Research Synthesis: Challenges & Opportunities of an Evolving Methodology

Conducting a Useful and Usable Review

Kris Morrissey, Ph.D. June 17, 2021

The assumptions, expectations, and potential of the literature review as a methodology has evolved significantly in recent decades. With advancements in sophisticated and accessible analytical software, combined with the use of systematic protocols, reviews are increasingly generating results that can advance knowledge and practice. But, while *reviews*, *synthesis* or *meta-analysis* have the capacity to inform practice in unique ways, they are also fraught with their own methodological, ethical, and practical issues.

The National Science Foundation now recognizes "Literature Reviews, Syntheses, or Meta-Analyses" as a distinct category of funding for research in the AISL field. *Addressing Societal Issues through STEM* (ASCs) was part of the first round of studies funded in that category. Committed to advancing practices that engage the public in societal issues, the project aimed to produce results that would be useful and usable by practitioners, researchers, funders, and others engaged in the work of informal learning. Drawing on the experience of the ASCs project and other experiences with reviews, this paper suggests that for reviews to be optimal, they need to be conducted in ways that are:

1. Systematic: Protocols and strategies for searching, selecting, analyzing literature should be systematic and transparent.
2. Pragmatic: A practical, organized and flexible management system is necessary to access, document, and keep track of studies that are included, as well as those excluded.
3. Fair: Interpretations and claims based on the work of other researchers should be fair.

The paper discusses these qualities and provides examples from practice and then closes with a summary of the value of literature synthesis to advance practice and research, and suggestions for further research and experimentation.

Introduction

Conducting or reading a synthesis can be a shortcut to answering questions of what works and what doesn't work, or what's been tried and what hasn't. Fields as diverse as medicine, education, and public policy often rely on research reviews and synthesis to inform strategic decisions about policies, funding priorities, and practices (Cooper et al., 2019). For the field of informal learning, research syntheses hold particular promise. The field values evidence-based practice but given the disparate paths and disciplinary backgrounds that bring individuals to the field, there is often a lack of a shared knowledge base. For those involved in the work of informal learning, reviews can provide a "more comprehensive and stronger picture based on many studies and settings than a single study" (p. 3). This is particularly true as the protocols and strategies for conducting literature reviews continue to evolve, supported by increasingly sophisticated and accessible analytical software (Pickard, 2013; Booth, Sutton, and Papaioannou, 2016; Cooper et al., 2019).

In 2017, the National Science Foundation introduced a new category of AISL funding for "Literature Reviews, Syntheses, or Meta-Analyses." *Addressing Societal Issues through STEM* (ASCs) was one of the projects funded in that new category. I was the Primary Investigator; John Fraser (Knology) was the Co-PI. As colleagues involved in the launch of the journal Museums & Social Issues, and as professionals and academics involved in publications over the past two decades, we saw a tremendous opportunity to review the growing body of literature about the role of the informal learning field in societal issues. In 2018, we enthusiastically proposed a review of that work. The proposal was rated *Fair* (i.e., not so good), and was politely "declined." Although our problem statement was strong, we had not clearly described the protocols we would use for accessing, selecting, and analyzing studies. That lack of clarity is common in reviews and is the motivation for writing this paper. The following year, we proposed the same project but defined our terms, connected to theory, described the protocols, and we provided evidence that the literature existed and could be accessed and analyzed with the resources and tools we had available. Reviewers gave the proposal *Excellent* ratings, commenting on the "rigorous methodology" and the "theoretical underpinnings."

The "rigorous methodology" was often challenged, compromised, or revised during the two years of the study, and questions pushed us to continually revisit, revise, or dig deeper. We learned a lot about the engagement of museums in social issues (Morrissey & Ball, in review; Fraser, Norlander & Nock, in review; Morrissey, Fraser & Ball, in review), and we learned a lot about the challenges of conducting a review. Synthesizing disparate studies into a cohesive narrative that is usable and useful is difficult. Condensing nuanced and subtle aspects of the original research studies for the sake of brevity can inadvertently ignore important aspects of the original research; but including too much detail can overwhelm the reader. Conducting a synthesis also faces practical challenges dealing with a potentially large set of studies that may vary in length, type of data, audience, and other variables. This paper discusses the unique challenges of conducting a review that translates research findings into applicable formats (usable) and leads to "transformational and observable impact" on practices, policies, and expectations (useful) (Pan & Pee, 2020, p. 407). Drawing on the ASCs

project and other experiences, this paper proposes three qualities or characteristics of a usable and useful synthesis:

1. **Systematic:** Protocols for selecting, evaluating, analyzing and synthesizing literature should be systematic and transparent.
2. **Pragmatic**: A practical, organized, and flexible management system is necessary to access, document, and keep track of studies that are included as well as those excluded.
3. **Fair:** Interpretations and claims based on the work of other researchers should be fair.

