Journal of Planning Education and Research
1–20

DOI: 10.1177/0739456X17723971
jper.sagepub.com
SAGE

Guidance on Conducting a Systematic Literature Review

Yu Xiao[1] and Maria Watson[1]

Abstract

Literature reviews establish the foundation of academic inquires. However, in the planning field, we lack rigorous systematic reviews. In this article, through a systematic search on the methodology of literature review, we categorize a typology of literature reviews, discuss steps in conducting a systematic literature review, and provide suggestions on how to enhance rigor in literature reviews in planning education and research.

Keywords

literature review, methodology, synthesis, typology

Introduction

Literature review is an essential feature of academic research. Fundamentally, knowledge advancement must be built on prior existing work. To push the knowledge frontier, we must know where the frontier is. By reviewing relevant literature, we understand the breadth and depth of the existing body of work and identify gaps to explore. By summarizing, analyzing, and synthesizing a group of related literature, we can test a specific hypothesis and/or develop new theories. We can also evaluate the validity and quality of existing work against a criterion to reveal weaknesses, inconsistencies, and contradictions (Paré et al. 2015).

As scientific inquiries, literature reviews should be valid, reliable, and repeatable. In the planning field, we lack rigorous systematic reviews, partly because we rarely discuss the methodology for literature reviews and do not provide sufficient guidance on how to conduct effective reviews.

The objective of this article is to provide guidance on how to conduct systematic literature review. By surveying publications on the methodology of literature review, we summarize the typology of literature review, describe the procedures for conducting the review, and provide tips to planning scholars.

This article is organized as follows: The next section presents the methodology adopted by this research, followed by a section that discusses the typology of literature reviews and provides empirical examples; the subsequent section summarizes the process of literature review; and the last section concludes the paper with suggestions on how to improve the quality and rigor of literature reviews in planning.

Research Methodology for This Study

Literature Search and Evaluation

Inclusion criterion. We only included studies that provide guidance on the methodology of conducting a literature review. Literature reviews on a specific topic were excluded from this study. We included studies from all disciplines, ranging from medical and health science to information systems, education, biology, and computer science. We only included studies written in English.

Literature identification. We started the literature search by using the keywords "how to conduct literature review", "review methodology," "literature review," "research synthesis," and "synthesis." For each manuscript, preliminary relevance was determined by title. From the title, if the content seemed to discuss the methodology of the literature review process, we obtained its full reference, including author, year, title, and abstract, for further evaluation.

We searched Google Scholar, Web of Science, and EBSCOhost, three frequently used databases by researchers across various disciplines. Because technological advancement changes methods for archiving and retrieving

Initial submission, November 2016; revised submission, February 2017; final acceptance, June 2017

[1]Texas A&M University, College Station, TX, USA

Corresponding Author:
Yu Xiao, Department of Landscape Architecture and Urban Planning, Texas A&M University, 3137 TAMU, College Station, TX 77843-3137, USA.
Email: yuxiao@tamu.edu

information, we limit the publication date to 1996 and 2016 (articles published in the past twenty years), so that we can build our review on the recent literature considering information retrieval and synthesis in the digital age. We first searched Google Scholar using broad keywords "how to conduct literature review" and "review methodology." After reviewing the first twenty pages of search results, we found a total of twenty-eight potentially relevant articles. Then, we refined our keywords. A search on Web of Science using keywords "review methodology," "literature review," and "synthesis" yielded a total of 882 studies. After initial screening of the titles, a total of forty-seven studies were identified. A search on EBSCOhost using keywords "review methodology," "literature review," and "research synthesis" returned 653 records of peer-reviewed articles. After initial title screening, we found twenty-two records related to the methodology of literature review. Altogether, three sources combined, we identified ninety-seven potential studies, including five duplicates that we later excluded.

Screening for inclusion. We read the abstracts of the ninety-two studies to further decide their relevance to the research topic—the methodology of literature review. Two researchers performed parallel independent assessments of the manuscripts. Discrepancies between the reviewers' findings were discussed and resolved. A total of sixty-four studies were deemed relevant and we obtained the full-text article for quality assessment.

Quality and eligibility assessment. We skimmed through the full-text articles to further evaluate the quality and eligibility of the studies. We deemed journal articles and books published by reputable publishers as high-quality research, and therefore, included them in the review. Most of the technical reports and on-line presentations are excluded from the review because of the lack of peer-review process. We only included very few high-quality reports with well-cited references.

The quality and eligibility assessment task was also performed by two researchers in parallel and independently. Any discrepancies in their findings were discussed and resolved. After careful review, a total of eighteen studies were excluded: four were excluded because they lacked guidance on review methodology; four were excluded because the methodology was irrelevant to urban planning (e.g., reviews of clinical trials); one was excluded because it was not written in English; six studies were excluded because they reviewed a specific topic. We could not find the full text for three of the studies. Overall, forty-six studies from the initial search were included in the next stage of full-text analysis.

Iterations. We identified an additional seventeen studies through backward and forward search. We also utilized the forward and backward search to identify literature review methods. Once the article establishing the review methodology was found, we identified best-practice examples by searching articles that had referenced the methodology paper. Examples were chosen based on their adherence to the method, after which preference was given to planning or planning-related articles. Overall, thirty-seven methods and examples were also included in this review.

Altogether, we included a total of ninety-nine studies in this research.

Data Extraction and Analysis

From each study, we extracted information on the following two subtopics: (1) the definition, typology, and purpose of literature review and (2) the literature review process. The literature review process is further broken down into subtopics on formulating the research problem, developing and validating the review protocol, searching the literature, screening for inclusion, assessing quality, extracting data, analyzing and synthesizing data, and reporting the findings. All data extraction and coding was performed using NVivo software.

At the beginning, two researchers individually extracted information from articles for cross-checking. After reviewing a few articles together, the two researchers reached consensus on what to extract from the articles. Then, the researchers split up the work. The two researchers maintained frequent communication during the data extraction process. Articles that were hard to decide were discussed between the researchers.

Typology of Literature Reviews

Broadly speaking, literature reviews can take two forms: (1) a review that serves as background for an empirical study and (2) a stand-alone piece (Templier and Paré 2015). Background reviews are commonly used as justification for decisions made in research design, provide theoretical context, or identify a gap in the literature the study intends to fill (Templier and Paré 2015; Levy and Ellis 2006). In contrast, stand-alone reviews attempt to make sense of a body of existing literature through the aggregation, interpretation, explanation, or integration of existing research (Rousseau, Manning, and Denyer 2008). Ideally, a systematic review should be conducted before empirical research, and a subset of the literature from the systematic review that is closely related to the empirical work can be used as background review. In that sense, good stand-alone reviews could help improve the quality of background reviews. For the purpose of this article, when we talk about literature reviews, we are referring to the stand-alone literature reviews.

The stand-alone literature review can be categorized by the purpose for the review, which needs to be determined before any work is done. Building from Paré et al. (2015) and Templier and Paré (2015), we group literature reviews into

four categories based on the review's purpose: describe, test, extend, and critique. This section provides a brief description of each review purpose and the related literature review types. Review methodology differentiates the literature review types from each other—hence, we use "review type" and "review methodology" interchangeably in this article. Table 1 can be used as a decision tree to find the most suitable review type/methodology. Based on the purpose of the review (column 1 in Table 1) and the type of literature (column 2), researchers can narrow down to possible review type (column 3). We listed the articles that established the specific type of review in column 4 of Table 1 and provided an example literature review in column 5.

It should be noted that some of the review types have been established and practiced in the medical sciences, so there are very few examples of literature reviews utilizing these methods in the field of urban planning and the social sciences, in general. Because the goal of this article is to make urban planners aware of these established methods to help improve review quality and expand planners' literature review toolkit, we included those rarely used but potentially useful methods in this paper.

Describe

The first category of review, whose aim is descriptive, is the most common and easily recognizable review. A descriptive review examines the state of the literature as it pertains to a specific research question, topical area, or concept. What distinguishes this category of review from other review categories is that descriptive reviews do not aim to expand upon the literature, but rather provide an account of the state of the literature at the time of the review.

Descriptive reviews probably have the most variation in how data are extracted, analyzed, and synthesized. Types of descriptive reviews include the narrative review (Green, Johnson, and Adams 2001), textual narrative synthesis (Popay et al. 2006; Lucas et al. 2007), metasummary (Sandelowski, Barroso, and Voils 2007), meta-narrative (Greenhalgh et al. 2005), and scoping review (Arksey and O'Malley 2005).

Narrative review. The narrative review is probably the most common type of descriptive review in planning, being the least rigorous and "costly" in terms of time and resources. Kastner et al. (2012) describes these reviews as "less concerned with assessing evidence quality and more focused on gathering relevant information that provides both context and substance to the authors' overall argument" (4). Often, the use of narrative review can be biased by the reviewer's experience, prior beliefs, and overall subjectivity (Noordzij et al. 2011, c311). The data extraction process, therefore, is informal (not standardized or systematic) and the synthesis of these data is generally a narrative juxtaposition of evidence. These types of reviews are common in the planning literature. A well-cited example would be Gordon and Richardson (1997), who use narrative review to explore topics related to the issue of compact development. The review is a persuasive presentation of literature to support their overall conclusions on the desirability of compact development as a planning goal.

Textual narrative synthesis. Textual narrative synthesis, outlined and exemplified by Popay et al. (2006) and Lucas et al. (2007), is characterized by having a standard data extraction format by which various study characteristics (quality, findings, context, etc.) can be taken from each piece of literature. This makes it slightly more rigorous than the standard narrative review. Textual narrative synthesis often requires studies to be organized into more homogenous subgroups. The synthesis will then compare similarities and differences across studies based on the data that was extracted (Lucas et al. 2007). Because of the standardized coding format, the review may include a quantitative count of studies that has each characteristic (e.g., nine of sixteen studies were at the neighborhood level) and a commentary on the strength of evidence available on the research question (Lucas et al. 2007). For example, Rigolon (2016) uses this method to examine equitable access to parks. The author establishes the coding format in table 2 of the article, and presents a quantitative count of articles as evidence for sub-topics of park access, such as acreage per person or park quality (e.g., the author shows that seven articles present evidence for low-socioeconomic status groups having more park acreage versus twenty studies showing that high- and mid-socioeconomic status groups have more acreage) (Rigolon 2016, 164, 166).

Metasummary. A metasummary, outlined by Sandelowski, Barroso, and Voils (2007), goes beyond the traditional narrative review and textual narrative synthesis by having both a systematic approach to the literature review process and by adding a quantitative element to the summarization of the literature. A metasummary involves the extraction of findings and calculation of effect sizes and intensity effect sizes based on these findings (more broadly known as vote counting) (Popay et al. 2006). For data extraction, findings from each included study (based on the study author's interpretation of the data, not the data itself) are taken from each study as complete sentences. These findings are then "abstracted" as thematic statements and summarized (Sandelowski, Barroso, and Voils 2007). "Effect sizes" are calculated based on the frequency of each finding; "intensity of effect sizes" is calculated by the number of findings in each article divided by the total number of findings across the literature (Sandelowski, Barroso, and Voils 2007).

