

University of Texas at Tyler

Scholar Works at UT Tyler

Human Resource Development Faculty
Publications and Presentations

Human Resource Development

2023

(Semi)automated approaches to data extraction for systematic reviews and meta-analyses in social sciences: A living review protocol

Amanda Legate

University of Texas at Tyler, ALegate@patriots.uttyler.edu

Kim Nimon

University of Texas at Tyler, knimon@uttyler.edu

Follow this and additional works at: https://scholarworks.uttyler.edu/hrd_fac



Part of the [Business Commons](#)

Recommended Citation

Legate, Amanda and Nimon, Kim, "(Semi)automated approaches to data extraction for systematic reviews and meta-analyses in social sciences: A living review protocol" (2023). *Human Resource Development Faculty Publications and Presentations*. Paper 32.

<http://hdl.handle.net/10950/4236>

This Article is brought to you for free and open access by the Human Resource Development at Scholar Works at UT Tyler. It has been accepted for inclusion in Human Resource Development Faculty Publications and Presentations by an authorized administrator of Scholar Works at UT Tyler. For more information, please contact tgullings@uttyler.edu.



STUDY PROTOCOL

REVISED (Semi)automated approaches to data extraction for systematic reviews and meta-analyses in social sciences: A living review protocol [version 2; peer review: 2 approved, 1 approved with reservations]

Amanda Legate , Kim Nimon

Human Resource Development, University of Texas at Tyler, Tyler, Texas, 75799, USA

V2 First published: 12 Sep 2022, 11:1036
<https://doi.org/10.12688/f1000research.125198.1>
 Latest published: 27 Jan 2023, 11:1036
<https://doi.org/10.12688/f1000research.125198.2>

Abstract

Background: An abundance of rapidly accumulating scientific evidence presents novel opportunities for researchers and practitioners alike, yet such advantages are often overshadowed by resource demands associated with finding and aggregating a continually expanding body of scientific information. Across social science disciplines, the use of automation technologies for timely and accurate knowledge synthesis can enhance research translation value, better inform key policy development, and expand the current understanding of human interactions, organizations, and systems. Ongoing developments surrounding automation are highly concentrated in research for evidence-based medicine with limited evidence surrounding tools and techniques applied outside of the clinical research community. Our objective is to conduct a living systematic review of automated data extraction techniques supporting systematic reviews and meta-analyses in the social sciences. The aim of this study is to extend the automation knowledge base by synthesizing current trends in the application of extraction technologies of key data elements of interest for social scientists.

Methods: The proposed study is a living systematic review employing a partial replication framework based on extant literature surrounding automation of data extraction for systematic reviews and meta-analyses. Protocol development, base review, and updates follow PRISMA standards for reporting systematic reviews. This protocol is preregistered in OSF: (Semi)Automated Approaches to Data Extraction for Systematic Reviews and Meta-Analyses in Social Sciences: A Living Review Protocol on August 14, 2022.

Conclusions: Anticipated outcomes of this study include: (a) generate insights supporting advancement in transferring existing reliable methods to social science research; (b) provide a foundation for protocol development leading to enhancement of comparability and

Open Peer Review

Approval Status

	1	2	3
version 2 (revision) 27 Jan 2023	 view	 view	
	↑	↑	
version 1 12 Sep 2022	 view	 view	 view

1. **Frederick L. Oswald** , Rice University, Houston, USA
2. **Michèle B. Nuijten** , Tilburg University, Tilburg, The Netherlands
3. **Sean Rife** , Murray State University, Murray, USA

Any reports and responses or comments on the article can be found at the end of the article.

benchmarking standards across disciplines; and (c) uncover exigencies that spur continued value-adding innovation and interdisciplinary collaboration for the benefit of the collective systematic review community.

Keywords

Automated data extraction, systematic review, meta-analysis, evidence synthesis, social science research, APA Journal Article Reporting Standards (JARS)



This article is included in the **Artificial Intelligence and Machine Learning** gateway.



This article is included in the **Living Evidence** collection.

Corresponding author: Amanda Legate (alegate@patriots.uttyler.edu)

Author roles: **Legate A:** Conceptualization, Data Curation, Investigation, Methodology, Project Administration, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Nimon K:** Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project was partially supported through research funding support provided by The Office of Research and Scholarship at The University of Texas at Tyler.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2023 Legate A and Nimon K. This is an open access article distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Legate A and Nimon K. **(Semi)automated approaches to data extraction for systematic reviews and meta-analyses in social sciences: A living review protocol [version 2; peer review: 2 approved, 1 approved with reservations]** F1000Research 2023, 11:1036 <https://doi.org/10.12688/f1000research.125198.2>

First published: 12 Sep 2022, 11:1036 <https://doi.org/10.12688/f1000research.125198.1>