What is a Research Synthesis?

Literature review is a broad umbrella term for any survey of existing literature that "summarizes and evaluates the existing knowledge on a particular topic" (Machi & McEvoy, 2012, p. 2). Most peoples' experience with a literature review is as a precursor to a research study or perhaps the design of an exhibit or proposal. However, there are many types of literature reviews that go beyond summarizing existing literature, with the meta-analysis perhaps the most familiar. A *meta-analysis* is a specific type of a literature review that is most commonly understood to be a quantitative analysis of quantitative data (Cooper et al., 2019; Fraenkel, Wallen, Hyun, 2015; Gough, Oliver & James, 2018). But often the body of literature to be reviewed is not quantitative, or the questions addressed are not best served by a quantitative analysis. A range of strategies and methodologies have evolved to address this need, referred to with different terms such as *research synthesis* (Cooper et al., 2019), *systematic review* (Gough, Oliver, & Thomas, 2012; Jesson et al., 2011), *configurative* or *aggregative review* (Gough, Oliver, & Thomas, 2017), and *comprehensive literature review* (Onwuegbuzie, & Frels, 2016). The lack of consensus around vernacular can confuse the reader and, confounding that challenge, some of the terms are also used as descriptive adjectives. For example, "systematic" can be used to refer to any process that follows a clear and transparent process such as "protocols should be systematic". But the term *systematic review* also refers to a type of review associated with specific assumptions and protocols such as those described in the book, An Introduction to Systematic Reviews (Gough, Oliver, Thomas, 2017). For the sake of clarity, in this paper, terms are italicized when they refer to a specific methodology such as *meta-analysis*, or *systematic review*. In the ASCs proposal, we described our protocol this way:

> *We will use protocols associated with the pedagogy of a configurative literature review, distinguished from a meta-analysis in the nature of the literature reviewed and the goal of the review. A meta-analysis aims to "add up" the findings across homogeneous studies, most often through quantitative assessments. Medical studies often employ this methodology to balance different findings about a drug or treatment.*
>
> *In contrast, ASCs research will follow the protocol of a configurative literature review because of the preponderance of qualitative research in ISL research, the lack of homogeneity in ways research is reported, and our interest in answering broad research questions and theory-building rather than measuring size of effect. The approach for a configurative review grows out of grounded theory work*

(Creswell & Poth, 2018; Devlin, 2018; Merriam & Tisdell, 2016) and is inquiry-based and qualitative in nature.

(Morrissey & Fraser, Unpublished grant proposal, 2018)

During the course of the ASCs project, I more often used the broader term *research synthesis*, in part because I have come to appreciate the challenge and critical importance of the process of synthesis. The term is also recommended by the "Handbook of Research Synthesis and Meta-Analysis " (Cooper et al. 2019), which is described as "the definitive vade mecum for behavioral and social scientists intent on applying the synthesis craft" (Cooper et al., 2019, p 6). At this point when definitions are murky, the transparency and clarity of a particular word choice is probably more important than the specific word.

The "Handbook of Research Synthesis and Meta-Analysis" (Handbook) includes a taxonomy of the dimensions or distinctions that review authors often use to describe their work (Cooper et al., 2019, p. 5). Table 1 identifies these dimensions and typical options for each dimension. Clarity on each of these at the outset guides the design and application of protocols for the literature that is searched, the language used to interpret the data, and the usefulness of the results for the intended audience. The ASCs project used this structure to organize the assumptions and expectations of the research. The focus was *practice* for a primary audience of *practitioners*, as well as scholars and individuals involved in policies and standards across the field. The goal was the *identification of central themes and issues*; protocols were designed to yield a corpus that was *representative* rather than exhaustive; and the organization was *conceptual*. The research perspective was motivated by our *position* that museums can and must engage with social issues in order to fulfill their mission and their responsibilities to the public.