To illustrate this, Limkakeng et al. (2014) used the metasummary techniques to extract themes regarding either patients' refusal or participation in medical research in emergency settings. First, a systematic search of the literature was

Table 1. Typology and Example of Literature Reviews by Purpose of Review.

Purpose	Literature Type	Review Type	Methodology Description	Example	Example Description
Describe	All literature types	Narrative review	Green, Johnson, and Adams 2001	Gordon and Richardson 1997[a]	Gordon and Richardson's (1997) example of a narrative review discusses the issue of whether or not compact cities are a desirable planning goal. The authors do not attempt to summarize the entire scope of literature, but rather identify key topics related to the research question and provide a descriptive account of the evidence in support of their conclusion.
		Textual narrative synthesis	Popay et al. 2006; Lucas et al. 2007	Rigolon 2016[a]	Rigolon's (2016) literature review of equitable access to urban parks exemplifies narrative synthesis as a result of its standard coding format (see table 2). Park access is broken down into subgroups, including park proximity, park acreage, and park quality to organize the literature.
		Metasummary	Sandelowski, Barroso, and Voils 2007	Limkakeng et al. 2014	Limkakeng et al.'s (2014) review uses metasummary techniques to extract themes regarding patients' refusal or participation in medical research in emergency settings. The prevalence and magnitude of these themes are conveyed using effect sizes and intensity sizes, and the process is clearly summarized in tables throughout the paper.
		Meta-narrative	Greenhalgh et al. 2005	Greenhalgh et al. 2004	Greenhalgh et al.'s (2004) review is categorized as a meta-narrative because of its consideration of the overarching research traditions affecting the conceptualization of innovation diffusion (see their table 1). The article then continues with a systematic review methodology to describe the literature regarding diffusion in health service organizations.
		Scoping review	Arksey and O'Malley 2005	Malekpour, Brown, and de Haan 2015[a]	Malekpour, Brown, and de Haan's (2015) study uses the scoping review methodology to broadly and systematically search literature regarding long-range strategic planning for infrastructure. The information extracted from the papers is organized and synthesized by year (see table 1); this also allows the studies to be placed in historical and academic context. In that regard, this study also exemplifies a meta-narrative.
Test	Quantitative	Meta-analysis	Glass 1976)	Ewing and Cervero 2010[a]	Ewing and Cervero's (2010) review extracted elasticities from their included studies to create weighted elasticities (their common measure of effect size in transportation studies; see tables 3–5). These are used to create broader and more generalized statements about the relationship between travel and the built environment.
	Mixed	Bayesian meta-analysis	Spiegelhalter et al. 1999; Sutton and Abrams 2001	Roberts et al. 2002	Roberts et al. (2002) used Bayesian meta-analysis to understand factors behind the uptake of child immunization. Prior probability was calculated through the opinions of the reviewers as well as the qualitative studies; posterior probability was calculated by adding the quantitative studies (see tables 1 and 2).
		Realist review	Pawson et al. 2005	S. M. Harden et al. 2015	S. M. Harden et al.'s (2015) article takes care to ask "what works for whom under what conditions" when it comes to group-based physical activity interventions. The authors separate each portion into distance tables (see tables 2 and 3) and comment on the effectiveness of this type of intervention.
	Qualitative	Ecological triangulation	Banning 2005	Fisher et al. 2014; Sandelowski and Leeman 2012[b]	Fisher et al. (2014) use ecological triangulation to provide context to research involving breast cancer risk communication between mothers and daughters. The reported thematic findings tables are structured as action-oriented statements; these are provided for topics of discussion, challenging factors, and strategies during mother-daughter interactions.

(continued)

Table 1. (continued)

Purpose	Literature Type	Review Type	Methodology Description	Example	Example Description
Extend	Qualitative	Meta-ethnography	Noblit and Hare 1988	Britten et al. 2002	Britten et al. (2002) uses meta-ethnography to draw conclusions about patients' medicine-taking behavior, and gives a nice example of creating second-order interpretations from concepts, and then extending those to third order interpretations (see table 2).
		Thematic synthesis	Thomas and Harden 2008	Neely, Walton, and Stephens 2014	Neely, Walton, and Stephens (2014) sought to answer their research question of whether young people use food practices to manage social relationships through a thematic synthesis of the literature. The authors generate analytical themes, similar to third-order interpretations, but their aim is to answer a particular research question rather than be exploratory, differentiating it from meta-ethnography.
		Meta-interpretation	Weed 2005	Arnold and Fletcher 2012	Arnold and Fletcher's (2012) article follows the meta-interpretation method very precisely (see figure 1 in their paper) and allows for iteration in the systematic review process. The authors identify sub-categories and, from there, create a taxonomy of stressors facing sport performers (see figures 3–6 in their paper).
		Meta-study	Zhao 1991; Paterson and Canam 2001	Anthony, Gucciardi, and Gordon 2016	Anthony, Gucciardi, and Gordon (2016) use meta-method, meta-theory analysis, and meta-data analysis to create an overall meta-synthesis of the development of mental toughness in sport and performance.
	Mixed	Critical interpretive synthesis	Dixon-Woods et al. 2006	Flemming 2010	Flemming (2010) does a very clear job in explaining the critical interpretive synthesis method, especially how it deviates from meta-ethnography. After coding, translating qualitative and quantitative research into each other through an integrative grid (table 5 and figure 2), and forming synthetic constructs (table 6), Flemming creates a synthesizing argument (figure 3) to examine the use of morphine to treat cancer related pain.
		Framework synthesis	Dixon-Woods 2011	Lindell and Perry 2000[a]	Lindell and Perry (2000) look to update the Protective Action Decision Model (PADM) with a new conceptual model (see figure 1) based on emerging evidence on household adjustment to earthquake hazard. The coded items are initially interpreted within the PADM framework, but are adjusted in the face of new evidence.
Critique	All literature types	Critical review	Grant and Booth 2009	Dieckmann, Malle, and Bodner 2009	Dieckmann, Malle, and Bodner (2009) provide an example of a critical review examining the reporting and practices of meta-analyses in psychology and related fields. The review compares the amassed group of literature in the subfields to a set of recommendations/ criteria for reporting and practice standards for meta-analysis.

[a]Planning or planning-related example.
[b]Although the commonly cited example for ecological sentence synthesis is by Banning (2005), Sandelowski and Leeman (2012) provide an example of the method in tables 1 and 2 of their article.

Table 2. Review Procedure by Purpose of Review.

Procedure	Describe	Test		Extend		Critique
Search the literature	Comprehensive	Comprehensive		Comprehensive		Selective and representative
Screen for inclusion	Only include articles that fit the inclusion criteria, which is based on the research question(s)					
Assess quality	Include all, rated if appropriate	Include high-quality articles		Include high-quality articles		Include all
Type of literature/data	All	Quantitative	Mixed	Qualitative	Qualitative	Quantitative/Mixed	All
Extract data	Create form or use content analysis to extract items related to the topic of interest (varying degrees of rigor)	Create form or use content analysis to extract items of interest	Extract data related to probability	Populate attribute matrices	Extract key concepts, second-order interpretations, and theoretical constructs	Extract evidence in support or in opposition to a prior conceptual model	Create criteria form or use content analysis to evaluate items of interest
Analyze and synthesize	Tables, charts, and narrative	Meta-regression	Bayesian statistical model	Contextual sentences and summary matrices/tables	Use existing theory and evidence to create third-order interpretations	Report findings and adjust conceptual model	Description and summary tables

conducted and themes were interpreted and extracted from the studies (e.g., themes favoring participation included "personal health benefit," "altruism," and "participant comfort with research"). These themes were counted as an expression of their frequency (e.g., help society/others was mentioned in nine [64 percent] papers), and papers were given an intensity score based on the number of themes included in that paper and the "intensity" of those themes (i.e., the themes in that paper common to other papers) (Limkakeng et al. 2014, 402, 406).

Meta-narrative. A meta-narrative, following the work of Kuhn (1970) on research paradigms and diffusion of innovations, distinguishes itself as a synthesis method by identifying the research traditions relevant to the research question and included studies (Greenhalgh et al. 2005). This way, "meta-narrative review adds value to the synthesis of heterogeneous bodies of literature, in which different groups of scientists have conceptualized and investigated the 'same' problem in different ways and produced seemingly contradictory findings" (Greenhalgh et al. 2005, 417). Studies are grouped by their research tradition and each study is judged (and data extracted) by criteria set by experts within that tradition. Synthesis includes identifying all dimensions of the research question, providing a description of the contributions made by each research tradition, and explaining all contradictions in context of the different paradigms (Greenhalgh et al. 2005). Greenhalgh et al.'s (2004, 587) article is categorized as a meta-narrative because of its consideration of the overarching research traditions affecting the conceptualization of innovation diffusion.

Scoping review. Similar to textual narrative synthesis, a scoping review (Arksey and O'Malley 2005) aims to extract as much relevant data from each piece of literature as possible—including methodology, finding, variables, etc.—since the aim of the review is to provide a snapshot of the field and a complete overview of what has been done. Because its goal is to be comprehensive, research quality is not a concern for scoping reviews (Peters et al. 2015). Scoping reviews can identify the conceptual boundaries of a field, the size of the pool of research, types of available evidence, and any research gaps. For example, when scoping current literature on long-range strategic planning for infrastructure, Malekpour, Brown, and de Haan (2015, 70, 72) summarize their findings based on year, research focus, approaches, methodologies, techniques for long-range planning, historical context, intellectual landscape, etc.

Test

A testing review looks to answer a question about the literature or test a specific hypothesis. A testing review can be broken into subcategories based on the type of literature being analyzed. Testing reviews of quantitative literature involve statistical analysis, whereas qualitative testing reviews look at results in various contexts to determine generalizability. Efforts have also been made to statistically combine quantitative and qualitative research in testing reviews. Types of testing reviews include meta-analysis (Glass 1976), Bayesian meta-analysis (Spiegelhalter et al. 1999; Sutton and Abrams 2001), realist review (Pawson et al. 2005), and ecological triangulation (Banning 2005).

Meta-analysis. Meta-analysis, established by Glass (1976), requires the extraction of quantitative data necessary to conduct a statistical combination of multiple studies. This includes extracting a summary statistic common to each study to serve as the dependent variable (this is usually "effect size") and moderator variables to serve as independent variables (Stanley 2001). The synthesis for a meta-analysis will include a meta-regression and an explanation of the results. For example, Ewing and Cervero (2010) used meta-analysis to test the relationship between travel variables (walking, vehicle miles traveled, transit use, etc.) and the built environment. They extracted or calculated elasticities from their included studies to create weighted elasticities (their common measure of effect size in transportation studies) (Ewing and Cervero 2010, 273–75). These were then used to create broader and more generalized statements about the entire set of studies; for example, from their analysis they concluded that "vehicle miles traveled (VMT) is most strongly related to measures of accessibility to destinations and secondarily to street network design variables" and "walking is most strongly related to measures of land use diversity, intersection density, and the number of destinations within walking distance" (Ewing and Cervero 2010, 265).