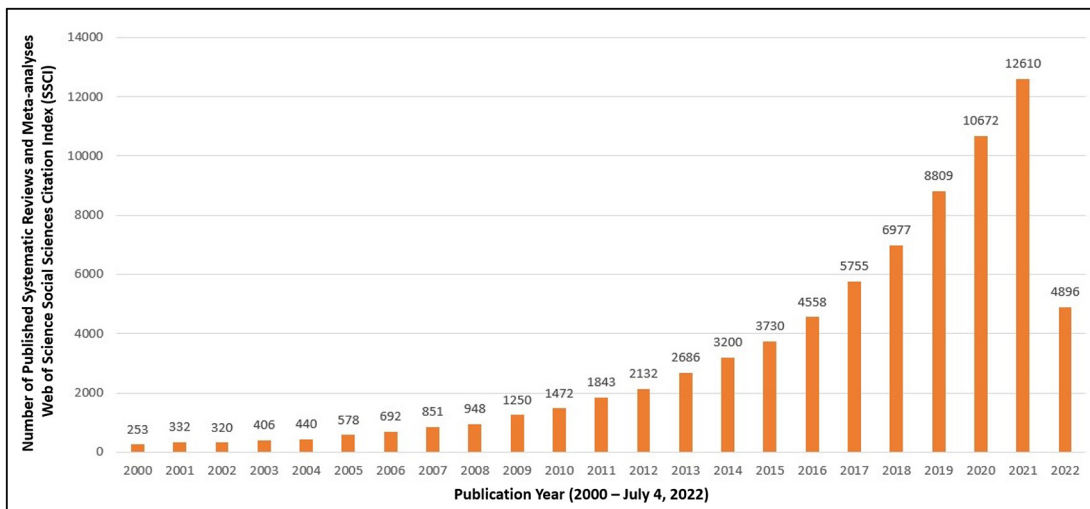
REVISED Amendments from Version 1

In response to feedback from all peer review reports, the following adjustments have been incorporated into the revised version of the protocol. Sections describing research objectives and methodological rationale have been modified to enhance clarity surrounding (a) objectives and scope of the proposed review, (b) study aims in light of similar efforts targeting medical research domains, and (c) identification of existing and/or emerging tools. Key items of interest have been updated to reflect potential outcomes associated with (a) benchmarking and performance assessment, (b) summary recommendations related to domain specific challenges, and (c) identification of code repositories. Revisions to search strategy and outcome reporting include additional description to better explicate (a) goals related to eligibility criteria, (b) rationale for exclusion criteria over and above inclusion criteria, (c) procedures for assessing and reporting reliability of screening and coding activities, and (d) plan for presentation of results. Adjustments to extended data include the addition of a datafile containing tabled data elements for comparison of reporting guidelines and a revised extraction techniques document incorporating a more comprehensive list of systems architectures. Figure 2 has been reformatted to enhance clarity surrounding APA guidance for applying tables/modules based on research design. As recommended by two reviewers, Figures 3 and 4 have been removed. Minor edits include word or phrasing amendments to improve accuracy and clearness.

Any further responses from the reviewers can be found at the end of the article

Introduction

Across disciplines, systematic reviews and meta-analyses are integral to exploring and explaining phenomena, discovering causal inferences, and supporting evidence-based decision making. The concept of metascience represents an array of evidence synthesis approaches which support combining existing research results to summarize what is known about a specific topic (Davis *et al.*, 2014; Gough *et al.*, 2020). Researchers use a variety of systematic review methodologies to synthesize evidence within their domains or to integrate extant knowledge bases spanning multiple disciplines and contexts. When engaging in quantitative evidence synthesis, researchers often supplement the systematic review with meta-analysis (a principled statistical process for grouping and summarizing quantitative information reported across studies within a research domain; Shamseer *et al.*, 2015). As technology advances, in addition to greater access to data, researchers are presented with new forms and sources of data to support evidence synthesis (Bosco *et al.*, 2017; Ip *et al.*, 2012; Wagner *et al.*, 2022). An abundance of accumulated scientific evidence presents novel opportunities for translational value, yet advantages are often overshadowed by resource demands associated with locating and aggregating a continually expanding body of information. In the social sciences, the number of published systematic reviews and meta-analyses have experienced continual growth over the past 20 years, with an annual increase approximating 21% based on citation reports from *Web of Science* (see Figure 1).



Source: Web of Science Core Collection (<https://clarivate.com/webofsciencegroup/solutions/web-of-science-core-collection/>)

Figure 1. Social sciences systematic review and meta-analysis publications by year. Note. Figure was generated using the Web of Science Core Collection database. A title search was conducted in the Social Sciences Citation Index (SSCI) for articles and reviews published between 2000-2022 including variations of the terms "Systematic Review" and "Meta-analysis". Search Syntax: ((TI=("meta-analy*" or "meta analy*" or metaanaly* or "system* review" or "literature review"))) AND PY=(2000-2022)) AND DT=(Article OR Review).

Background

Comprehensive data extraction activities associated with evidence synthesis have been described as time-consuming to the point of critically limiting the usefulness of existing approaches (Holub *et al.*, 2021). Moreover, research indicates that it can take several years for original studies to be included in a new review due to the rapid pace of new evidence generation (Jonnalagadda *et al.*, 2015). As such, research communities are increasingly interested in the application of automation technologies to reduce the workload associated with systematic reviews. Tsafnat *et al.* (2014, p. 2) delineated fifteen tasks associated with systematic reviews and meta-analyses, illuminating the automation potential for each — including the steps involved in repetitive data extraction. Recent studies and conference proceedings have outlined critical factors influencing the development and adoption of automation efforts across social and behavioral sciences. Including, but not limited to, (a) an absence of tools developed for use outside of medical science research (Marshall & Wallace, 2019); (b) a lack of universal terminology (Gough *et al.*, 2020); and (c) nonuniformity in presenting and reporting data (Yarkoni *et al.*, 2021). Notwithstanding these contributions, important questions related to how social scientists are addressing known challenges remain unanswered.

The need for this review in social sciences

Social sciences encompass a broad range of research disciplines, however, what social scientists share is an interest in expanding a collective understanding of human behaviors, interactions, systems, and organizations (National Institute of Social Sciences, n.d.). Systematic reviews and meta-analyses are fundamental to supporting reproducibility and generalizability of research surrounding social and cultural aspects of human behavior, however, the process of extracting data from primary research is a labor-intensive effort, fraught with the potential for human error (see Pigott & Polanin, 2020; Yu *et al.*, 2018). In contrast with the more defined standards that have evolved throughout the clinical research domain, within and across social sciences, substantial variation exists in research designs, reporting protocols, and even publication outlet standards (Davis *et al.*, 2014; Short *et al.*, 2018; Wagner *et al.*, 2022). Notwithstanding that application of automation technologies in the social sciences could benefit from greater standardization of reporting protocols and terminology, understanding of the current state of (semi) automated extraction across these disciplines is largely speculative.

In the clinical research community, automation technologies are rapidly evolving for data extraction. Tools applying intelligent technologies for the purpose of data extraction are increasingly common for research involving Randomized Control Trials (RCT; see Schmidt *et al.*, 2021). As data elements targeted for extraction from clinical studies and healthcare interventions often differ from those targeted by social scientists, transferability of technological solutions remains constrained. Figure 2 presents a general overview of methodologies covered by quantitative reporting guidelines applicable to social sciences per the American Psychological Association (APA, 2020). As of 2018, the APA Journal Article Reporting Standards (JARS) were updated to include clinical trial reports (represented in Figure 2 by the block labeled “Clinical Trials Module C”), incorporating elements also identified by the Cochrane Handbook for Systematic Reviews of Interventions; Higgins *et al.*, 2022).