Table 1. Characteristics of a Research Synthesis

Dimension	Option	ASCs Design
Focus	Findings, Methods, Theories, or Practice	Practice
Goal	Criticism, Identification of Central issues, or Integration	Identification of Central Issues
Perspective	Neutral or Position	Position
Coverage	Exhaustive, Selective, Representational, Pivotal	Representational
Organization	Historical, Conceptual, Methodological	Conceptual
Audience	Scholars, Practitioners, Public, Policy Makers	Practitioners

Note: Adapted from Cooper, Hedges, & Valentine (2019)

Quality 1: SYSTEMATIC PROTOCOLS

The term "systematic" is ubiquitous in the literature about conducting a review or synthesis, and inextricably linked to the term "protocol." A systematic protocol is explicit, transparent, and driven by the research questions. Another researcher should be able to use the same protocols and arrive at roughly the same results, which is the bedrock of the scientific method (Cooper et al., 2019; Gough et al., 2012; Cresswell & Poth, 2018). Lack of protocols or inconsistent adherence to protocols can easily lead to inconsistent decisions that may impact the results. The Handbook (Cooper, et al., 2019) describes these risks:

- Variations in the definitions of terms might lead to differences in analytical procedures.
- Variations in searched sources might lead to differences in what is retrieved.
- Variation in the information recorded might lead to differences in hypotheses and conclusions.
- When evaluating the studies, variations in criteria might lead to differences in which studies remain in the synthesis.
- When analyzing the studies, variations in procedures can lead to differences in cumulative results.
- When interpreting the results, variations in criteria might lead to differences in what is deemed important and what is acknowledged.
- Variations in reporting might lead readers to place more or less trust in the synthesis.

All of these variations can be ameliorated, if not avoided, by developing and adhering to systematic protocols. Protocols are not recipes with exact directions for selecting or analyzing studies, but they provide a gauge to help the researcher calibrate decisions and to minimize variations. Two of the protocols used in the ASCs project are described below. The first protocol provided an operational definition of a social issue (or social problem) which guided the search and selection process. The second protocol guided the evaluation of the studies to determine selection and utility of the studies.

Example of a Protocol for Defining Terms

The ASCs research started with a definition of a social issue that was based on a previous effort to synthesize evaluation reports (Morrissey et al., 2014). In that project , we spent countless hours looking at evaluation reports and debating whether the topic was a social issue or not. We eventually realized we couldn't rely on our opinions but needed a definition that was reasonable and operational. We turned to the work of sociologists who describe a social problem not as a condition or phenomenon that is objectively harmful, but rather a condition that society has acknowledged is harmful and needs to be addressed (Best 2013; Best 2016) .

Based on that experience, the ASCs project defined social issues as conditions that are: *(1) publicly acknowledged as harmful or limiting to society; (2) complex, systemic, and often enduring;*

and (3) characterized by a lack of public agreement on the nature of the problem or the nature of the solution. The key phrase "publicly acknowledged as harmful or limiting" shifted the focus from identifying topics that *we* considered to be social issues to identifying topics that were *publicly recognized* or acknowledged to be problems. Topics were identified by examining a selection of public opinion polls, drawing from polls that had been reviewed and rated for reliability and political leanings (NateSilver, 2017). A spreadsheet (Figure 1) was used to list each topic the polls that ranked or measured the topic as important to the public. That list was used to generate keywords to search for relevant literature. We also conducted searches with terms such as "social issues" or "social problems". We considered this as an internal rather than external claim of a social issue, meaning that the topic wasn't identified on public opinion polls, but the author of the article made and supported claims that the topic was a social issue.

Topic	Gallup (B)	Pew (B-)	ABC/Washington Post (A+)	Associated press	NBC/Wall street (A-)	NPR/PBS/Marist	Quinnipiac Univ	CNN (B+)
Federal budget deficit/Federal del	☑	☐	☐	☐	☑	☐	☐	☐
Gap between rich and poor	☑	☐	☐	☐	☐	☐	☐	☐
Race relations/Racism	☑	☑	☐	☐	☐	☐	☐	☐
Environment/Pollution/Climate chang	☑	☐	☐	☐	☐	☐	☐	☐
Poverty/Hunger/Homelessness	☑	☐	☐	☑	☐	☐	☐	☐
Abortion	☑	☐	☐	☐	☑	☐	☐	☐
Border wall	☐	☑	☑	☐	☐	☐	☐	☐
Immigration	☑	☑	☑	☐	☑	☐	☑	☑
Bipartisan cooperation	☐	☑	☐	☐	☐	☐	☐	☐
the economy & jobs	☑	☑	☑	☑	☑	☐	☑	☑
terrorism/ national security	☐	☑	☐	☐	☑	☐	☐	☐
climate change	☐	☑	☑	☑	☑	☐	☐	☐
Health care	☐	☑	☑	☑	☑	☑	☑	☑

Figure 1. Selection of topics listed on public opinion polls.