Bayesian meta-analysis. Bayesian statistics have recently been used to include qualitative studies in meta-analysis (Spiegelhalter et al. 1999; Sutton and Abrams 2001). Bayesian meta-analysis is a unique method that relies on calculating prior and posterior probabilities to determine the importance of factors (variables) on an outcome. Experts in the field of interest record their judgment of what they believe will be the important factors on the outcome (ranked). They then review the qualitative literature and revise their ranked factors. The ranking from each reviewer for each factor creates the prior probability. Data are then coded and extracted from the quantitative literature; the prior probability is statistically combined with the quantitative evidence to create the posterior probability, thereby combining both literature types (Roberts et al. 2002).

As an example, Roberts et al. (2002) used this form of Bayesian meta-analysis to understand factors behind the uptake of child immunization. Factors with the highest prior probabilities were, in order or importance, "lay beliefs about immunization," "advice from health professions," "child's health," and "structural issues"; after including the quantitative studies, "child's health" was the highest probability,

followed by "lay beliefs," "advice from health professionals," and "structural issues" (see table 2 in Roberts et al. 2002, 1598). Mays, Pope, and Popay (2005, 15) also summarize the Bayesian meta-analysis process in their paper.

Realist review. A realist review is commonly used to evaluate policy in practice and looks to answer the question of what works for whom, under what circumstances/conditions, and how. Summary sentences, therefore, will have a format such as "if A, then B' or 'in the case of C, D is unlikely to work'" (Pawson et al. 2005, 24). Although many reviews call for a standardized form for extracting data, the variety of literature types and the "many-sided hypothesis" of a realist review may lend itself better to more flexible extraction forms. For instance, Pawson suggests completing different sections of an extraction form for different sources or simply highlighting relevant sentences (Pawson et al. 2005, 30).

This type of review may be very useful for planners evaluating policies that may have differential or inequitable impacts, especially if the underlying mechanism or mediating factors are unclear. S. M. Harden et al. (2015, 2) explain, "While the *a priori* aim of measuring effectiveness of interventions is warranted, such an approach is typically insufficient for knowledge translation within complex systems." Their realist review examined "the environmental (e.g., location), situational (e.g., context), and implementation (e.g., delivery agent) factors" that influence the success of promoting exercise through group dynamics (S. M. Harden et al. 2015, 2). They organized the synthesis and presentation of the literature based on these factors; their summary tables clearly demonstrate for whom the interventions are successful, the conditions under which the intervention is successful, *how* the intervention is successful, and the intervention in context of the findings (S. M. Harden et al. 2015, 6–11).

Ecological triangulation. Very similar to a realist review, ecological triangulation looks to answer the question of what works for what persons, under what conditions, with what outcomes (Banning 2005). Data extraction, therefore, is guided by these questions and organized in a matrix of study attributes (such as study participant on one axis and the contextual study attributes on the other). Ecological triangulation is also like a meta-study in that meta-method, meta-data, and meta-theory should be considered. The focus of the analysis and synthesis, then, according to Banning "is to determine what evidence across cases (articles) do theory, method, and the analysis of persons and conditions support interventions with positive results" (2005, 1). This is sometimes referred to as ecological sentence synthesis. Sandelowski and Leeman (2012) give a clear example of how information in an ecological sentence synthesis can be presented in tables 1 and 2 of their paper (1408). Another example can be found in Fisher et al. (2014, 521), where they use ecological sentences to report how mothers and daughters discuss breast cancer risk.

Extend

An extending review goes beyond a summary of the data and attempts to build upon the literature to create new, higher-order constructs. This category of review lends itself to theory-building. Like the testing review, there are several types of extending reviews based on the type of literature used in the review. For qualitative literature, often these techniques involve extracting concepts and second-order constructs from the literature and transforming them into third-order constructs. This allows studies to be translated into each other and overarching hypotheses and concepts to be explored. However, because of this, not all literature can be included in this type of review—studies must be similar enough to be able to be synthesized and not lose the integrity of the individual study (Mays, Pope, and Popay 2005).

Extending reviews are often done through qualitative or mixed literature because of their theory-building nature. The qualitative methods are under a larger umbrella of what the literature refers to as "meta-synthesis" (Korhonen et al. 2013; Ludvigsen et al. 2016). Types of extending reviews include meta-ethnography (Noblit and Hare 1988), thematic synthesis (Thomas and Harden 2008), meta-interpretation (Weed 2005), meta-study (Zhao 1991; Paterson and Canam 2001), critical interpretive synthesis (Dixon-Woods et al. 2006), and framework synthesis (Dixon-Woods 2011).

Meta-ethnography. Meta-ethnography, presented by Noblit and Hare (1988), has seven major steps: getting started, deciding what is relevant to the initial interest, reading the studies, determining how the studies are related, translating the studies into one another, synthesizing the translations, and expressing the synthesis. The authors suggest creating a list of concepts in each study and juxtaposing them to understand their relationship. From there, the studies are translated into each other. Noblit and Hare (1988) explain, "Translations are especially unique syntheses, because they protect the particular, respect holism, and enable comparison. An adequate translation maintains the central metaphors and/or concepts of each account in their relation to other key metaphors or concepts in that account" (28). This is done using three techniques: reciprocal translational analysis (similar to contents analysis, concepts from each study are translated into one another), refutational synthesis (identify contradictions, create refutations, and attempt to explain them), and line of argument synthesis (similar to the constant comparative method, findings from each study are used to create a general interpretation) (Dixon-Woods et al. 2005). Meta-ethnography has evolved from a method of combining qualitative research to a literature synthesis method (Barnett-Page and Thomas 2009). This technique is listed first because several of the authors in this section have built upon meta-ethnography or deviated from it in some way to create their own methods.

Britten et al. (2002, 210) operationalize meta-ethnography to examine the influence of people's belief about the

meaning of medicine on their medicine-taking behavior and interaction with healthcare professionals. Meta-ethnography is often done using Schutz's (1962) idea of first-order constructs (everyday understandings, or participants' understandings) and second-order constructs (interpretations of first-order constructs, usually done by the researchers); third-order constructs are therefore the synthesis of these constructs into a new theory (Schutz 1962). Tables 1 and 2 in Britten et al. (2002) outline the processes of extracting first- and second-order constructs from the literature and generating third-order constructs to conduct a meta-analysis. First, the papers were read to understand the main concepts (behaviors) in the body of literature (in this case, adherence/compliance, self-regulation, aversion, alternative coping strategies, and selective disclosure). These are first-order constructs. Each paper was then coded and an explanation/theory for the behaviors was given by the reviewers for each (e.g., "patients carry out a 'cost-benefit' analysis of each treatment, weighing up the costs/risks of each treatment against the benefits as they perceive them" —these are the second-order constructs. Third-order constructs are developed through the second-order constructs and a line of argument is made to synthesize the studies (see their table 2 and the first full paragraph on p. 213 in Britten et al. 2002).

Thematic synthesis. Thematic synthesis is very similar to meta-ethnography. The data extraction and synthesis process for thematic synthesis utilizes thematic analysis; themes are extracted from the literature, clustered, and eventually synthesized into analytical themes (Thomas and Harden 2008). These analytical themes, similar in their construction to third order constructs, are then used to answer the research question. In theory, this is the key difference between the two. Thomas and Harden (2008) explain, "It may be, therefore, that analytical themes are more appropriate when a specific review question is being addressed (as often occurs when informing policy and practice), and third order constructs should be used when a body of literature is being explored in and of itself, with broader, or emergent, review questions" (9). For example, Neely, Walton, and Stephens (2014) use thematic synthesis to answer their research question of how young people use food practices to manage social relationships. Unlike Britten et al. (2002), who created second-order constructs from themes present in each paper (and then looked to see how the papers were related to one another), Neely, Walton, and Stephens (2014) used all themes from all papers to create theme clusters from which they draw their conclusions about the group of papers as a whole.

Meta-interpretation. Meta-interpretation, put forward by Weed (2005), looks to improve upon the systematic review to allow it to fit within an interpretive approach in order to remain "true" to the epistemology of the synthesized research (Weed 2005). To accomplish this, a research area rather than a research question is chosen. "Maximum variation sampling" is utilized to find an initial few contrasting studies. Using a focus of "meaning in context," conceptual issues that emerge from analyzing these studies will lead to more iterations of literature selection until theoretical saturation is reached (Weed 2005). Weed (2005) notes that a "statement of applicability" must be written to clearly identify the boundaries and scope of the synthesis. Arnold and Fletcher's (2012) article follows the meta-interpretation method very precisely and allow for iteration in the systematic review process (401). They used this method to identify subcategories and, from there, create a taxonomy of stressors facing sport performers (outlined in their figure 3 and detailed in their figures 4–6).

Meta-study. Meta-study, conceived by Zhao (1991) and further operationalized by Paterson and Canam (2001), is composed of the combination of meta-data-analysis, meta-method, and meta-theory. Paterson and Canam (2001) advocate meta-ethnography for the meta-data-analysis portion, though not exclusively. Meta-method extracts methodological information from each study (including sampling and data collection) while considering the relationship between outcomes, ideology, and types of methods used; for example, Paterson and Canam describe this as exploring whether the methods were "liberating or limiting" (2001, 90). Meta-theory examines the philosophical and theoretical assumptions and orientations for each paper, the underlying paradigms, and the quality of the theory (Paterson and Canam 2001; Barnett-Page and Thomas 2009). Paterson and Canam (2001) discuss the synthesis of these three aspects as being iterative and dynamic, but take care not to offer a standardized procedure (Barnett-Page and Thomas 2009). Anthony, Gucciardi, and Gordon (2016) very clearly detail their own process of utilizing meta-study to synthesize literature on the development of mental toughness, and can be referenced as an example.

Critical interpretive synthesis. Critical interpretive synthesis (Dixon-Woods et al. 2006) arose as a modification to the meta-ethnography procedure in order to accommodate diverse literature with a variety of methodologies. Data extraction is more formal or informal based on the paper. There is no standard "quality appraisal" as a formal stage in literature review—rather, each piece of literature is judged by different criteria (based on other literature of its type), and reviewers consider the theoretical context and research traditions that could affect the evidence. It should be noted that this method modifies the entire literature review process by making it more iterative, reflexive, and exploratory (and therefore less formal and standardized), but checks and balances are instead established through utilizing a research team as opposed to relying on individual interpretations of the literature (Dixon-Woods et al. 2006).

Flemming (2010) does a very clear job in explaining the critical interpretive synthesis method, especially in how the

analytic process deviates from meta-ethnography to incorporate a mix of literature types. After coding, translating qualitative and quantitative research into each other through an integrative grid (see table 5 and figure 2 in the paper), and forming synthetic constructs, the author creates a synthesizing argument to examine the use of morphine to treat cancer-related pain (Flemming 2010). Second-order constructs reported in the literature and third-order constructs created by the reviewers (called synthetic constricts in critical interpretive synthesis) can be used equally when creating the synthesizing argument in a critical interpretive synthesis, marking a difference between this method and meta-ethnography (Dixon-Woods et al. 2006, 6).