To elaborate, in health intervention research, targeted data elements generally include Population (or Problem), Intervention, Control, and Outcome (i.e., PICO; see Eriksen *et al.*, 2018; Tsafnat *et al.*, 2014). In social science research, elements targeted for extraction are similarly a function of study design, but targets can take numerous forms based on research questions considered. Researchers in social sciences often rely on APA JARS guidelines, which delineate key elements and respective reporting locations for authors to follow when presenting results of qualitative (JARS-Qual) and quantitative (JARS-Quant) research (APA, 2020; Appelbaum *et al.*, 2018; see also Purdue Online Writing Lab, n.d.). For example, in addition to descriptive statistics (e.g., sample size, mean, standard deviation), meta-analytic efforts typically aim to extract and aggregate inferential elements such as effect sizes and *p*-values. Where structural equation models are involved, a researcher may be interested in extracting model fit indices; or when conducting a reliability generalization or psychometric meta-analysis (see Hunter & Schmidt, 1996), extraction of instrument psychometric properties would be imperative (Appelbaum *et al.*, 2018; see “Extended Data” for supplementary files containing target data elements).

Given that evidence-based medicine is often associated with superior protocol standards and systematic guidelines (i.e., gold standards; Grimmer *et al.*, 2021), the task of transferring even the most reliable automation technologies to social science research presents a substantial challenge. Even within more technical disciplines, such as Information Systems, researchers grapple with automation challenges associated with a lack of uniformity in description and presentation of constructs and measurement items (Wagner *et al.*, 2022, p. 12). While discourse surrounding the delayed uptake of automation tools in the social sciences is occurring, the question of application transferability to domains outside of clinical research remains underexplored. Despite known barriers, delays in interdisciplinary methodological progress inhibit opportunities for collaborative knowledge synthesis both within and across fields (Gough *et al.*, 2020). If automation techniques experiencing rapid growth in clinical research hold potential for transferability to the range of

Quantitative Studies <i>Quantitative studies regardless of design (Table 1, p. 7)</i>		
Experimental Manipulations <i>Studies involving experimental manipulation (Table 2, p. 11)</i>		Nonexperimental Manipulations <i>Studies that do not involve experimental manipulation (Table 3, p. 14)</i>
Randomized <i>Module A: use random assignment</i>	Nonrandomized <i>Module B: do not use random assignment</i>	Clinical Trials <i>Module C: evaluate effects of health-related interventions</i>
Single Design (N-of-1) <i>Studies conducted on a single individual (Table 5, p. 16)</i>		
Longitudinal <i>Studies that collect data on more than one occasion (Table 4, p. 13)</i>		
Replication <i>Studies intended to replicate a prior study (Table 6, p. 17)</i>		
Structural Equation Modeling <i>Studies involving specification of structural and/or measurement models (Table 7, p. 18)</i>		
Bayesian Techniques <i>Studies applying Bayesian analysis techniques (Table 8, p. 20)</i>		
Meta-Analyses <i>Studies reporting research syntheses and meta-analysis (Table 9, p. 21)</i>		

Figure 2. The APA Journal Article Reporting Standards (JARS) by study design summary. *Note.* Figure adapted from Appelbaum *et al.* (2018, Tables 1-9). APA JARS recommendations are outlined in a series of tables and modules addressing varying quantitative study designs; a singular table or combination of tables may apply to a given research report. Please note that Figure 2 includes current APA guidance released as of 2018; new tables may be added over time, please visit <https://apastyle.apa.org/jars/quantitative> for more details.

study designs prevalent throughout social and behavioral sciences, benefits could be far reaching for the development and validation of theoretical models, measurement scales, and much more.

Objectives of this living review

The purpose of this study is to conduct a living systematic review (LSR) to extend the automation knowledge base by identifying existing and emergent systems for (semi) automated extraction of data used by social science researchers conducting systematic reviews and meta-analyses. We aim to uncover and present evidence that can serve as a companion project to ongoing research by Schmidt *et al.* (2021) who are summarizing the state of (semi)automated data extraction technologies for systematic reviews in medical research. As such, the present study holds potential to complement extant scholarship in systematic review extraction technologies by lending support for side-by-side comparison of evidence emerging from the social sciences domains with existing evidence from research medical domains. Following Schmidt *et al.* (2020b, 2021), who are conducting a review of data extraction techniques for healthcare interventions (i.e., RCTs, case studies, and cohort studies; see Schmidt *et al.*, 2020a), we apply an adapted version of their methodological strategy for social science disciplines where observational research is widespread practice.¹ This effort entails targeting extraction of JARS data elements identified by the APA (Appelbaum *et al.*, 2018; see “Extended Data”).

Employing a differentiated replication framework, we apply the LSR methodology to iteratively aggregate and report: (a) the extant state of technology-assisted data extraction in social science research; (b) application trends in automation tools/techniques for extraction of data from abstracts and full text documents outside of biomedical and clinical research corpora; (c) evidence synthesis stages and tasks for which automation technologies are predominantly applied across

¹The authors of the present study do not intend to utilize techniques developed by Schmidt *et al.* (2020a, 2020b, 2021) for automating search, retrieval, and relevance screening tasks associated with the LSR.

social science disciplines; (d) specific data elements and structures targeted for automated extraction efforts by social science researchers; and (e) applied benchmarking standards for performance evaluation.

Related research

To inform this protocol and assess the extent to which our questions have been addressed in prior literature, we explored existing (semi) automated data extraction reviews. We identified six literature reviews (three static, one living, one scoping, and one cross-sectional pilot survey), two software user surveys, and one conference proceeding report. [Table 1](#) provides a summary of scoped studies. Where some efforts focused on software applications, or “tools” that perform or assist with systematic review tasks ([Harrison *et al.*, 2020](#); [Scott *et al.*, 2021](#)), others directed attention to underlying methods or techniques (e.g., machine learning algorithms) or reviewed multiple categorizations (see [Blaizot *et al.*, 2022](#); [Schmidt *et al.*, 2021](#); [O'Connor *et al.*, 2019](#)).

The extant knowledge base and ongoing developments surrounding systematic review automation are highly concentrated in research for evidence-based medicine (e.g., medical research, clinical trials, healthcare interventions) with limited evidence supporting how automation techniques are applied outside of the medical community (see [O'Connor *et al.*, 2019](#)). This is not surprising given the unique relevance of systematic reviews for informing healthcare practice and policy development ([Moher *et al.*, 2015](#)). However, while technologies to support data extraction from primary literature have advanced rapidly, many existing tools were not developed for application outside of research on the effectiveness of health-related interventions. [O'Mara-Eves *et al.* \(2015\)](#), for example, reported that text-mining techniques for classifying and prioritizing (i.e., ranking) relevant studies had undergone substantial methodological advancement, yet also highlighted that where assessment methods could be implemented with relatively high confidence in clinical research, much work was needed to determine how systems might perform in other disciplines. Other researchers similarly noted issues such as heterogeneity in testing and performance metrics ([Blaizot *et al.*, 2022](#); [Jonnalagadda *et al.*, 2015](#); [Tsafnat *et al.*, 2014](#)) as well as risk of systemic biases resulting from inconsistent annotations in training corpora ([Schmidt *et al.*, 2021](#)). Across projects reviewed, calls resounded for additional assessment of automation methods, including testing methods across different datasets and domains and testing the same datasets across different automation methods ([Schmidt *et al.*, 2021](#), [O'Mara-Eves *et al.*, 2015](#); [O'Connor *et al.*, 2019](#); [Jonnalagadda *et al.*, 2015](#)). Despite research presenting evidence of trends toward more complete reporting (i.e., past five years; [Schmidt *et al.*, 2021](#)), dialogue

Table 1. Summary of existing reviews.

