., & Rousseau, D. M. (2009). Evidence-based management: Concept cleanup time? The Academy of Management Perspectives, 23(4), 19-32.

Card, N. A. (2012). Applied meta-analysis for social science research. New York, NY: Guilford Press.

Carnwell, R., & Daly, W. (2001). Strategies for the construction of a critical review of the literature. Nurse Education in Practice, 1(2), 57-63.

Chen, W., & Hirschheim, R. (2004). A paradigmatic and methodological examination of information systems research from 1991 to 2001. Information Systems Journal, 14(3), 197-235.

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. Knowledge in Society, 1(1), 104-126.

Cooper, H. M. (2009). Research synthesis and meta-analysis: A step-by-step approach (4th ed.). Thousand Oaks, CA: Sage.

Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. Review of Educational Research, 52(2), 291-302.

Coryn, C. L. S. (2007). The "holy trinity" of methodological rigor: A skeptical view. Journal of Multidisciplinary Evaluation, 4(7), 26-31.

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: A step-by-step approach. British Journal of Nursing, 17(1), 38-43.

Davies, D., & Dodd, J. (2002). Qualitative research and the question of rigor. Qualitative Health Research, 12(2), 279-289.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. Information Systems Research, 3(1), 60-95.

Dixon-Woods, M., Agarwal, S., Young, B., Jones, D., & Sutton, A. (2004). Integrative approaches to qualitative and quantitative evidence. London, UK: Health Development Agency.

Dubé, L., & Paré, G. (2003). Rigor in information systems positivist case research: Current practices, trends, and recommendations. MIS Quarterly, 27(4), 597-636.

Fichman, R. G. (1992). Information technology diffusion: A review of empirical research. Proceedings of the 13th International Conference on Information Systems, 195–206.

Garfield, E. (1982). ISI's "new" index to scientific reviews: Applying Research front specialty searching to the retrieval of the review literature. Current Contents, 39, 5-12.

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. Systematic Reviews, 1(28), 1-9.

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. Health Information & Libraries Journal, 26(2), 91-108.

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. Journal of Chiropractic Medicine, 5(3), 101-117.

Hammersley, M. (1987). Some notes on the terms "validity" and "reliability". British Educational Research Journal, 13(1), 73-81.

Hart, C. (1999). Doing a literature review: Releasing the social science research imagination. London, United Kingdom: Sage.

Hauge, Ø., Ayala, C., & Conradi, R. (2010). Adoption of open source software in software-intensive organizations—a systematic literature review. Information and Software Technology, 52(11), 1133-1154.

Higgins, J. P. T., & Altman, D. G. (2008). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), Cochrane handbook for systematic reviews of interventions (pp. 187–241). Chichester, UK: John Wiley & Sons.

Higgins, J. P. T., & Green, S. (2008). Cochrane handbook for systematic reviews of interventions. Chichester, United Kingdom: John Wiley & Sons.

Hoon, C. (2013). Meta-synthesis of qualitative case studies: An approach to theory building. Organizational Research Methods, 16(4), 522-556.

Jesson, J., Matheson, L., & Lacey, F. M. (2011). Doing your literature review: Traditional and systematic techniques. London, UK: Sage.

Jeyaraj, A., Rottman, J. W., & Lacity, M. C. (2006). A review of the predictors, linkages, and biases in IT innovation adoption research. Journal of Information Technology, 21(1), 1-23.

Joseph, D., Ng, K.-Y., Koh, C., & Ang, S. (2007). Turnover of information technology professionals: A narrative review, meta-analytic structural equation modeling, and model development. MIS Quarterly, 31(3), 547-577.

Jüni, P., Altman, D. G., & Egger, M. (2001). Assessing the quality of controlled clinical trials. British Medical Journal, 323(7303), 42-46.

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. Information & Management, 43(6), 740-755.

King, W. R., & He, J. (2005). Understanding the role and methods of meta-analysis in IS research. Communications of the Association for Information Systems, 16, 665-686.

Kirkevold, M. (1997). Integrative nursing research—an important strategy to further the development of nursing science and nursing practice. Journal of Advanced Nursing, 25(5), 977-984.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (EBSE Technical Report EBSE-2007-01).

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. Implementation Science, 5(69), 1-9.

Levy, Y., & Ellis, T. J. (2006a). A systems approach to conduct an effective literature review in support of information systems research. Informing Science: International Journal of an Emerging Transdiscipline, 9, 181-212.

Levy, Y., & Ellis, T. J. (2006b). Towards a framework of literature review process in support of information systems research. Paper presented at the Proceedings of the 2006 Informing Science and IT Education Joint Conference, Fort Lauderdale, FL.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Loannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. PLoS Med, 6(7), e1000100.

Liu, Z., Min, Q., & Ji, S. (2008). A comprehensive review of research in IT adoption. Proceedings of the 4th International Conference on Wireless Communications, Networking and Mobile Computing, 1-5.

Ma, Q., & Liu, L. (2004). The technology acceptance model: A meta-analysis of empirical findings. Journal of Organizational and End User Computing, 16(1), 59-72.

Montori, V. M., & Guyatt, G. H. (2008). Progress in evidence-based medicine. The Journal of the American Medical Association, 300(15), 1814-1816.

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. International Journal of Qualitative Methods, 1(2), 13-22.

Noblit, G. W., & Hare, R. D. (1988). Meta-ethnography: Synthesizing qualitative studies. Newbury Park, CA: Sage.