Framework synthesis. Framework synthesis, sometimes referred to as "best fit" framework synthesis (a derivative of the method), involves establishing an a priori conceptual model of the research question by which to structure the coding of the literature (Carroll et al. 2013; Dixon-Woods 2011). The conceptual model (framework) will then be modified based on the collected evidence. Therefore, "the final product is a revised framework that may include both modified factors and new factors that were not anticipated in the original model" (Dixon-Woods 2011, 1). Although initially meant for exclusively qualitative literature, the authors believe this can be applied to all literature types. For example, the review of household hazard adjustment by Lindell and Perry (2000), although not specifically identified as a framework synthesis, interprets the findings of the review with a previously established conceptual model (the Protective Action Decision Model) and suggests how the findings required modification to the previous theory. The updated model is then presented (Lindell and Perry 2000, 489).

Critique

A critiquing review or *critical review* (Paré et al. 2015) involves comparing a set of literature against an established set of criteria. Works are not aggregated or synthesized with respect to each other, but rather judged against this standard and found to be more or less acceptable (Grant and Booth 2009; Paré et al. 2015). Data extraction will be guided by the criteria chosen by the reviewers and synthesis could include a variety of presentation formats. For example, reviewers could set a threshold for overall acceptability as a composite of the individual criterion and report how many studies meet the minimum requirement. Reviewers could also simply report the statistics for each criterion and give a narrative summary of the overall trends. Dieckmann, Malle, and Bodner (2009) provide an example of a critical review examining the reporting and practices of meta-analyses in psychology and related fields. The review compares the amassed group of literature in the subfields to a set of recommendations/criteria for reporting and practice standards for meta-analysis. Statistics for how well the literature compared against the criteria are presented in tables throughout the paper (Dieckmann, Malle, and Bodner 2009).

Hybrid Reviews

The review types/methodologies presented above can be mixed and combined in a review. It is entirely possible to create a literature review through a hybridization of these methods. In fact, Paré et al. (2015) found that 7 percent of their sample of literature review were hybrid reviews. For example, Malekpour, Brown, and de Haan (2015) is an example of a scoping review with elements of a meta-narrative—by organizing their included studies by year, they examined the historical and academic context in which the studies were published. Reviewers should not be constrained by or "siloed" into the synthesis methodologies. Rather choose elements that will best answer the research question. In the case of Malekpour, Brown, and de Haan (2015), the organization of studies chronologically and including their paradigms fits the nature of a scoping review very well; it would be a detriment to exclude that information simply because it is confined in another method.

Process of Literature Review

A successful review involves three major stages: planning the review, conducting the review, and reporting the review (Kitchenham and Charters 2007; Brereton et al. 2007). In the planning stage, researchers identify the need for a review, specify research questions, and develop a review protocol. When conducting the review, the researchers identify and select primary studies, extract, analyze, and synthesize data. When reporting the review, the researchers write the report to disseminate their findings from the literature review.

Despite differences in procedures across various types of literature reviews, all the reviews can be conducted following eight common steps: (1) formulating the research problem; (2) developing and validating the review protocol; (3) searching the literature; (4) screening for inclusion; (5) assessing quality; (6) extracting data; (7) analyzing and synthesizing data; and (8) reporting the findings (Figure 1). It should also be noted that the literature review process can be iterative in nature. While conducting the review, unforeseeable problems may arise that requires modifications to the research question and/or review protocol. An often-encountered problem is that the research question was too broad and the researchers need to narrow down the topic and adjust the inclusion criterion. Various types of reviews do differ in the review protocol, selection of literature, and techniques for extracting, analyzing, and summarizing data. We summarized these differences in Table 2. The following paragraphs discuss each step in detail.

Step 1: Formulate the Problem

As discussed earlier, literature reviews are research inquiries, and all research inquiries should be guided by research

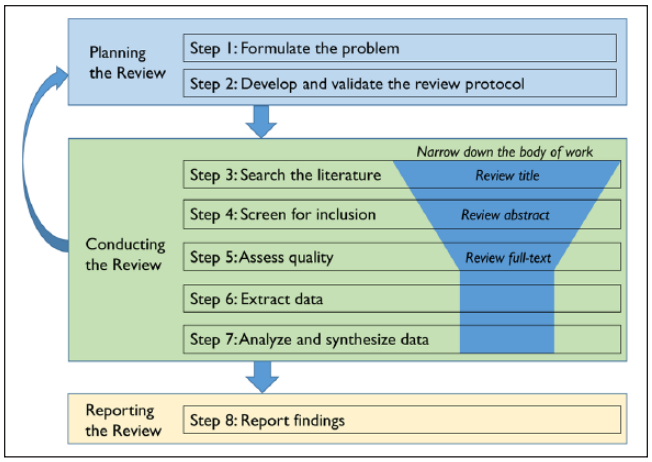

Figure 1. Process of systematic literature review.

questions. Research questions, therefore, drive the entire literature review process (Kitchenham and Charters 2007). The selection of studies to be included in the review, methodology for data extraction and synthesis, and reporting, should all be geared toward answering the research questions.

A common mistake for novices is to select too broad of a research question (Cronin, Ryan, and Coughlan 2008). A broad research question can result in a huge amount of data identified for the review, making the review unmanageable. If this happens, the researchers should narrow down the research topic—for example, choosing a subtopic within the original area for the review.

Identifying the appropriate research question can be an iterative process. Brereton et al. (2007) suggested using pre-review mapping to help identify subtopics within a proposed research question. After an initial search of literature on the research question, the researchers can conduct a quick mapping procedure to identify the kinds of research activities related to the research question, for instance, the range of subtopics, the number of studies within each subtopic, and the years the studies were carried out. Pre-review mapping helps researchers decide whether it is feasible to review the bulk of materials or they need to narrow down to a more specific research question.

Step 2: Develop and Validate the Review Protocol

The review protocol is comparable to a research design in social science studies. It is a preset plan that specifies the methods utilized in conducting the review. The review protocol is absolutely crucial for rigorous systematic reviews (Okoli and Schabram 2010; Brereton et al. 2007). It is necessary for enhancing the quality of review because it reduces the possibility of researcher bias in data selection and analysis (Kitchenham and Charters 2007). It also increases the reliability of the review because others can use the same protocol to repeat the study for cross-check and verification.

The review protocol should describe all the elements of the review, including the purpose of the study, research questions, inclusion criteria, search strategies, quality assessment criteria and screening procedures, strategies for data extraction, synthesis, and reporting (Gates 2002; Gomersall et al. 2015). Including a project timetable in the review protocol is also useful for keeping the study on track (Kitchenham and Charters 2007).

It is very important to validate the review protocol carefully before execution (Okoli and Schabram 2010; Brereton et al. 2007). In medicine, review protocols are often submitted for peer review (Kitchenham and Charters 2007). Because literature review lays the foundation for knowledge advancement, in planning education and research, we should carefully evaluate and critique the review protocols to increase the rigor of studies in our field. If possible, research teams should establish external review panels for validating their literature review protocols. Master's and doctoral students should work with their advisors to lay out and polish the review protocols before conducting the literature review. We suggest master's thesis and doctoral dissertation committees review students' literature review protocols as part of the proposal defense.

Step 3: Search the Literature

The quality of literature review is highly dependent on the literature collected for the review—"Garbage-in, garbage-out." The literature search finds materials for the review; therefore, a systematic review depends on a systematic search of literature.

Channels for literature search. There are three major sources to find literature: (1) electronic databases; (2) backward searching; and (3) forward searching.

Nowadays, electronic databases are a typical first stop in the literature search. Electronic databases constitute the predominant source of published literature collections (Petticrew and Roberts 2006). Because no database includes the complete set of published materials, a systematic search for literature should draw from multiple databases. Web of Science, EBSCO, ProQuest, IEEE Xplore are among the typically used databases in urban planning. Google Scholar is a very powerful open access database that archives journal articles as well as "gray literature," such as conference proceedings, thesis, and reports. Norris, Oppenheim, and Rowland (2008) compared Google and Google Scholar to other open access search engines such as OAIster and OpenDOAR. They found that Google and Google Scholar performed the best. These two sources combined were capable of finding more than three-quarters of open access publications identified in their study, although the performance of search engines varies across disciplines. Moreover, with an increase in open access journal publications, one can search the Directory of Open Access Journals to find open access articles. It should be noted that not all publications have been digitized and are

findable online. If the study requires literature published before the Internet age, going through the archive at the library is still necessary.

To obtain a complete list of literature, researchers should conduct a *backward search* to identify relevant work cited by the articles (Webster and Watson 2002). Using the list of references at the end of the article is a good way to find these articles.

Also should be conducted is a *forward search* to find all articles that have since cited the articles reviewed (Webster and Watson 2002). Search engines such as Google Scholar and the ISI Citation Index allow forward search of articles (Levy and Ellis 2006).

One can also perform backward and forward searches by author (Levy and Ellis 2006). By searching the publications by the key authors who contribute to the body of work, the researchers can make sure that their relevant studies are included. A search of the authors' CVs, Google Scholar pages, and listed publications on researcher's network such as ResearchGate.net are good ways to find their other publications. Contacting the authors by email and phone is an alternative approach.

Finally, consulting experts in the field has been proposed as a way to evaluate and cross-check the completeness of the search (Petticrew and Roberts 2006; Okoli and Schabram 2010). Going through the list generated from the searches, one can identify the scholars who make major contributions to the body of work. They are the experts in the field. Also, it is often useful to find the existing systematic reviews as a starting point for the forward and backward searches (Kitchenham and Charters 2007).

Keywords used for the search. The keywords for the search should be derived from the research question(s). Researchers can dissect the research question into concept domains (Kitchenham and Charters 2007). For example, the research question is "what factors affect business continuity after a natural disaster?" The domains are "business," "continuity," and "natural disaster."

A trial search with these keywords could retrieve a few documents crudely and quickly. For instance, a search of "business" + "continuity" + "natural disaster" on Google Scholar yielded lots of articles on business continuity planning, which do not answer the research question. This tells us we need to adjust the keywords.

One can also take the concepts in the search statement and extend them by synonyms, abbreviations, alternative spellings, and related terms (Rowley and Slack 2004; Kitchenham and Charters 2007). The synonyms of "business" can be "enterprise" and "firm." The terms related to "continuity" are "impact," "recovery," and "resilience"/ "resiliency." "Natural disaster" can be further broken down into "flood," "hurricane," "earthquake," "drought," "hail," "tornado," etc.