Reference	Method/Article Type (Sample Size)	Discipline(s) or Field(s) of Interest	Primary focus of project
Blaizot <i>et al.</i> (2022)	Systematic Review ($n=12$)	Health Science Research	Use of AI methods in healthcare reviews
Harrison <i>et al.</i> (2020)	Feature Analysis ($n=15$), User Survey ($n=6$)	Healthcare Research	Software tools supporting T&Ab screening for healthcare research
Holub <i>et al.</i> (2021)	Systematic Review & Cross-sectional Pilot Survey ($n=78$)	Clinical Trials (RCT)	Data extraction according to tabular structures
Jonnalagadda <i>et al.</i> (2015)	Systematic Review ($n=26$)	Biomedical Research, Clinical Trials (RCT)	Data extraction from full text; Biomedical information extraction algorithms
O'Connor <i>et al.</i> (2019)	Report/Conference Proceeding (ICASR); approx. 50 participants	Interdisciplinary	Maximizing use of technology for transfer of scientific research findings to practice
O'Mara-Eves <i>et al.</i> (2015)	Systematic Review ($n=44$)	Multidisciplinary (not specified)	Text mining technologies for (semi) automating citation/T&Ab screening
Schmidt <i>et al.</i> (2021)	Living Systematic Review ($n=53$)	Medical/ Epidemiological Research, Clinical Trials	Methods/tools for (semi) automating data extraction in SR research
Scott <i>et al.</i> (2021)	User Survey ($n=253$)	Human Health Interventions	SR automation tool use
Tsafnat <i>et al.</i> (2014)	Scoping Review/Survey of Literature	Evidence-based medicine (RCT)	Support or automate processes of SR and/or each task of SR

Note. AI = Artificial Intelligence; T&Ab = Title and Abstract; RCT = Randomized Control Trials, ICASR = International Collaboration for Automation of Systematic Reviews; SR = Systematic Review.

emerging from the systematic review community indicates that the time is ripe for dedicating more attention toward enhancing interdisciplinary comparability and benchmarking standards (O'Connor *et al.*, 2019).

Existing platforms are available to support research teams in a range of time-consuming manual tasks (Blaizot *et al.*, 2022). Even with these expediciencies, not all key activities within the overall review process have received equal attention in application and technique development (O'Connor *et al.*, 2019; Scott *et al.*, 2021). Only a few years ago (semi) automated screening approaches such as text-mining for processing full-texts were not commonly available (O'Mara-Eves *et al.*, 2015). As relevant study details were not always included in abstracts and often appeared throughout and across various sections of a given study (including tables and figures) discussion turned toward development of data extraction methods supporting full-text corpora (Tsafnat *et al.*, 2014). Today, researchers supporting evidence-based medicine benefit from more robust data extraction techniques; especially efforts targeting PICO-related elements (Schmidt *et al.*, 2021). Software tools are available for data extraction (e.g., Abstracktr, Robot Reviewer, SWIFT-Review; see Blaizot *et al.*, 2022, p. 359), however, they have received mixed reviews related to their respective effectiveness. Notwithstanding substantial methodological strides in recent years, limited multidisciplinary reviews evaluating application effectiveness in non-clinical contexts may offer some explanation for the reported delays in uptake outside of evidence-based medicine. Further, the nominal extant research comparing techniques applied in both clinical and social contexts suggests that existing tools may not “perform as well on ‘messy’ social science datasets” (Miwa *et al.*, 2014; as cited in O'Mara-Eves *et al.*, 2015, p. 16). Even within structured tabular reporting contexts (i.e., tables), our understanding of technology applicability and transferability across disciplines is limited (Holub *et al.*, 2021).

Serving as a model for the present study, Schmidt *et al.* (2021) reviews tools and techniques available for (semi) automated extraction of data elements pertinent to synthesizing the effects of healthcare interventions (see Higgins *et al.*, 2022). Their noteworthy living review is exploring a range of data-mining and text classification methods for systematic reviews. The authors uncovered that early often employed approaches (e.g., rule-based extraction) gave way to classical machine-learning (e.g., naïve Bayes and support vector machine classifiers), and more recently, trends indicate increased application of deep learning architectures such as neural networks and word embeddings (for yearly trends in reported systems architectures, see Schmidt *et al.*, 2021, p. 8). Overall, the future of automated data extraction for systematic reviews and meta-analytic research is very bright. As the earlier (i.e., preliminary) stages of the systematic review process have experienced rapid advancement in functionality and capability, development of techniques for all stages is foreseeable in the near future. Just as software tools and data extraction techniques vary in scope, purpose, and financial commitment, so too will research questions, goals, and study designs. Interdisciplinary groups and applied researchers alike call for increased collaboration to spur innovation and further advance the state of computer-assisted evidence synthesis (O'Mara-Eves *et al.*, 2015; O'Connor *et al.*, 2019). Though it can be inferred that not all developments spawned by the medical sciences community are easily transferrable to social sciences, necessity in fields inundated with new evidence production has carved a path for other disciplines; a path in which challenges and opportunities are openly displayed to serve as a foundation for the entire systematic review community to build upon. Additional inquiries surrounding approaches applied in social sciences may introduce previously unencountered demands that spur innovation and create valuable contributions for the entire systematic review community.

Protocol

A LSR involves similar resource demands as would a static review, but is ongoing (i.e., continually reprised). The methodological rationale for selecting LSR for the proposed study is based predominantly on the pace of emerging evidence (Khamis *et al.*, 2019). Given the uncertainty surrounding existing evidence, and the rapid pace of technological advancement, continual surveillance will allow for faster presentation of new and emergent information that may impact findings and offer value for readers (Elliott *et al.*, 2014, 2017). Further, as this review targets published articles, the LSR methodology provides for continual search and retrieval to identify newly developed tools or techniques for which associated publications may not yet be available during previous searches. The following sections present the planned methodological approach of our living review.