Example of a Protocol for Evaluating Quality

Most textbooks about reviews devote at least a chapter to articulating the importance and the challenges of evaluating the quality of the studies selected for the review. As one states, "Assessing the quality and relevance of individual studies within a review is a crucial part of the review process and contributes to the quality and credibility of a review itself" (Gough, Oliver, & Thomas, 2012, p. 154). But deciding what criteria to use for appraising the quality of an individual study is difficult because ideas about quality are context dependent. As one text explains:

> *"A teacher, human resource manager, or clinical psychologist might say that a good study is one that results in knowledge that improves practice. A journal editor might say that a good study is one that is often cited; a journal peer reviewer might say that a good study is one that makes a theoretical contribution to the literature."*
>
> *(Valentine, 2019, p. 130)*

The handbook describes three options for evaluating quality. The synthesis can: (1) avoid dealing with quality, (2) rely on the peer review process (i.e., use only published studies) or, (3) rely on scores derived from quality scales. The ASCs researchers weren't comfortable ignoring quality or creating our own quality scales which can be difficult and is generally discouraged (Valentine, 2019, p. 130). We chose to rely on a variation of peer-review, which is the most common strategy for addressing study quality in a research synthesis (p. 132). We selected our literature from three sources that had already undergone some type of formal

review of the value of the work to the field or the validity of the research: peer-reviewed literature; reports posted on informalscience.org; and doctoral and master's theses published through the ProQuest international research database. Graduate research is overseen by faculty who are often recognized experts on the topic of study; publications in peer-reviewed journals have been reviewed by experts on the topic; and nationally-funded projects are typically subjected to external review by experts in the field before being funded. We hypothesized that including these sources would provide a standard of quality, and that the combination of the three might provide more breadth and representation than a single source of studies.

We also assessed the utility of the studies to our research questions by reviewing studies and noting if the research questions, methodologies and results were clearly identified. Those studies were classified as *empirical* and coded and were used to address most of the questions. Other studies were classified as either *descriptive* or *expository* and used selectively as described in Table 2.

Table 2. Classifying Studies by Utility

Category	Description	Used for
Empirical *n*=150	Describes research or evaluation questions; described a method of collecting or analyzing data; and presented data.	Answering all research questions.
Descriptive *n*=71	Describes an exhibit, a program, or other product or practice. Data from an evaluation might be referenced, but not the focus of the publication.	Answering research questions about topics addressed, types of collaborations. Not used for research questions about impacts.
Expository *n*=16	States opinions about the field or recommendations for practices.	Answering research questions about topics, arguments for engagement. Not used for research questions about impacts or any questions related to "what is" rather than "what should be."

Quality 2: PRAGMATIC DATA MANAGEMENT

Regardless of how compelling the research question is, the ability to answer it through a literature review depends upon the existence of appropriate literature, access to that literature, and a management system to keep track of the literature that has been included and excluded. The book *Introduction to Systematic Reviews* (Gough et al., 2012) emphasizes that "The reliability of a review rests in part on the ability of the systematic reviewers to keep track of the studies they have found and the plans they have for them" (p. 137). After conducting several reviews involving hundreds of studies, I would expand that claim to state that the likelihood of even *completing* a review, much less producing results that are useful and usable, is dependent on strategies and opportunities for accessing, organizing, and managing the data. Maintaining a system for keeping the data secure, accessible, and usable is "a fundamental part of good research practice" (McLeon, Childs & Lomas, p. 71).

We were fortunate to have access to the significant online inventory and subscriptions of the University of Washington, rated as one of the top ten research libraries in the country. This advantage provided access to literature, as well as sophisticated search tools and access to ProQuest, which is the largest international repository of graduate research. We documented each search on a spreadsheet identifying the search terms, any conditions we put on the search, and the number of citations the search yielded as shown in Figure 2. The citations of the studies found in each of the searches were listed on a second spreadsheet showing the searches that yielded that particular article (Figure 3). This process linked each document considered for inclusion to the search terms that yielded that citation. This tracking of searches was tedious but it provided a record that could be used to explain gaps or questions about the inventory.