Many search engines allow the use of Boolean operators in the search. It is important to know how to construct the search strings using Boolean "AND" and "OR" (Fink 2005). Oftentimes, ''AND'' is used to join the main terms and ''OR'' to include synonyms (Breretona et al. 2007). Therefore, a possible search string can be—("business" OR "firm" OR "enterprise") AND ("continuity" OR "impact" OR "recovery" OR "resilience" OR "resiliency") AND ("natural disaster"). One can also search within the already retrieved result to further narrow down to a topic (Rowley and Slack 2004).

There are a few things to consider when selecting the correct keywords. First, researchers should strike a balance between the degree of exhaustiveness and precision (Wanden-Berghe and Sanz-Valero 2012). Using broader keywords can retrieve more exhaustive and inclusive results but more irrelevant articles are identified. In contrast, using more precise keywords can improve the precision of search but might result in missing records. At this early stage, being exhaustive is more important than being precise (Wanden-Berghe and Sanz-Valero 2012).

Second, researchers doing cross-country studies should pay attention to the cultural difference in terminology. For instance, "eminent domain" is called "compulsory acquisition" and "parking lot" called "car park" in Australia and New Zealand. "Urban revitalization" is typically called "urban regeneration" in the United Kingdom. The search can only be successful if we use the correct vocabulary from the culture of study.

Third, Bayliss and Beyer (2015) brought up the issue of the evolving vocabulary. For example, the interstate highway system was originally called "interstate and defense highways" because it was constructed for defense purposes in the cold war era (Weingroff 1996). The term "defense" was then dropped from the name. Therefore, researchers should be conscious of the vocabulary changes over time. In the search of literature dated back in history, one should use the correct vocabulary from that period of time.

Fourth, to know whether the keywords are working, Kitchenham and Charters (2007) suggested researchers check results from the trial search against lists of already known primary studies to know whether the keywords can perform sufficiently. They also suggested consultation with experts in the field.

Last but not least, it is very important to document the date of search, the search string, and the procedure. This allows researchers to backtrack the literature search and to periodically repeat the search on the same database and sources to identify new materials that might have shown up since the initial search (Okoli and Schabram 2010).

Sampling strategy. All literature searches are guided by some kind of sampling logic and search strategies adopted by the reviewers (Suri and Clarke 2009). The sampling and search strategies differ across various types of literature reviews. Depending on the purpose of the review, the search can be exhaustive and comprehensive or selective and representative (Bayliss and Beyer 2015; Suri and Clarke 2009; Paré

et al. 2015). For example, the purpose of a scoping review is to map the entire domain and requires an exhaustive and comprehensive search of literature. Gray literature, such as reports, theses, and conference proceedings, should be included in the search. Omitting these sources could result in publication bias (Kitchenham and Charters 2007). Other descriptive reviews are not so strict in their sampling strategy, but a good rule of thumb is that the more comprehensive the better. Testing reviews with the goal of producing generalizable findings, such as the meta-analysis or realist review, require a comprehensive search. However, they are more selective in terms of quality. Grey literature might not be employed in such syntheses because they are usually deemed inferior in quality compared to peer-reviewed studies. Reviews with the purpose of extending the existing body of work can be selective and purposeful. They don't require identification of all the literature in the domain, but do require representative work to be included. Critical reviews are flexible in their sampling logic. It can be used to highlight the deficiencies in the existing body of work, thus being very selective and purposive. It can also serve as an evaluation of the entire field, thus requiring comprehensiveness.

Refining results with additional restrictions. Other practical criteria might include the publication language, date range of publication, and source of financial support (Kitchenham and Charters 2007; Okoli and Schabram 2010). First, reviewers can only read publications in a language they can understand. Second, date range of publication is often used to limit the search to certain publication periods. We can rarely find all the studies published in the entirety of human history; even if we can, the bulk of work may be too much to review. The most recent research may be more relevant to the current situation and therefore can provide more useful insights. Lastly, in the case of health care research, researchers may only include studies receiving nonprivate funds because private funding may be a source of bias in the results (Fink 2005). This can be of concern to planners as well.

Stopping rule. A rule of thumb is that the search can stop when repeated searches result in the same references with no new results (Levy and Ellis 2006). If no new information can be obtained from the new results, the researchers can call the search to an end.

Step 4: Screen for Inclusion

After compiling the list of references, researchers should further screen each article to decide whether it should be included for data extraction and analysis. An efficient way is to follow a two-stage procedure: first start with a coarse sieve through the articles for inclusion based on the review of abstracts (described in this section), followed by a refined quality assessment based on a full-text review (described in step 5). The purpose of this early screening is to weed out articles with content inapplicable to the research question(s) and/or established criteria. At this stage, reviewers should be inclusive. That is to say, if in doubt, the articles should be included (Okoli and Schabram 2010). The overall methodology for screening is the same across different types of literature reviews.

Criteria for inclusion/exclusion. Researchers should establish inclusion and exclusion criteria based on the research question(s) (Kitchenham and Charters 2007). Any studies unrelated to the research questions(s) should be excluded. For instance, this article answers the research question of how to conduct an effective systematic review; therefore, only articles related to the methodology of literature review are included. We excluded literature reviews on specific topics that provide little guidance on the review methodology.

Inclusion and exclusion criteria should be practical (Kitchenham and Charters 2007; Okoli and Schabram 2010). That is to say, the criteria should be capable of classifying research, can be reliably interpreted, and can result in the amount of literature manageable for the review. The criteria should be piloted before adoption (Kitchenham and Charters 2007).

The inclusion and exclusion criteria can be based on research design and methodology (Okoli and Schabram 2010). For instance, studies may be restricted to those carried out in certain geographic areas (e.g., developed vs. developing countries), of certain unit of analyses (e.g., individual business vs. the aggregate economy; individual household vs. the entire community), studying a certain type of policy or event (e.g., Euclidean zoning vs. form-based codes; hurricanes vs. earthquakes), adopting a specific research design (e.g., quantitative vs. qualitative; cross-sectional vs. time-series; computer simulation vs. empirical assessment), obtaining data from certain sources (e.g., primary vs. secondary data) and of certain duration (e.g., long-term vs. short-term impacts), and utilizing a certain sampling methodology (e.g., random sample vs. convenience sample) and measurement (e.g., subjective vs. objective measures; self-reported vs. researcher-measured) in data collection. Studies might be excluded based on not satisfying any of the methodological criteria although not all criteria must be used for screening.

Screening procedure. When it comes to the overall screening procedure, many suggest at least two reviewers work independently to appraise the studies matching the established review inclusion and exclusion criteria (Gomersall et al. 2015; Kitchenham and Charters 2007; Brereton et al. 2007; Templier and Paré 2015). Wanden-Berghe and Sanz-Valero (2012) recommended that at least one reviewer be well versed on the issue of the review; however, having a nonexpert second reviewer could be beneficial for providing a fresh look at the subject matter.

The appraisal is commonly based on the *abstracts* of the studies (Brereton et al. 2007). In case the abstract does not

provide enough information, one could also read the conclusion section (Breretona et al. 2007). The individual assessment should be inclusive—if in doubt, always include the studies.

In case of discrepancies in the assessment results, which are quite common, the two reviewers should resolve the disagreement through discussion or by a third party (Gomersall et al. 2015; Kitchenham and Charters 2007; Breretona et al. 2007). Again, if in doubt, include the studies for further examination (Breretona et al. 2007).

Finally, the list of excluded papers should be maintained for record keeping, reproducibility, and crosschecking (Kitchenham and Charters 2007). This is particularly important for establishing interrater reliability among multiple reviewers (Okoli and Schabram 2010; Fink 2005).

Step 5: Assess Quality

After screening for inclusion, researchers should obtain full texts of studies for the quality assessment stage. Quality assessment acts as a fine sieve to refine the full-text articles and is the final stage in preparing the pool of studies for data extraction and synthesis. Ludvigsen et al. (2016) saw quality appraisal as a means for understating each study before proceeding to the steps of comparing and integrating findings.

Quality standards differ across various types of reviews (Whittemore and Knafl 2005). For example, quality assessment is not crucial for some types of descriptive reviews and critical reviews: descriptive reviews such as scoping reviews are concerned with discovering the breadth of studies, not the quality, and critical reviews should include studies of all quality levels to reveal the full picture. However, quality assessment is important for reviews aiming for generalization, such as testing reviews. With this said, Okoli and Schabram (2010) recognized that quality assessment does not necessarily need to be used as a yes-or-no cutoff, but rather serve as a tool for reviewers to be aware of and acknowledge differences in study quality.

There is no consensus on how reviewers should deal with quality assessment in their review (Dixon-Woods et al. 2005). Some researchers suggested that studies need to be sufficiently similar or homogenous in methodological quality to draw meaningful conclusions in review methods such as meta-analyses (Okoli and Schabram 2010; Gates 2002); others thought excluding a large proportion of research on the grounds of poor methodological quality might introduce selection bias and thus diminish the generalizability of review findings (Suri and Clarke 2009; Pawson et al. 2005). Stanley (2001) argued that differences in quality provide the underlying rationale for doing a meta-analysis; thus, they do not provide a valid justification for excluding studies from the analysis. Therefore, reviewers in a research team should jointly decide what their decision on quality assessment is based on their unique circumstance. The most important consideration for this stage is that the criteria be reasonable and defendable.

Criteria for quality assessment. The term "quality assessment" often refers to checking the "internal validity" of a study for systematic reviews (Petticrew and Roberts 2006). A study is internally valid if it is free from the main methodological biases. Reviewers can judge the quality of study by making an in-depth analysis of the logic from the data collection method, to the data analysis, results, and conclusions (Fink 2005). Some researchers also include "external validity" or generalizability of the study in the quality assessment stage (Rousseau, Manning, and Denyer 2008; Petticrew and Roberts 2006).

Ranking studies based on a checklist is a common practice for quality assessment. For example, Okoli and Schabram (2010) suggest ranking the studies based on the same methodological criteria used for inclusion/exclusion. Templier and Paré (2015) recommend using recognized quality assessment tools, for example, checklists, to evaluate research studies. Because of the differences in research design, qualitative and quantitative studies usually require different checklists (Kitchenham and Charters 2007). Checklists have been developed to evaluate various subcategories of qualitative and quantitative studies. For example, Myers (2008) produced a guide to evaluate case studies, ethnographies, and ground theory. Petticrew and Roberts (2006) provided a collection of checklists for evaluating randomized controlled trials, observational studies, case-control studies, interrupted time-series, and cross-sectional surveys. Research institutions, such as the Critical Appraisal Skills Programme (CASP) within the Public Health Resource Unit in United Kingdom and Joanna Briggs Institute (JBI), also provide quality checklists that can be adapted to evaluate studies in planning.

The ranking result from quality assessment can be used in two ways. One is to "weight" the study qualitatively by placing studies into high, medium, and low categories (Petticrew and Roberts 2006). One should then rely on high-quality studies to construct major arguments and research synthesis before moving on to the medium-quality studies. Low-quality studies can be used for supplement, but not be used as foundational literature. The other way to use quality assessment rankings is to "weight" each study quantitatively. For example, in a meta-analysis, one can run the regression analysis using quality scores as "weights"—this way, the higher-quality work gets counted more heavily than the lower-quality work (Petticrew and Roberts 2006; Haddaway et al. 2015).