Protocol registration and guidelines

This protocol is pre-registered in the Open Science Framework (OSF), an openly accessible repository facilitating the management, storage, and sharing of research processes and pertinent data files (Soderberg, 2018). This protocol adheres to the PRISMA-P guidelines (Moher *et al.*, 2015; Shamseer *et al.*, 2015). A completed PRISMA-P checklist is available at (Semi) Automated Approaches to Data Extraction for Systematic Reviews and Meta-Analyses in Social Sciences: A Living Review Protocol (<https://osf.io/j894w>). No human subjects are involved in this study.

Search sources

Search strategy for this review follows existing research with protocol strategy adapted to fit goals and key elements of interest in social science domains. The model study initiated a LSR of processes supporting the (semi) automation of data

extraction from research studies (e.g., clinical trials, epidemiological research; Schmidt *et al.*, 2021, p. 26). Drawing upon the successful search strategy implemented by Schmidt *et al.*, (2020b, 2021), we will conduct searches via the *Web of Science Core Collection*, *IEEE Xplore Digital Library*, and the *DBLP Computer Science Bibliography*. Databases and collections specific to clinical, medical, and biomedical literature are excluded from the search strategy (i.e., MEDLINE and PubMed). A preliminary search of *Web of Science* per protocol was conducted; 4,835 records were identified from Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (A&HCI), Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH), and Emerging Sources Citation Index (ESCI). Based on the goals of the proposed study, several adjustments were made to the replicated search syntax. See “Extended Data” for relevant search strategy details, including syntax adjustments and preliminary search results. For the base review, this strategy will be replicated (to the extent possible) for remaining databases and any deviations openly reported in the project repository and subsequent publications.

Workflow

The workflow structure follows existing research and guidance developed by Elliott *et al.* (2017) for transitioning to living status, frequency of monitoring, and incorporation of new evidence (see Figure 3). Transparent reporting of the base review and updates will follow PRISMA guidelines (Page *et al.*, 2021). We intend to report results from the base review and later searches separately (Kahale *et al.*, 2022). As quantity of new citations is unknown, necessary adjustments to the workflow described in this protocol will be detailed in future versions of the manuscript, noted in corresponding PRISMA reporting framework, and made available via the project repository.

Search and updates

The base review will begin upon publication and peer approval of this protocol. The review will be continually updated via living methodological surveys of newly published literature (Khamis *et al.*, 2019). Updates will include search and

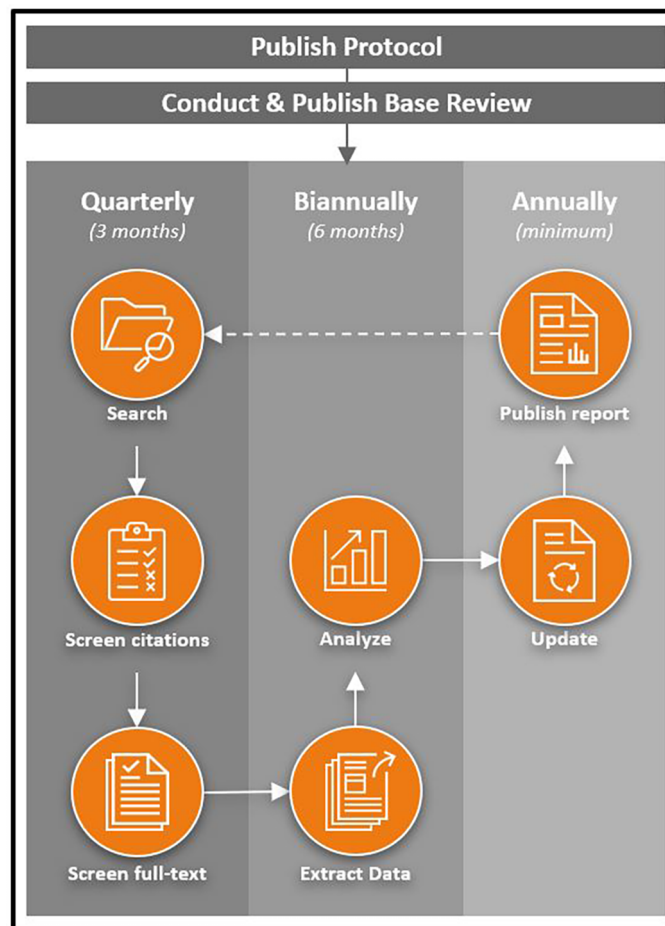


Figure 3. Living review workflow. Note. Arrows represent stages involved in a static systematic review; the dotted line (from “Publish Report” to “Search”) represents the stage at which the review process is repeated from the beginning while the review remains in living status.

screening of new evidence quarterly (every three months) with a cross-sectional analysis of relevant full texts at intervals of no less than twice per year (Khamis *et al.*, 2019). Synthesis and publication of new evidence arising from continual surveillance will occur no less than once per year or until the review is no longer in living status.

Citation and abstract screening will be coordinated using Rayyan (Ouzzani *et al.*, 2016). All citations (i.e., titles and abstracts) identified by the search(es) and all full-text documents retrieved will be screened in duplicate, with blinded and independent study selection by each researcher to reduce the risk of bias. Data extraction and coding of full text articles meeting inclusion criteria will follow the same procedure.

As screening and coding decisions required for this investigation will involve subjective judgment, the researchers will make concerted effort to strengthen transparency and replicability by conducting and reporting intercoder (i.e., interrater) reliability (IRR) assessments at multiple points throughout the study workflow (Belur *et al.*, 2018). IRR assessment and discussion will take place immediately following completion of initial title and abstract screening, upon completion of full text screening, and again following coding of included studies. IRR assessment will be repeated for each search and update phase of the living review as detailed in Figure 3. Reliability estimates will be reported for percent agreement ($\alpha_0=0.15$) and Gwet's AC_1 chance-adjusted index (see Zhao *et al.*, 2022). In the event of unresolvable disagreement(s) between researchers where consensus cannot be reached via discussion, additional qualified reviewer(s) will be consulted. All relevant details pertaining to coding decisions, resolutions and/or procedural adjustments, and underlying data will be uploaded to the project repository (see "Data Availability Statement").