Oates, B. (2011). Evidence-based information systems: A decade later. Paper presented at the Proceedings of the 19th European Conference on Information Systems, Helsinki, Finland.

Oates, B., Edwards, H., & Wainwright, D. (2012). A model-driven method for the systematic literature review of qualitative empirical research. Proceedings of the 33rd International Conference on Information Systems, 1-18.

Ogawa, R. T., & Malen, B. (1991). Towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. Review of Educational Research, 61(3), 265-286.

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. Sprouts, 10(26), 1-46.

Oxman, A. D. (1994). Checklists for review articles. British Medical Journal, 309(6955), 648-651.

Oxman, A. D., & Guyatt, G. H. (1988). Guidelines for reading literature reviews. Canadian Medical Association Journal, 138(8), 697-703.

Paré, G., Cameron, A. F., Poba-Nzaou, P., & Templier, M. (2013). A systematic assessment of rigor in information systems ranking-type Delphi studies. Information & Management, 50(5), 207-217.

Paré, G., Trudel, M. C., & Jaana, M. (2012). Stand-alone literature reviews in information systems research: Development of a taxonomy of review types and assessment of current practices. Paper presented at the Proceedings of the 2012 ASAC Conference—Information Systems Division, St John's, Canada.

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.

Parker, C. M., & Castleman, T. (2009). Small firm e-business adoption: A critical analysis of theory. Journal of Enterprise Information Management, 22(1/2), 167-182.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review—a new method of systematic review designed for complex policy interventions. Journal of Health Services Research & Policy, 10(S1), 21-34.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. In Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering, 71-80.

Pfeffer, J., & Sutton, R. I. (2006). Evidence-based management. Harvard Business Review, 84(1), 62-74.

Pinsonneault, A., & Kraemer, K. L. (1993). Survey research methodology in management information systems: An assessment. Journal of Management Information Systems, 10(2), 75-105.

Reay, T., Berta, W., & Kohn, M. K. (2009). What's the evidence on evidence-based management? Academy of Management Perspectives, 23(4), 5-18.

Robey, D., Im, G., & Wareham, J. D. (2008). Theoretical foundations of empirical research on interorganizational systems: Assessing past contributions and guiding future directions. Journal of the Association for Information Systems, 9(9), 497-518.

Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in management and organizational science: Assembling the field's full weight of scientific knowledge through syntheses. The Academy of Management Annals, 2(1), 475-515.

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

Sandelowski, M., Voils, C. I., & Barroso, J. (2006). Defining and designing mixed research synthesis studies. Research in the Schools, 13(1), 29-40.

Sandelowski, M., Voils, C. I., Leeman, J., & Crandell, J. L. (2012). Mapping the mixed methods–mixed research synthesis terrain. Journal of Mixed Methods Research, 6(4), 317-331.

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. Database for Advances in Information Systems, 38(3), 29-49.

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. BMC Medical Research Methodology, 7(1), 10-17.

Tobin, G. A., & Begley, C. M. (2004). Methodological rigour within a qualitative framework. Journal of Advanced Nursing, 48(4), 388-396.

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. Human Resource Development Review, 4(3), 356-367.

Turner, M., Kitchenham, B., Brereton, P., Charters, S., & Budgen, D. (2010). Does the technology acceptance model predict actual use? A systematic literature review. Information and Software Technology, 52(5), 463-479.

Varey, R. J., Wood-Harper, T., & Wood, B. (2002). A theoretical review of management and information systems using a critical communications theory. Journal of Information Technology, 17(4), 229-239.

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the giant: On the importance of rigour in documenting the literature search process. Proceedings of the 17th European Conference on Information Systems, 2206–2217.

Watson, R. T. (2001). Introducing MISQ review—a new department in MIS Quarterly. MIS Quarterly, 25(1), 103-106.

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

Whittemore, R. (2005). Combining evidence in nursing research: Methods and implications. Nursing Research, 54(1), 56-62.

Williams, M. D., Dwivedi, Y. K., Lal, B., & Schwarz, A. (2009). Contemporary trends and issues in IT Adoption and diffusion research. Journal of Information Technology, 24(1), 1-10.

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. (2013). Using grounded theory as a method for rigorously reviewing literature. European Journal of Information Systems, 22(1), 45-55.

Wu, J., & Lederer, A. (2009). A meta-analysis of the role of environment-based voluntariness in information technology acceptance. MIS Quarterly, 33(2), 419-432.

Yang, H., & Tate, M. (2009). Where are we at with cloud computing?: A descriptive literature review. Proceedings of the 20th Australasian Conference on Information Systems, 807–819.

About the Authors

Mathieu Templier is Assistant Professor of Management Information Systems at the Faculty of Business Administration at Université Laval, Québec City, Canada. His research interests focus on the adoption and implementation of information systems, the impacts of information technology in organizations, and research methods. His work has appeared in *Information & Management*, *IEEE Transactions on Engineering Management*, *Journal of Telemedicine and Telecare*, the Hawaii International Conference on System Sciences (HICSS) proceedings and the Administrative Sciences Association of Canada Conference proceedings.

Guy Paré is Professor of Information Technologies and holder of the Research Chair in Information Technology in Health Care at HEC Montréal. His current research interests involve the implementation challenges and the impacts of electronic medical records systems and various forms of telemedicine. His publications have appeared in top-ranked journals in the fields of information systems and medical informatics. Professor Paré's expertise has been sought by several leading organizations including the World Health Organization, the Department of Health in France, Canada Health Infoway, and the Quebec Ministry of Health and Social Services. He was elected a member of the Royal Society of Canada in 2012.