Quality assessment procedure. Similar to the inclusion screening process, it is recommended that two or more researchers perform a parallel independent quality assessment (Breretona et al. 2007; Noordzij et al. 2009). All disagreements should be resolved through discussion or consultation with an independent arbitrator (Breretona et al. 2007; Noordzij et al. 2009). The difference is that reviewers will read through the *full text* to carefully examine each study against the quality criteria. The full-text review also provides an opportunity

for a final check on inclusion/exclusion. Studies that do not satisfy the inclusion criteria specified in step 4 should also be excluded from the final literature list. Like in step 4, the list of excluded papers should be maintained for record keeping, reproducibility, and crosschecking (Kitchenham and Charters 2007).

Step 6: Extracting Data

There are several established methods for synthesizing research, which were discussed in the third section (Kastner et al. 2012; Dixon-Woods et al. 2005; Whittemore et al. 2014; Barnett-Page and Thomas 2009). Many of these synthesis methods have been compiled from the medical field, where quality synthesis is paramount to control the influx of new research, as well as qualitative methods papers looking for appropriate ways to find generalizations and overarching themes (Kastner et al. 2012; Dixon-Woods et al. 2005; Whittemore et al. 2014; Barnett-Page and Thomas 2009). The different literature review typologies discussed earlier and the type of literature being synthesized will guide the reviewer to appropriate synthesis methods. The synthesis methods, in turn, will guide the data extraction process—for example, if one is doing a meta-analysis, data extraction will be centered on what's needed for a meta-regression whereas a metasummary will require the extraction of findings. The authors refer the readers to the examples in Table 1 for more detailed advice on the conduct of these extraction and synthesis methods.

In general, the process of data extraction will often involve coding, especially for extending reviews. It is important to establish whether coding will be inductive or deductive (i.e., whether or not the coding will be based on the data or preexisting concepts) (Suri and Clarke 2009). The way in which studies are coded will have a direct impact on the conclusions of the review. For example, in extending reviews such as meta-ethnography and thematic synthesis, conclusions and generalizations are made based on the themes and concepts that are coded. If this is done incorrectly or inconsistently, the review is less reliable and valid. In the words of Stock, Benito, and Lasa (1996, 108), "an item that is not coded cannot be analyzed." Stock, Benito, and Lasa (1996) encourage the use of a codebook and tout the benefits of having well-designed forms. Well-designed forms both increase efficiency and lower the number of judgments an individual reviewer must make, thereby reducing error (Stock, Benito, and Lasa 1996).

If researchers are working in a team, they need to code a few papers together before splitting the task to make sure everyone is on the same page and coding the papers similarly (Kitchenham and Charters 2007; Stock, Benito, and Lasa 1996). However, it is preferred that at least two researchers code the studies independently (Noordzij et al. 2009; Gomersall et al. 2015). Additionally, it is important for researchers to review the entire paper, and not simply rely on the results or the main interpretation. This is the only way to provide context for the findings and prevent any distortion of the original paper (Onwuegbuzie, Leech, and Collins 2012).

Step 7: Analyzing and Synthesizing Data

Once the data extraction process is complete, the reviewer will organize the data according to the review they have chosen. Often, this will be some combination of charts, tables, and a textual description, though each review type will have slightly different reporting standards. For example, a meta-analysis will have a results table for the regression analysis, a metasummary will report effect and intensity sizes, and a framework synthesis will include a conceptual model (Dixon-Woods 2011; Glass 1976; Sandelowski, Barroso, and Voils 2007). Again, the authors refer the reader to Table 1 for suggestions on how to conduct specific reviews.

In addition to the specific review types, a few papers have offered helpful insights into combining mixed methods research and combining qualitative and qualitative studies (Heyvaert, Maes, and Onghena 2011; Sandelowski, Voils, and Barroso 2006). Rousseau, Manning, and Denyer (2008) discuss the hazards of synthesizing different types of literature due to varying epistemological approaches, political and cultural contexts, and political and scientific infrastructure (the authors give the example of a field valuing novelty over accumulation of evidence). More specifically, Sandelowski, Voils, and Barroso (2006, 3–4) discuss the problems faced when combining qualitative literature (such as differences in ontological positions, epistemological positions, paradigms of inquiry, foundational theories and philosophies, and methodologies) and quantitative literature (such as study heterogeneity). Mixed study synthesis, therefore, opens up the entire range of error, and some scholars argue it should not be done (Mays, Pope, and Popay 2005; Sandelowski, Voils, and Barroso 2006). In general, however, there are three types of mixed method review designs: segregated design, integrated design, and contingent design (Sandelowski, Voils, and Barroso 2006).

A segregated design involves synthesizing qualitative and quantitative studies separately according to their respective synthesis traditions and textually combining both results (Sandelowski, Voils, and Barroso 2006). This is exemplified by the mixed methods review/synthesis by A. Harden and Thomas (2005). Qualitative studies were analyzed by finding descriptive themes and distilling them into analytic themes whereas quantitative studies were combined using meta-analyses. The analytic themes were the framework for combining the findings of the quantitative studies into a final synthesis. An integrated design, by contrast, analyzes and synthesizes quantitative and qualitative research together (Sandelowski, Voils, and Barroso 2006; Whittemore and Knafl 2005). This can be done by transforming one type into the other—qualitizing quantitative data or quantitizing qualitative data—or by combining them through Bayesian synthesis or critical interpretive synthesis, as described in the third section (Sandelowski, Voils, and Barroso 2006; Dixon-Woods et al.

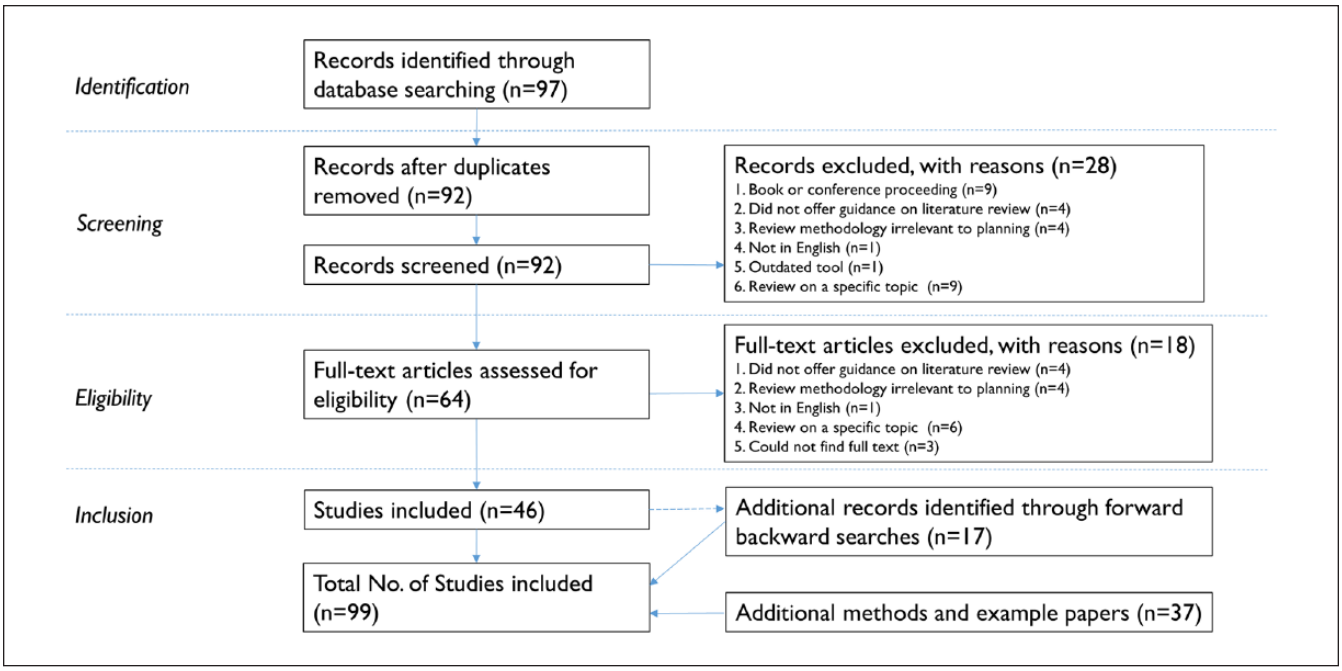

Figure 2. Literature search and evaluation for inclusion.

2006). Lastly, contingent design is characterized by being a cycle of research synthesis—a group of qualitative or quantitative studies is used to answer one specific research question (or subresearch question) and then those results will inform the creation of another research question to be analyzed by a separate group of studies, and so on (Sandelowski, Voils, and Barroso 2006). Although groups of studies may end up being exclusively quantitative and qualitative, "the defining feature of contingent designs is the cycle of research synthesis studies conducted to answer questions raised by previous syntheses, not the grouping of studies or methods as qualitative and quantitative" (Sandelowski, Voils, and Barroso 2006, 36). Realist reviews or ecological triangulation could be classified as a contingent review: groups of literature may be analyzed separately to answer the questions of "what works," "for which groups of people," and "why?"

Step 8: Report Findings

For literature reviews to be reliable and independently repeatable, the process of systematic literature review must be reported in sufficient detail (Okoli and Schabram 2010). This will allow other researchers to follow the same steps described and arrive at the same results. Particularly, the inclusion and exclusion criteria should be specified in detail (Templier and Paré 2015) and the rationale or justification of each of the criteria should be explained in the report (Peters et al. 2015). Moreover, researchers should report the findings from literature search, screening, and quality assessment (Noordzij et al. 2009), for instance, in a flow diagram as shown in Figure 2.

The literature review should follow a clear structure that ties the studies together into key themes, characteristics or subgroups (Rowley and Slack 2004). In general, no matter how rigorous or flexible your methods for review are, make sure the process is transparent and conclusions are supported by the data. Whittemore and Knafl (2005) suggest, even, that any conclusion from an integrative review be displayed graphically (be it a table or diagram) to assure the reader that interpretations are well-grounded. Each review will have varying degrees of subjectivity and going "beyond" the data. For example, descriptive reviews should be careful to present the data as it is reported, whereas extending reviews will, by nature of the review, move beyond the data. Make sure you are aware of where your review lies on this spectrum and report findings accordingly. In general, all novel findings and unexpected results should be highlighted (Okoli and Schabram 2010). The literature review should also point out opportunities and directions for future research (Okoli and Schabram 2010; Rowley and Slack 2004). And lastly, the draft of the review should be reviewed by the entire review team for checks and balances (Andrews and Harlen 2006).

Discussion and Conclusions

Literature reviews establish the foundation for academic inquires. Stand-alone reviews can summarize prior work, test hypotheses, extend theories, and critically evaluate a body of work. Because these reviews are meant to exist as their own contribution of scholarly knowledge, they should be held to a similar level of quality and rigor in study design as we would hold other literature (Okoli and Schabram 2010). Additionally, stand-alone literature reviews can serve as valuable overviews

of a topic for planning practitioners looking for evidence to guide their decisions, and therefore their quality can have very real-world implications (Templier and Paré 2015).