Eligibility criteria

Papers considered for inclusion are refereed publications and conference proceedings in social sciences and related disciplines which describe the application of automation techniques or tools to support tasks associated with the extraction of data elements from primary research studies. As in prior reviews, English language reports, published 2005 or later will be considered for inclusion (Jonnalagadda *et al.*, 2015; O'Mara-Eves *et al.*, 2015; Schmidt *et al.*, 2020b, 2021). The model article includes secondary goals related to reproducibility, transparency, and assessment of evaluation methods (Schmidt *et al.*, 2021). The present study will also consider and synthesize reported evaluation metrics; however, we will not exclude studies omitting robust performance tests. To refine and test eligibility criteria, the complete list of *Web of Science* subject categories was reviewed, and inclusion decisions determined jointly by both researchers. Each category was evaluated based on the scientific branches and academic activity boundaries described by Cohen (2021). See the project supplementary data files for category selection procedures and criteria. Subjects deemed appropriate for inclusion in the initial search, title, and abstract screening stages (see Tsafnat *et al.*, 2014) include foundational and applied formal sciences, social sciences, and social science related disciplines. Excluded subjects include natural sciences and applied clinical or medicinal science categories. In all cases, over-inclusion is prioritized to maximize search recall. See "Extended Data" for comprehensive search strategy details.

A key concern when extracting data from research corpora lies in defining the elements to be extracted; a concern equally relevant for technology-supported extraction efforts. Based on results from their ongoing review, Schmidt *et al.* (2021) concluded that extant data extraction literature focuses on (semi)automating the extraction of PICO elements from RCT research reports. By adapting the search strategy used by Schmidt *et al.* (2021), we aim to uncover discussion surrounding the application of technologies supporting data extraction from studies representing alternative research designs which do not rely on PICO reporting standards. To this end, we attempt to identify technologies that are being or have been applied across a broad range of social sciences. Our search strategy applies subject category inclusion filters to promote identification of relevant literature while mitigating potential for redundancy. The full list of included subject categories is available in the project repository (see "Extended Data"). Where an amount of overlap in extraction targets across domains is anticipated, a goal of the present study is to retrieve literature exploring the use of (semi)automated techniques which demonstrate potential for extracting APA defined reporting elements (Appelbaum *et al.*, 2018). The eligibility criteria outlined in the following sections apply to research reports targeted for this review. Tools, technologies, and/or system architectures identified will be included in our review regardless of domain(s) of origin or domain(s) in which they are predominantly applied, given that the citing article meets eligibility criteria.

Screening decisions require exercising subjective judgment; even when coding for predetermined inclusion and exclusion of articles, decision-making can vary by coder. According to SR reliability literature, variation in coding behavior is influenced by multiple factors, including (but not limited to) subject matter expertise, academic background, research experience, and even interpersonal dynamics (Belur *et al.*, 2018). The inclusion and exclusion criteria outlined below provide a coding framework to promote consistency, accuracy, and reproducibility in coding behavior. Variation in level of detail included in abstracts may result in a preliminary inclusion decision based on title and abstract screening and a subsequent exclusion decision based on full-text review. Where some overlap may exist across the inclusion and

exclusion criteria, we elected to use a high level of specificity in developing the coding framework to facilitate detailed documentation of screening decisions for IRR assessment and reporting.

Included papers

Eligible records include those which:

- employ an evidence-synthesis method (e.g., systematic reviews, psychometric meta-analysis, meta-analysis of effect sizes, etc.) and/or present a proof of concept, tool tests, or otherwise review automation technologies.
- apply an existing or proposed tool or technique for the purpose of technology-assisted data extraction from the abstracts or full-text of a literature corpus.
- report on any automated approach to data extraction (e.g., NLP, ML, TM), provided that at least one entity is extracted semi-automatically and sufficient detail is reported for:
 - entities (i.e., data elements) targeted for automated extraction (per APA JARS)
 - location of the extracted entities or data elements (e.g., abstract, methods, results)
 - the automation tool and/or technique used to support data extraction

Excluded papers

Studies considered ineligible for inclusion in this review are those which:

- apply tools or techniques to synthesize evidence exclusively from medical, biomedical, clinical (e.g., RCTs), or natural science research.
- present guidelines, protocols, or user surveys without applying and/or testing at least one automation technique or tool.
- are labeled as editorials, briefs, or opinion pieces.
- do not apply an existing, proposed, or prototype tool or technique for the purpose of technology-assisted data extraction from the abstracts or full-text of a literature corpus (e.g., extraction of citation data only, narrative discussion that is not accompanied by application or testing).
- do not apply automation for the extraction of data from scientific literature (e.g., web scraping, electronic communications, transcripts, or alternative data sources).

Key items of interest

O'Connor *et al.*, (2019) described data extraction activities as the process of “extracting the relevant content data from a paper’s methods and results and the meta-data about the paper” (p. 4), therefore, we primarily target key reporting items for methods and results sections recommended by the APA. Because systematic evidence synthesis involves multiple stages, it is possible to apply multiple extraction techniques within the context of a single study and/or use a single technology to support review tasks across various stages of the same project (Blaizot *et al.*, 2022; Jonnalagadda *et al.*, 2015). Therefore, to support an exhaustive review and accommodate anticipated variation across automation approaches and reporting formats, a secondary area of interest includes identifying all paper sections and SR stages for which data extraction technologies have been applied (see “Extended Data”).

(Semi)automation, as defined by Marshall and Wallace (2019, p. 2) involves “using machine learning to expedite tasks, rather than complete them.” A pervasive theme throughout (semi)automation literature related to performance benchmarking is limited between-study comparability. Based on extant research, we anticipate that most included reports will incorporate basic evaluation metrics (e.g., true/false positives, true/false negatives, error) as well as other commonly reported performance measures (e.g., precision, recall, F1 scores). Though the data may prove otherwise, it is plausible that findings will mirror recent and ongoing reports revealing substantial variation in not only the type of evaluation metrics reported, but in how they are reported. For example, Schmidt *et al.* (2021) highlighted variety in both method and

presentation of recall-precision trade off assessment (e.g., plots, cut offs, probability thresholds). They also noted that underlying algorithms represent different approaches to data extraction at the entity level, adding nuance to performance comparability (e.g., data labels, entity length, pre-classification features, training requirements). We anticipate that literature reporting tool reviews or user surveys may include measures associated with workload, such as burden, efficiency, and utility, or even more subjective assessments such as user experience, cost effectiveness, or intuitiveness of software (Harrison *et al.*, 2020; Scott *et al.*, 2021).