Figure 2. Selection from the Log of Searches within the InformalScience.org Repository

Search Number	Keywords	Refinements	Results
319	Race AND Museum*	Evaluations, Reports, and Peer-Reviewed Articles Only	26
320	Racism OR race OR rac*	Evaluations, Reports, and Peer-Reviewed Articles Only	4
323	Controver*	Evaluations, Reports, and Peer-Reviewed Articles Only	30
325	"Social Challenge"	Evaluations, Reports, and Peer-Reviewed Articles Only	1
326	"Societal Challenge"	Evaluations, Reports, and Peer-Reviewed Articles Only	0
327	"Social Concern"	Evaluations, Reports, and Peer-Reviewed Articles Only	2
328	"Societal Concern"	Evaluations, Reports, and Peer-Reviewed Articles Only	2
329	"Social Issue"	Evaluations, Reports, and Peer-Reviewed Articles Only	10
330	"Societal Issue"	Evaluations, Reports, and Peer-Reviewed Articles Only	4
331	"Social Problem"	Evaluations, Reports, and Peer-Reviewed Articles Only	3
332	"Societal Problem"	Evaluations, Reports, and Peer-Reviewed Articles Only	0
333	Museum* AND Energy	Evaluations, Reports, and Peer-Reviewed Articles Only	19
334	"Science Center" AND Energy	Evaluations, Reports, and Peer-Reviewed Articles Only	4
335	Informal AND Energy	Evaluations, Reports, and Peer-Reviewed Articles Only	20
336	Zoo* AND Energy	Evaluations, Reports, and Peer-Reviewed Articles Only	1
337	Aquarium* AND Energy	Evaluations, Reports, and Peer-Reviewed Articles Only	0
338	Botanical AND Energy	Evaluations, Reports, and Peer-Reviewed Articles Only	1
339	Museum* AND Religion	Evaluations, Reports, and Peer-Reviewed Articles Only	3
340	"Science Center" AND Religion	Evaluations, Reports, and Peer-Reviewed Articles Only	0
341	Informal AND Religion	Evaluations, Reports, and Peer-Reviewed Articles Only	4
342	Zoo* AND Religion	Evaluations, Reports, and Peer-Reviewed Articles Only	0

Figure 3. Selection from the log of citations that resulted from searches.

Search number	Title	Abstract
10	Emotions at Play: Assessing Emotional Responses of Teenagers after they play "Papers, Please"	The purpose of this study was to assess and describe emoti
5 & 12	Empathy and Its Potential in Museum Practices	This research explores the museum community's interest in
4	Engaging Latino audiences in informal science education	Environment for the Americas (EFTA), a non-profit organizati
2	Exploring the Effects of Communication Framed by Environmental Concern in Informal Science Education Contexts	Informal science education (ISE) venues such as zoos, natu
12	Eyes like this	I came to the United States to attend high school at the age
1	Framing diversity: An ethnography of educational outreach at the Getty Museum, Los Angeles	This dissertation investigates institutional efforts to promote
2 & 5 & 12	From Exhibit to Action: The Impact of Museum Experiences on Visitors' Social Justice Actions	Over the last few decades, many museums have embraced
12	From visit to action: How zoo visitor characteristics influence environmentally-responsible behavior	Over the last 30 years, AZA-accredited zoos and aquariums
2	Furthering the Discussion: Museums and Controversial Issues	During studies at the University of the Arts an industry shift in
10	Gatecrashers: The First Generation of Outsider Artists in America	Although interest in the work of untrained artists has surged
5 & 12	Generation Digital	Generation Digital explores how digital technology, particular
1 & 5	Green Museums Waking up the World: Indigenous and Mainstream Approaches to Exploring Sustainability	Mainstream and Indigenous Museums are ideally situated, b
1 & 5	Green Up: Design Proposal for an Exhibition Exploring the Important Role of Communities' Involvement in Sustainable Practices Through Art and Food	This paper presents a proposal for an exhibition on ecology t

Once the inventory was stable, each study was numbered with the peer-reviewed articles starting with 001 and the suffix "AR" to designate it as an article (e.g., 001.AR, 002. AR); documents drawn from informalscience.org started with the number 200, with the suffix "IS" (e.g., 200.IS, 201.IS) and the doctoral dissertations and master's theses located through the ProQuest Database started with 300 followed by "PQ" (e.g., 300.PQ, 301.PQ). All items were listed on a spreadsheet with the associated meta-data such as author, year of publication, keywords, usually the abstract, and other metadata. PDF's of all the studies were uploaded to a shared Google drive as the main repository of documents and also imported into an NVIVO file.