The planning field needs to increase its rigor in literature reviews. Other disciplines, such as medical sciences, information systems, computer sciences, and education, have engaged in discussions on how to conduct quality literature reviews and have established guidelines for it. This paper fills the gap by systematically reviewing the methodology of literature reviews. We categorized a typology of literature reviews and discussed the steps necessary in conducting a systematic literature review. Many of these review types/methodologies were developed in other fields but can be adopted by planning scholars.

Conducting literature reviews systematically can enhance the quality, replicability, reliability, and validity of these reviews. We highlight a few lessons learned here: first, start with a research question. The entire literature review process, including literature search, data extraction and analysis, and reporting, should be tailored to answer the research question (Kitchenham and Charters 2007).

Second, choose a review type suitable for the review purpose. For novice reviewers, Table 1 can be used as a guide to find the appropriate review type/methodology. Researchers should first decide what they want to achieve from the review: is the purpose to describe or scope a body of work, to test a specific hypothesis, to extend from existing studies to build theory, or to critically evaluate a body of work? After deciding the purpose of the review, researchers can follow the typologies in Table 1 to select the appropriate review methodology(ies).

Third, plan before you leap. Developing a review protocol is a crucial first step for rigorous systematic reviews (Okoli and Schabram 2010; Breretona et al. 2007). The review protocol reduces the possibility of researcher bias in data selection and analysis (Kitchenham and Charters 2007). It also allows others to repeat the study for cross-check and verification, and thus increases the reliability of the review. In planning education, we suggest dissertation and thesis committees establish a routine of reviewing students' literature review protocols as part of their dissertation and thesis proposals. This will allow the committee to ensure the literature search is comprehensive, the inclusion criteria are rational, and the data extraction and synthesis methods are appropriate.

Fourth, be comprehensive in the literature search and be aware of the quality of literature. For most types of reviews, the literature search should be comprehensive and identify up-to-date literature, which means that the researcher should search in multiple databases, conduct backward and forward searches, and consult experts in the field if necessary. Assessing rigor and quality, and understanding how arguments were developed, is also a vital concern in literature review. Before proceeding to comparing and integrating findings, we need to first understand each study (Ludvigsen et al. 2016).

Fifth, be cautious, flexible, and open-minded to new situations and ideas that may emerge from the review. Literature review can be an iterative process. Researchers need to pilot the review and decide what is manageable. Sometimes, the research question needs to be narrowed down. Deeper understanding can be gained during the review process, requiring a change in keywords and/or analytical methods. In a sense, the literature review protocol is a living document. Changes can be made to it in the review process to reflect new situations and new ideas.

Sixth, document decisions made in the review process. Systematic reviews should be reliable and repeatable, which requires the review process to be documented and made transparent. We suggest outlets for planning literature reviews, such as the *Journal of Planning Literature*, to make it a requirement for authors to report the review process and major decisions in sufficient detail.

Seventh, teamwork is encouraged in the review process. Many researchers suggest at least two reviewers work independently to screen literatures for inclusion, conduct quality assessment, and code studies for analysis (Gomersall et al. 2015; Kitchenham and Charters 2007; Breretona et al. 2007; Templier and Paré 2015). Some recommended that a well-versed senior researcher (such as an academic advisor) be paired up with a nonexpert junior researcher (student) in the review (Wanden-Berghe and Sanz-Valero 2012). Any discrepancies in opinion should be reconciled by discussion. In an academic setting, it is very effective for an advisor to screen and code at least a few papers together with students to establish standards and expectations for the review before letting the students conduct their independent work.

And finally, we offer a remark on the technology and software available for facilitating systematic reviews. Researchers can rely on software for assistance in the literature review process. Software such as EndNote, RefWorks, and Zotero can be used to manage bibliographies, citations, and references. When writing a manuscript, reference management software allows the researcher to insert in-text citations and automatically generate a bibliography. Many databases (e.g., Web of Science, ProQuest, and Scopus) allow search results to be downloaded directly into reference management software. Software such as NVivo and ATLAS.ti can be used to code qualitative and quantitative studies through nodes or by topical area. The researcher can then pull out the relevant quotes in a topical area or node from all papers for analysis and synthesis. Joanna Briggs Institute (JBI) developed a SUMARI (System for the Unified Management, Assessment and Review of Information) that allows multiple reviewers to conduct, manage, and document a systematic review. Planning researchers can take advantage of the technology that is available to them.

Acknowledgment

We thank the National Institute of Standards and Technology (NIST) for providing funding support for this research via the Center for Risk-Based Community Resilience Planning.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the Center for Risk-Based Community Resilience Planning, a Center of Excellence funded through a cooperative agreement between the U.S. National Institute of Science and Technology and Colorado State University [NIST Financial Assistance Award Number: 70NANB15H044].

References

Andrews, Richard, and Wynne Harlen. 2006. "Issues in Synthesizing Research in Education." *Educational Research* 48 (3): 287–99.

Anthony, David R., Daniel F. Gucciardi, and Sandy Gordon. 2016. "A Meta-study of Qualitative Research on Mental Toughness Development." *International Review of Sport and Exercise Psychology* 9 (1): 160–90.

Arksey, Hilary, and Lisa O'Malley. 2005. "Scoping Studies: Towards a Methodological Framework." *International Journal of Social Research Methodology* 8 (1):19–32.

Arnold, Rachel, and David Fletcher. 2012. "A Research Synthesis and Taxonomic Classification of the Organizational Stressors Encountered by Sport Performers." *Journal of Sport and Exercise Psychology* 34 (3): 397–429.

Banning, James H. 2005. "Ecological Triangulation: An Approach for Qualitative Meta-Synthesis." *What works for youth with disabilities' project, US Department of Education, School of Education, Colorado: Colorado State University*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.152.5185&rep=rep1&type=pdf. Accessed July 21, 2016.

Barnett-Page, Elaine, and James Thomas. 2009. "Methods for the Synthesis of Qualitative Research: A Critical Review." *BMC Medical Research Methodology* 9 (1): 1–11.

Bayliss, H. R., and F. R. Beyer. 2015. "Information Retrieval for Ecological Syntheses." *Research Synthesis Methods* 6 (2): 136–48.

Breretona, Pearl, Barbara A. Kitchenhama, David Budgenb, Mark Turnera, and Mohamed Khalilc. 2007. "Lessons from Applying the Systematic Literature Review Process within the Software Engineering Domain." *Journal of Systems and Software* 80 (4): 571–83.

Britten, Nicky, Rona Campbell, Catherine Pope, Jenny Donovan, Myfanwy Morgan, and Roisin Pill. 2002. "Using Meta Ethnography to Synthesise Qualitative Research: A Worked Example." *Journal of Health Services Research & Policy* 7 (4): 209–15.

Carroll, Christopher, Andrew Booth, Joanna Leaviss, and Jo Rick. 2013. "'Best Fit' Framework Synthesis: Refining the Method." *BMC Medical Research Methodology* 13 (1): 1–16.

Cronin, P., F. Ryan, and M. Coughlan. 2008. "Undertaking a Literature Review: A Step-by-Step Approach." *British Journal of Nursing* 17 (1): 38–43.

Dieckmann, N. F., B. F. Malle, and T. E. Bodner. 2009. "An Empirical Assessment of Meta-Analytic Practice." *Review of General Psychology* 13 (2): 101–15.

Dixon-Woods, Mary. 2011. "Using Framework-Based Synthesis for Conducting Reviews of Qualitative Studies." *BMC Medicine* 9 (1): 1.

Dixon-Woods, Mary, Shona Agarwal, David Jones, Bridget Young, and Alex Sutton. 2005. "Synthesising Qualitative and Quantitative Evidence: A Review of Possible Methods." *Journal of Health Services Research & Policy* 10 (1): 45–53.

Dixon-Woods, Mary, Debbie Cavers, Shona Agarwal, Ellen Annandale, Antony Arthur, Janet Harvey, Ron Hsu, Savita Katbamna, Richard Olsen, and Lucy Smith. 2006. "Conducting a Critical Interpretive Synthesis of the Literature on Access to Healthcare by Vulnerable Groups." *BMC Medical Research Methodology* 6 (1): 1.

Ewing, Reid, and Robert Cervero. 2010. "Travel and the Built Environment: A Meta-analysis." *Journal of the American Planning Association* 76 (3): 265–94.

Fink, A. 2005. *Conducting Research Literature Reviews: From the Internet to Paper*, 2nd ed. Thousand Oaks, CA: Sage.

Fisher, Carla L., Erin Maloney, Emily Glogowski, Karen Hurley, Shawna Edgerson, Wendy G. Lichtenthal, David Kissane, and Carma Bylund. 2014. "Talking about Familial Breast Cancer Risk Topics and Strategies to Enhance Mother–Daughter Interactions." *Qualitative Health Research* 24 (4): 517–35.

Flemming, K. 2010. "Synthesis of Quantitative and Qualitative Research: An Example Using Critical Interpretive Synthesis." *Journal of Advanced Nursing* 66 (1): 201–17.

Gates, S. 2002. "Review of Methodology of Quantitative Reviews Using Meta-analysis in Ecology." *Journal of Animal Ecology* 71 (4): 547–57.

Glass, Gene V. 1976. "Primary, Secondary, and Meta-analysis of Research." *Educational Researcher* 5 (10): 3–8.

Gomersall, J. S., Y. T. Jadotte, Y. F. Xue, S. Lockwood, D. Riddle, and A. Preda. 2015. "Conducting Systematic Reviews of Economic Evaluations." *International Journal of Evidence-Based Healthcare* 13 (3): 170–78.

Gordon, Peter, and Harry W. Richardson. 1997. "Are Compact Cities a Desirable Planning Goal?" *Journal of the American Planning Association* 63 (1): 95–106.

Grant, M. J., and A. Booth. 2009. "A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies." *Health Information and Libraries Journal* 26 (2): 91–108.

Green, B. N., C. D. Johnson, and A. Adams. 2001. "Writing Narrative Literature Reviews for Peer-Reviewed Journals: Secrets of the Trade." *Journal of Sports Chiropractic & Rehabilitation* 15 (1): 5–19.

Greenhalgh, Trisha, Glenn Robert, Fraser Macfarlane, Paul Bate, and Olivia Kyriakidou. 2004. "Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations." *The Milbank Quarterly* 82 (4): 581–629.

Greenhalgh, Trisha, Glenn Robert, Fraser Macfarlane, Paul Bate, Olympia Kyriakidou, and Richard Peacock. 2005. "Storylines of Research in Diffusion of Innovation: A Meta-narrative Approach to Systematic Review." *Social Science & Medicine* 61 (2): 417–30.