Primary anticipated outcomes include identification of (a) tools/techniques applied to (semi) automate the extraction of data elements from research articles; (b) data elements targeted for extraction based on APA JARS standards; (c) systematic review and meta-analysis stages for which automation technologies are utilized; (d) evaluation metrics reported for applied automation technologies; and (e) where tools or technologies are presented, the potential for transferability across social science domains. Secondary anticipated outcomes include identification of (a) specific sections of research papers from which data is successfully extracted from primary corpora; (b) structure of content extracted using automation technologies; and (c) challenges reported by social science researchers related to the application of (semi) automated data extraction tools or technologies. Primary and secondary outcome items of interest are further described below, and supplementary data files referenced for readers to access additional information.

Primary items of interest

1. Techniques, tools, systems architectures, and/or automation approaches applied for data extraction from research documents (abstracts and full text).
 - A comprehensive list is provided in extended data files; see “Extraction Techniques Revised.docx”.
2. Data elements targeted for extraction using automation technologies as outlined by JARS (APA, 2020) and further explicated by Appelbaum *et al.* (2018, p. 6).
 - A table containing a comprehensive list of targeted data elements is provided in the extended data files, see “Target Data Elements.docx”.
3. Review tasks and stages for which extraction technologies are applied.
 - A list of fifteen tasks along with stage classifications is adapted from Tsafnat *et al.* (2014). See extended data files; “Review Classifications.docx”.
4. Evaluation metrics used to assess performance of the techniques or tools applied to support data extraction.
 - Extended data file “Extraction Techniques Revised.docx” provides a comprehensive list of metrics along with a description for each metric.
5. Transferability, availability, and accessibility of technique or tool. Target questions include:
 - Does the tool or technology easily transfer to research targeting the extraction of non-PICO prescribed elements?
 - Is the technology publicly available for use by social science researchers?
 - Where an established tool or platform was used, is it cataloged in the Systematic Review Toolbox (Marshall *et al.*, 2022)?
 - If a technique or tool is proprietary, is it open source code?
 - Where code is open source, is it maintained in a code repository (e.g., GitHub or GitLab)?

Secondary items of interest

1. Location (e.g., paper section) from which elements were extracted from research documents. To account for expected variation in reporting, paper sections of interest include, but are not limited to:

- (a) Title, abstract, and author notes; (b) Introduction and background; (c) Methods; (d) Results; and (e) Discussion
 - A table containing a comprehensive list of targeted data elements is provided in the extended data files, see “Comprehensive List of Eligible Data Elements.xlsx”.
2. Structure of content from which data entities were extracted (where named).
- Content formats may include unstructured (e.g., text); structured (e.g., tables), or images (e.g., figures).
3. Challenges identified by social science researchers when applying automation technology to support data extraction efforts. Target questions include:
- Are there conditions under which tools are perceived as more (or less) useful than others?
 - How might technologies or tools be enhanced to better support evidence synthesis efforts across social science domains?
 - Do researchers identify reporting practices that may promote or hinder the application of automation technologies?
 - What challenges are associated with varying degrees of human involvement and/or decision-making?

Presentation of results

Reporting of literature search and screening results will follow PRISMA guidelines (Page *et al.*, 2021) and tailored LSR flowchart recommendations by Kahale *et al.* (2022). To maximize comparability with the ongoing review of extraction technologies for medical research, we plan to present the results of data extraction following Schmidt *et al.* (2021), who reported results in tabular, graphical, and narrative formats. Data visualization will consist of tables and figures (e.g., bar and pie charts); each new version of the review will use the same reporting and presentation structure unless new information is uncovered between published reviews that necessitates additional formats. Where appropriate, descriptive statistics will be presented in table format within each published review. Relevant corpus details will be presented in table format; where size and/or graphic requirements limit inclusion in the published review, a table containing corpus details will be maintained in the project repository and referenced in the published manuscript. We will also provide comprehensive underlying data files supporting all reported results. Underlying data files will be maintained in the project repository and updated alongside each new version of the LSR.

Dissemination of information

Authors plan to submit the base review results and subsequent update(s) to F1000Research for publication. Following the FAIR principles (i.e., findable, accessible, interoperable, and reusable; Wilkinson *et al.*, 2016), all corresponding data will be available via the OSF project repository.

Study status

To support preparation of this protocol, a preliminary search was conducted in *Web of Science*. No formal search, screening, or review activities have been initiated. This protocol was preregistered via OSF Registries on August 14, 2022.

Data availability

Underlying data

No data are associated with this article.

Extended data

Repository: (Semi) Automated Approaches to Data Extraction for Systematic Reviews and Meta-Analyses in Social Sciences: A Living Review Protocol. <https://doi.org/10.17605/OSF.IO/YWTF9> (Legate & Nimon, 2022).

Original and revised supplemental data files are available in the linked repository: Updated Supplemental Files: (Semi) Automated Approaches to Data Extraction for Systematic Reviews and Meta-Analyses in Social Sciences: A Living Review Protocol Updated. <https://doi.org/10.17605/OSF.IO/EWFKP> (Legate & Nimon, 2023).

This project contains the following extended data:

- Extraction Techniques Revised.docx – categories and descriptions of data extraction techniques, architecture components, and evaluation metrics of interest
- Review Classifications.docx – review tasks and stages of interest
- Target Data Elements.docx – key elements of interest for targeted data elements
- Comprehensive List of Eligible Data Elements.xlsx – comprehensive list of elements with extraction potential per APA JARS
- Search Strategy.docx – search syntax for preliminary search in Web of Science
- APA & Cochrane Data Elements.xlsx – tabled data elements for Cochrane reviews, APA Module C (clinical trials), and APA (all study designs)

Data are available under the terms of the [Creative Commons Attribution 4.0 International Public License](https://creativecommons.org/licenses/by/4.0/) (CC-BY 4.0).

Reporting guidelines

This protocol follows PRISMA-P reporting guidelines (Moher *et al.*, 2015). Open Science Framework (OSF) Repository: (Semi) Automated Approaches to Data Extraction for Systematic Reviews and Meta-Analyses in Social Sciences: A Living Review Protocol. <https://doi.org/10.17605/OSF.IO/YWTF9> (Legate, & Nimon, 2022).