The data management relied on both spreadsheets and NVIVO qualitative data management software. Spreadsheets helped organize metadata and were easier to share for comments across the team and were sometimes used to analyze characteristics of studies. NVIVO provided extensive tools for analyzing, organizing, searching, coding and analyzing the documents as well as visualizations of data that often guided lines of inquiry and further analysis. NVIVO tools also allow the researcher to organize groups of studies into folders (i.e., empirical studies, descriptive studies, expository studies) or sort groups into sets (i.e., all studies that include professional development) and then analyze within or between those groups.

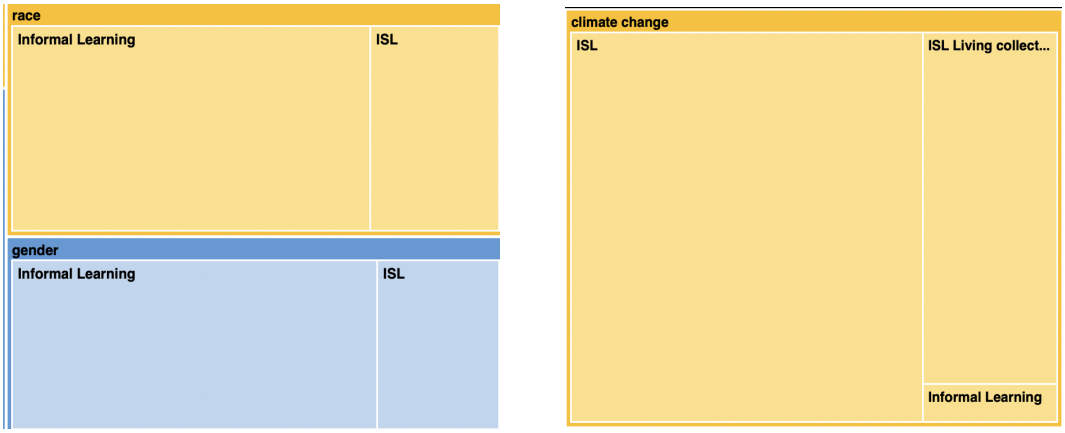
Using both of these management tools was often cumbersome and required diligence to keep both aligned, but each provided unique and important affordances to our work and neither was adequate independent of the other. And spreadsheets and NVIVO can both analyze data created and exported from the other which provides rich opportunities for

analyzing and visualizing data. For example, the inventory spreadsheet was imported into NVIVO and used as a classification sheet to connect the data in the spreadsheet to each PDF within the NVIVO file. Text frequencies identified within NVIVO were exported to spreadsheets to more easily create charts. See Figure 4 for examples of the crosswalking of data between NVIVO and spreadsheets. This paper does not argue for a particular data management tool, but instead suggests that the data management tools and systems should be appropriate for the data set, as well as the skills and preferences of the researchers.

Figure 4. Classification sheet with examples of images generated through NVIVO using the classification sheet as source of data or as a filter for studies that generated data for images.

Study Number	Category	Environment	Funders	Interventions	Topic 1
207.IS	Empirical	ISL Living collections	NOAA	PD	climate change
208.IS	Empirical	ISL Living collections	NOAA	PD	climate change
258.IS	Empirical	ISL	NOAA	media planetarium	climate change
052.ar	Empirical	ISL	NIH	combination	Health & Wellness
254.IS	Empirical	ISL	NIH	Exhibit	Health & Wellness
229.IS	Empirical	ISL	National Science Foundat	research	environment
038.ar	Empirical	ISL	National Science Foundat	PD	climate change

A spreadsheet can be used to classify all studies within NVIVO which then allows for sorting, filtering, and visualizing the corpus by any of the variables within the spreadsheet.



A hierarchy chart displays a visual representation of the data such as these charts that cluster studies by topic and then by environment. The topics of gender and race are addressed less often within informal STEM learning environments than other informal learning environments where as the topic of climate change is addressed most commonly within ISL settings (according to this set of data).

Quality 3: FAIR INTERPRETATION

A review carries more weight than a single study and therefore it is particularly important that results of the review fairly represent the aggregate of studies. When a review states "these studies suggest" or "the aggregate of these studies confirm," the reader needs to be able to trust that the claim represents the aggregate of the studies and not just the opinion of the person conducting the review. This puts a singular responsibility on researchers conducting reviews to be fair, and to be "free from self-interest, deception, injustice, or

favoritism" (Fair, Merriam-Webster, n.d.). Fairness may be particularly important in research about social problems where issues of equity that surfaced in the original research could be misunderstood or simply missed by the person conducting the synthesis, inadvertently perpetuating norms or inequities.