Haddaway, N. R., P. Woodcock, B. Macura, and A. Collins. 2015. "Making Literature Reviews More Reliable through Application of Lessons from Systematic Reviews." *Conservation Biology* 29 (6): 1596–605.

Harden, Angela, and James Thomas. 2005. "Methodological Issues in Combining Diverse Study Types in Systematic Reviews." *International Journal of Social Research Methodology* 8 (3): 257–71.

Harden, Samantha M., Desmond McEwan, Benjamin D Sylvester, Megan Kaulius, Geralyn Ruissen, Shauna M Burke, Paul A Estabrooks, and Mark R Beauchamp. 2015. "Understanding for Whom, under What Conditions, and How Group-Based Physical Activity Interventions Are Successful: A Realist Review." *BMC Public Health* 15 (1): 1.

Heyvaert, Mieke, Bea Maes, and Patrick Onghena. 2011. "Applying Mixed Methods Research at the Synthesis Level: An Overview." *Research in the Schools* 18 (1): 12–24.

Kastner, M., A. C. Tricco, C. Soobiah, E. Lillie, L. Perrier, T. Horsley, V. Welch, E. Cogo, J. Antony, and S. E. Straus. 2012. "What Is the Most Appropriate Knowledge Synthesis Method to Conduct a Review? Protocol for a Scoping Review." *BMC Medical Research Methodology* 12:114.

Kitchenham, Barbara, and Stuart Charters. 2007. "Guidelines for Performing Systematic Literature Reviews in Software Engineering." In *EBSE Technical Report*, Software Engineering Group, School of Computer Science and Mathematics, Keele University, Department of Computer Science, University of Durham.

Korhonen, A., T. Hakulinen-Viitanen, V. Jylha, and A. Holopainen. 2013. "Meta-synthesis and Evidence-Based Health Care—A Method for Systematic Review." *Scandinavian Journal of Caring Sciences* 27 (4): 1027–34.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions, International Encyclopedia of Unified Science*, vol. 2, no. 2. Chicago: University of Chicago Press.

Levy, Y., and T. J. Ellis. 2006. "A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research." *Informing Science Journal* 9:182–212.

Limkakeng, Jr., Alexander T., Lucas Lentini Herling de Oliveira, Tais Moreira, Amruta Phadtare, Clarissa Garcia Rodrigues, Michael B. Hocker, Ross McKinney, Corrine I. Voils, and Ricardo Pietrobon. 2014. "Systematic Review and Metasummary of Attitudes toward Research in Emergency Medical Conditions." *Journal of Medical Ethics* 40 (6): 401–8.

Lindell, Michael K., and Ronald W. Perry. 2000. "Household Adjustment to Earthquake Hazard: A Review of Research." *Environment and Behavior* 32 (4): 461–501.

Lucas, Patricia J., Janis Baird, Lisa Arai, Catherine Law, and Helen M. Roberts. 2007. "Worked Examples of Alternative Methods for the Synthesis of Qualitative and Quantitative Research in Systematic Reviews." *BMC Medical Research Methodology* 7 (1): 4.

Ludvigsen, M. S., E. O. C. Hall, G. Meyer, L. Fegran, H. Aagaard, and L. Uhrenfeldt. 2016. "Using Sandelowski and Barroso's Meta-Synthesis Method in Advancing Qualitative Evidence." *Qualitative Health Research* 26 (3): 320–29.

Malekpour, Shirin, Rebekah R. Brown, and Fjalar J. de Haan. 2015. "Strategic Planning of Urban Infrastructure for Environmental Sustainability: Understanding the Past to Intervene for the Future." *Cities* 46:67–75.

Mays, N., C. Pope, and J. Popay. 2005. "Systematically Reviewing Qualitative and Quantitative Evidence to Inform Management and Policy-Making in the Health Field." *Journal of Health Services Research & Policy* 10 (Suppl 1): 6–20.

Myers, P. M. D. 2008. *Qualitative Research in Business & Management*. London: Sage.

Neely, Eva, Mat Walton, and Christine Stephens. 2014. "Young People's Food Practices and Social Relationships. A Thematic Synthesis." *Appetite* 82:50–60.

Noblit, George W., and R. Dwight Hare. 1988. *Meta-ethnography: Synthesizing Qualitative Studies*, Vol. 11. Newbury Park, CA: Sage.

Noordzij, M., L. Hooft, F. W. Dekker, C. Zoccali, and K. J. Jager. 2009. "Systematic Reviews and Meta-analyses: When They Are Useful and When to Be Careful." *Kidney International* 76 (11): 1130–36.

Noordzij, M., C. Zoccali, F. W. Dekker, and K. J. Jager. 2011. "Adding up the Evidence: Systematic Reviews and Meta-Analyses." *Nephron Clinical Practice* 119 (4): C310–16.

Norris, M., C. Oppenheim, and F. Rowland. 2008. "Finding Open Access Articles Using Google, Google Scholar, OAIster and OpenDOAR." *Online Information Review* 32 (6): 709–15.

Okoli, C., and K. Schabram. 2010. "A Guide to Conducting a Systematic Literature Review of Information Systems Research." *Sprouts: Working Papers on Information Systems* 10 (26).

Onwuegbuzie, Anthony J., Nancy L. Leech, and Kathleen M. T. Collins. 2012. "Qualitative Analysis Techniques for the Review of the Literature." *Qualitative Report* 17.

Paré, Guy, Marie-Claude Trudel, Mirou Jaana, and Spyros Kitsiou. 2015. "Synthesizing Information Systems Knowledge: A Typology of Literature Reviews." *Information & Management* 52:183–99.

Paterson, Barbara L., and Connie Canam. 2001. *Meta-study of Qualitative Health Research: A Practical Guide to Meta-analysis and Meta-synthesis*, Vol. 3. Thousand Oaks, CA: Sage.

Pawson, R., T. Greenhalgh, G. Harvey, and K. Walshe. 2005. "Realist Review: A New Method of Systematic Review Designed for Complex Policy Interventions." *Journal of Health Services Research and Policy* 70 (1): 21–34.

Peters, M. D. J., C. M. Godfrey, H. Khalil, P. McInerney, D. Parker, and C. B. Soares. 2015. "Guidance for Conducting Systematic Scoping Reviews." *International Journal of Evidence-Based Healthcare* 13 (3): 141–46.

Petticrew, M., and H. Roberts. 2006. *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell.

Popay, Jennie, Helen Roberts, Amanda Sowden, Mark Petticrew, Lisa Arai, Mark Rodgers, Nicky Britten, Katrina Roen, and Steven Duffy. 2006. "Guidance on the Conduct of Narrative Synthesis in Systematic Reviews." *A product from the ESRC methods programme Version* 1:b92.

Rigolon, Alessandro. 2016. "A Complex Landscape of Inequity in Access to Urban Parks: A Literature Review." *Landscape and Urban Planning* 153:160–69.

Roberts, Karen A., Mary Dixon-Woods, Ray Fitzpatrick, Keith R. Abrams, and David R. Jones. 2002. "Factors Affecting Uptake of Childhood Immunisation: A Bayesian Synthesis of Qualitative and Quantitative Evidence." *The Lancet* 360 (9345): 1596–99.

Rousseau, D. M., J. Manning, and D. Denyer. 2008. "Evidence in Management and Organizational Science: Assembling the Field's Full Weight of Scientific Knowledge through Syntheses." In *AIM Research Working Paper Series*: Advanced Institute of Management Research.

Rowley, Jennifer, and Frances Slack. 2004. "Conducting a Literature Review." *Management Research News* 27 (6): 31–39.

Sandelowski, Margarete, Julie Barroso, and Corrine I. Voils. 2007. "Using Qualitative Metasummary to Synthesize Qualitative and Quantitative Descriptive Findings." *Research in Nursing & Health* 30 (1): 99–111.

Sandelowski, Margarete, and Jennifer Leeman. 2012. "Writing Usable Qualitative Health Research Findings." *Qualitative Health Research* 22 (10): 1404–13.

Sandelowski, Margarete, Corrine I. Voils, and Julie Barroso. 2006. "Defining and Designing Mixed Research Synthesis Studies." *Research in the Schools* 13 (1): 29–40.

Schutz, Alfred. 1962. "Collected Papers. Vol. 1. The Hague: Martinus Nijhoff. 1964." *Collected Papers* 2.

Spiegelhalter, David, Jonathan P. Myles, David R. Jones, and Keith R. Abrams. 1999. "An Introduction to Bayesian Methods in Health Technology Assessment." *British Medical Journal* 319 (7208): 508.

Stanley, T. D. 2001. "Wheat from Chaff: Meta-Analysis as Quantitative Literature Review." *Journal of Economic Perspectives* 15 (3): 131–50.

Stock, William A., Juana Gomez Benito, and Nekane Balluerka Lasa. 1996. "Research Synthesis: Coding and Conjectures." *Evaluation and the Health Professions* 19 (1): 104–17.

Suri, Harsh, and David Clarke. 2009. "Advancements in Research Synthesis Methods: From a Methodologically Inclusive Perspective." *Review of Educational Research* 79 (1): 395–430.

Sutton, Alex J., and Keith R. Abrams. 2001. "Bayesian Methods in Meta-analysis and Evidence Synthesis." *Statistical Methods in Medical Research* 10 (4): 277–303.

Templier, Mathieu, and Guy Paré. 2015. "A Framework for Guiding and Evaluating Literature Reviews." *Communications of the Association for Information Systems* 37, Article 6.

Thomas, James, and Angela Harden. 2008. "Methods for the Thematic Synthesis of Qualitative Research in Systematic Reviews." *BMC Medical Research Methodology* 8 (1): 1.

Wanden-Berghe, C., and J. Sanz-Valero. 2012. "Systematic Reviews in Nutrition: Standardized Methodology." *British Journal of Nutrition* 107:S3–S7.

Webster, Jane, and Richard T. Watson. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review." *MIS Quarterly* 26 (2): xiii–xiii.

Weed, Mike. 2005. "'Meta Interpretation': A Method for the Interpretive Synthesis of Qualitative Research." *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*.

Weingroff, Richard F. 1996. "Federal-Aid Highway Act of 1956, Creating the Interstate System." *Public Roads* 60 (1).

Whittemore, R., and K. Knafl. 2005. "The Integrative Review: Updated Methodology." *Journal of Advanced Nursing* 52 (5): 546–53.

Whittemore, Robin, Ariana Chao, Myoungock Jang, Karl E. Minges, and Chorong Park. 2014. "Methods for Knowledge Synthesis: An Overview." *Heart & Lung* 43 (5): 453–61.

Zhao, Shanyang. 1991. "Metatheory, Metamethod, Meta-data-Analysis: What, Why, and How?" *Sociological Perspectives* 34 (3): 377–90.

Author Biographies

Yu Xiao is an associate professor in the Department of Landscape Architecture and Urban Planning at Texas A&M University. Her research interests include community resilience, disaster management, and economic development.

Maria Watson is a PhD candidate in the Urban and Regional Science program at Texas A&M University. Her research interests include disaster recovery, public policy, and economic development.