There is also a tension between *representing* the individual studies and *synthesizing* the individual studies. The synthesis should produce new insights that are more than the sum of the parts and yet maintain a fidelity to the individual studies. Each individual study has a unique voice and message, and the researcher conducting the synthesis also brings their own assumptions, perspective and expectations for the synthesis (Onwuegbuzie & Frels, 2016). While it is often tempting, and sometimes appropriate to use blanket claims such as "these studies claim/prove/suggest," these types of statements can fail to distinguish between what the *aggregate of studies* claim, and what the *author of the synthesis* is claiming. This tension can be further complicated by the agonism and rhetorical traditions of publishing and academic cultures which often encourage and reward critique and big findings rather than small findings

A useful book on this topic, "They Say I Say" (Graff & Birkenstein, 2018), provides textual templates for distinguishing between what others say and what the author says. Deceptively simple templates that generally follow the structure of "They say … ; I Say …" help the reader understand the source of the claim and to hear the voice of the source. When I supervised doctoral dissertations or master's theses, I often saw examples of citations used to support claims that were perhaps valid, but were not the claims of the authors cited. I created an acronym (SEE) as a mnemonic device that: Any Statement should be followed by Evidence and Examples. A statement or a claim made in any review or synthesis should provide evidence and examples from the literature.

In the ASCs project, there were many claims that we deliberated over. Were we going too far astray from the literature or, conversely, were we avoiding potentially useful observations because we stayed too bound to the individual studies? We often asked, "Does the literature really support that claim? Is it only one or two studies that support that claim?" Below are a few sentences from our articles where we attempted to use quotations or textual analysis (in the third example) to clarify voices and claims.

- Example of using several short quotations from original studies to support a claim of the synthesis about social change (claim underlined):
    - "Most of these articles expressed the desire to inspire or effect change at the societal level, using phrases such as "working together to effect social change" (Cabrera & Gomberg-Munoz, 2010, p. 205); "oriented towards social change by facilitating new ways of thinking about climate change" (Cameron, 2012, p. 331); and "ISLCs have the potential to be a highly effective vector for effecting broad societal change" (Geiger et al, 2017, p. 222)."
- Example of using a single quotation of an author to support the synthesis' claim about addressing internal inequities:
    - "A number of studies and articles also spoke to the importance of addressing issues of equity internally. In a study about social justice

perspectives across the field, one professional stated, “For a long time, we have tried to have those conversations with the public without having those really important intelligent conversations internally first” (Filo 2017: 20).

- Example of using textual analysis to support the synthesis’ claim about science dominating discussions over technology, engineering, and math.
    - “In fact, the term “science” and variations of the word (e.g., scientific) occurred over 13,000 times compared to engineering or technology, which each had fewer than 500 occurrences or math, which had fewer than 100.”

Summary

The ASCs project provided an opportunity to see the breadth of work taking place in large and small institutions across the country that addresses social issues. The accountability reflected in the evaluation reports, the perceptive and important questions and aspirations addressed through research, and the innovative and evolving practices we saw was inspiring. The aggregate of these studies helped identify what we know and don’t know about engaging the public in social issues. It showed particular opportunities ISL field has to create social change. And it helped reveal what we don’t know about how to measure or document change at the individual or societal level. Those findings confirmed to us the value of reviews to inform policies and practices.

The process of engaging in the review also illustrated what we know and don’t know about how to conduct a review that can inform practice. In design work, there is a saying that “usable doesn’t mean useful” and “useful doesn’t mean usable.” I hope that the reflections and observations shared in this paper advance the understanding and application of a review as a methodology that is both useful and usable for researchers, practitioners and all those who support the work of the informal learning field.

Further Research

Further research on the value of reviews and syntheses to reflect and inform practice should explore ways to represent the myriad places where the field deliberates and reflects on practices that are and that could be. Drawing literature from peer-reviewed sources provides a distinct advantage for ensuring a level of quality control. However, publication bias and other factors influence what appears in peer-reviewed literature and significant discourse and documentation of emerging practices may take place in less formal forums such as presentations at conferences, social media and other forums. Those sources may provide a different perspective on the field.

Further research could also assess the uptake and impact of reviews. There is at least one framework that has been proposed to measure *usefulness*, *usability*, as well as *in-use* (Pan & Pee, 2020). Likely other frameworks exist or wait to be developed. Research that measures the *usefulness*, *usability*, and *in-use* of the results of research synthesis could inform practices and priorities in the field of informal learning.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1906556. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Deep appreciation to co-PI John Fraser and researchers Rebecca Joy Norlander, Kate Flinner, Joanna Laursen Brucker, Theresa Ball at Knology for their collaboration and reflections that led to this paper. Thanks also to the ASCs advisors, Joan LaFrance, Tonya Matthews, Stephanie Ratcliffe, and Troy Dow Sadler for their ongoing guidance and to the external reviewers Marjorie Bequette and Kevin Crowley, particularly for encouraging us to draw attention to the process of conducting a research synthesis.

References

Best, J. (2013). *Social Problems* (2nd ed.). W. W. Norton & Company.

Best, J. (2016). *Social Problems* (3rd ed.). W. W. Norton & Company.

Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic approaches to a successful literature review* (Second edition). Sage.

Cooper, H. M. (2017). *Research Synthesis and Meta-Analysis: A Step-by-Step approach* (5th ed.). SAGE.

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (Eds.), (2019). *Handbook of Research Synthesis and Meta-Analysis* (3rd ed.). Russell Sage Foundation.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches* (4th ed.). Sage.

Devlin, A. S. (2018). *The Research Experience: Planning, Conducting, and Reporting Research*. SAGE.

Fraser, J., Norlander, R. J., & Nock, K. (manuscript submitted for publication). Methods that Reveal and Conceal: What the Literature Says about Museums, STEM, & Social Issues

Fraenkel, J.R., Wallen, N. & Hyun, H., 2018. *How to Design and Evaluate Research in Education* (10th ed.). McGraw-Hill Education.

Gough, D., Oliver, S., & Thomas, J. (2012). *An Introduction to Systematic Reviews*. Sage.

Graff, G. & Birkenstein, C. (2018). *"They Say, I Say": The Moves that Matter in Academic Writing (4th ed.).* Norton & Company.

Jesson, J., Matheson, L. & Lacey, F, 2011. *Doing your Literature Review: Traditional and Systematic Techniques*. Sage.

Machi, L. A., & McEvoy, B. T. (2012). *The Literature Review: Six Steps to Success* (2nd ed). Corwin Press.

Matt, G. & Cook, T. (2019). Threats to the Validity of Generalized Inferences from Research Synthesis. In Cooper, (Eds.) *Handbook of Research Synthesis and Meta-Analysis* (3rd ed.). (pp. 490-522). Russell Sage Foundation.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation* (4th ed.). Jossey-Bass.

Morrissey, K. & Ball, T. (manuscript submitted for publication). The Role of the ISL Field in Addressing Social Issues: A Matter of Capacity and Intention.

Morrissey, K., Petrie, K., Danning, K., Wingleharth, T., & Montano, P. (2014). *Museums & Social Issues: A Research Synthesis of an Emerging Trend* (Building Informal Science Education) [Informal Science.org report].

Morrissey, K., & Dirk, G. (2020). Identity & Museum Practice: Promises, Practices, and a Broken Pipeline. *Curator: The Museum Journal*, *63*(4), 555–570.

Onwuegbuzie, A.J. & Frels, Rebecca, 2016. 7 Steps to a Comprehensive Literature Review: A Multimodal & Cultural Approach. SAGE Publications.

Pan, S. L., & Pee, L. G. (2020). Usable, in-use, and useful research: A 3U framework for demonstrating practice impact. *Information Systems Journal*, 30(2), 403–426. https://doi.org/10.1111/isj.12274

Pickard, A. J., & Childs, S. (2013). *Research methods in information*. Neal-Schuman.

NateSilver538. (2019, March 25). *Pollster Ratings*. FiveThirtyEight. https://projects.fivethirtyeight.com/pollster-ratings/.

Valentine, J. (2019). Incorporating Judgments about Study Quality into Research Synthesis. In Cooper, (Eds.) *Handbook of Research Synthesis and Meta-Analysis* (3rd ed.). (pp. 129-140). Russell Sage Foundation.

~

Knology Practical social science for a better world

Morrissey, K. (2021). *Research Synthesis: Challenges and Opportunities of an Evolving Methodology*. Knology. https://doi.org/*10.13140/RG.2.2.13810.63680/1*

Date of Publication June 17, 2021 Knology Publication # NSF.022.499.02

This material is based upon work supported by the National Science Foundation under Grant No. 1906556. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.