References

- American Psychological Association: *Publication manual of the American Psychological Association*. 7th ed. 2020.
- Appelbaum M, Cooper H, Kline RB, *et al.*: **Journal Article Reporting Standards for Quantitative Research in Psychology: The APA Publications and Communications Board Task Force report.** *Am. Psychol.* 2018; **73**(1): 3–25.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Belur J, Tompson L, Thornton A, *et al.*: **Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making.** *Soc. Methods Res.* 2018; **50**(2): 837–865.
[Publisher Full Text](#)
- Blaizot A, Veettil S, Saidoung P, *et al.*: **Using artificial intelligence methods for systematic review in health sciences: A systematic review.** *Res. Synth. Methods.* 2022; **13**(3): 353–362.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Cohen E: **The boundary lens: theorising academic activity.** *The university and its boundaries: Thriving or surviving in the 21st Century*. Routledge; 1st ed. 2021; pp. 14–41.
[Publisher Full Text](#)
- Bosco F, Uggerslev K, Steel P: **MetaBUS as a vehicle for facilitating meta-analysis.** *Hum. Resour. Manag. Rev.* 2017; **27**(1): 237–254.
[Publisher Full Text](#)
- Davis J, Mengersen K, Bennett S, *et al.*: **Viewing systematic reviews and meta-analysis in social research through different lenses.** *Springerplus.* 2014; **3**(1): 1–9.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Elliott J, Turner T, Clavisi O, *et al.*: **Living systematic reviews: An emerging opportunity to narrow the evidence-practice gap.** *PLoS Med.* 2014; **11**(2): E1001603.
[Publisher Full Text](#)
- Elliott J, Synnot A, Turner T, *et al.*: **Living systematic review: 1. Introduction—the why, what, when, and how.** *J. Clin. Epidemiol.* 2017; **91**: 23–30.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Eriksen MB, Frandsen TF: **The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review.** *J. Med. Libr. Assoc.* 2018; **106**(4): 420–431.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Gough D, Davies P, Jantvedt G, *et al.*: **Evidence Synthesis International (ESI): Position Statement.** *Syst. Rev.* 2020; **9**(1): 155.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Grimmer J, Roberts ME, Stewart BM: **Machine learning for social science: An agnostic approach.** *Annu. Rev. Polit. Sci.* 2021; **24**: 395–419.
[Publisher Full Text](#)
- Harrison H, Griffin S, Kuhn I, *et al.*: **Software tools to support title and abstract screening for systematic reviews in healthcare: An evaluation.** *BMC Med. Res. Methodol.* 2020; **20**(1): 7.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Higgins JPT, Thomas J, Chandler J, *et al.*: *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*. Cochrane; 2022.
[Reference Source](#)
- Holub K, Hardy N, Kallmes K: **Toward automated data extraction according to tabular data structure: Cross-sectional pilot survey of the comparative clinical literature.** *JMIR Form. Res.* 2021; **5**(11): E33124.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Hunter JE, Schmidt FL: **Cumulative research knowledge and social policy formulation: The critical role of meta-analysis.** *Psychol. Public Policy Law.* 1996; **2**(2): 324–347.
[Publisher Full Text](#)
- Ip S, Hadar N, Keefe S, *et al.*: **A Web-based archive of systematic review data.** *Syst. Rev.* 2012; **1**(1): 15.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Jonnalagadda S, Goyal P, Huffman M: **Automating data extraction in systematic reviews: A systematic review.** *Syst. Rev.* 2015; **4**(1): 78.
[PubMed Abstract](#) | [Publisher Full Text](#)

- Kahale L, Elkhoury R, El Mikati I, *et al.*: **Tailored PRISMA 2020 flow diagrams for living systematic reviews: a methodological survey and a proposal** [version 3; peer review: 2 approved]. *F1000Res.* 2022; **10**: 192.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Khamis A, Kahale L, Pardo-Hernandez H, *et al.*: **Methods of conduct and reporting of living systematic reviews: A protocol for a living methodological survey** [version 1; peer review: 2 approved]. *F1000 Res.* 2019; **8**: 221.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Legate A, Nimon K: **Updated Supplemental Files: (Semi)Automated Approaches to Data Extraction for Systematic Reviews and Meta-Analyses in Social Sciences: A Living Review Protocol.** 2023, January 12.
[Publisher Full Text](#)
- Legate A, Nimon K: **(Semi) automated approaches to data extraction for systematic reviews and meta-analyses in social sciences: A living review protocol.** *OSF. [Dataset].* 2022, August 14.
[Publisher Full Text](#)
- Li T, Higgins JP, Deeks JJ, editors. **Chapter 5: Collecting data.** Higgins JPT, Thomas J, Chandler J, *et al.*, editors. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022).* Cochrane; 2022.
[Reference Source](#)
- Marshall C, Sutton A, O'Keefe H, *et al.*, editors. *The Systematic Review Toolbox.* 2022.
[Reference Source](#)
- Marshall I, Wallace B: **Toward systematic review automation: A practical guide to using machine learning tools in research synthesis.** *Syst. Rev.* 2019; **8**(1): 110–163.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Miwa M, Thomas J, O'Mara-Eves A, *et al.*: **Reducing systematic review workload through certainty-based screening.** *J. Biomed. Inform.* 2014; **51**: 242–253.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Moher D, Shamseer L, Clarke M, *et al.*: **Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement.** *Syst. Rev.* 2015; **4**(1): 1.
[PubMed Abstract](#) | [Publisher Full Text](#)
- National Institute of Social Sciences: n.d. *What Is "Social Science?"*.
[Reference Source](#)
- O'Connor A, Tsafnat G, Thomas J, *et al.*: **A question of trust: Can we build an evidence base to gain trust in systematic review automation technologies?** *Syst. Rev.* 2019; **8**(1): 143.
[PubMed Abstract](#) | [Publisher Full Text](#)
- O'Mara-Eves A, Thomas J, McNaught J, *et al.*: **Using text mining for study identification in systematic reviews: A systematic review of current approaches.** *Syst. Rev.* 2015; **4**(1): 5.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Ouzzani M, Hammady H, Fedorowicz Z, *et al.*: **Rayyan-a web and mobile app for systematic reviews.** *Syst. Rev.* 2016; **5**(1): 210.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Page M, McKenzie J, Bossuyt P, *et al.*: **The PRISMA 2020 statement: An updated guideline for reporting systematic reviews.** *J. Clin. Epidemiol.* 2021; **88**: 105189–105906.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pigott T, Polanin J: **Methodological guidance paper: High-quality meta-analysis in a systematic review.** *Rev. Educ. Res.* 2020; **90**(1): 24–46.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Purdue Online Writing Lab: *APA Style Introduction.* Purdue Online Writing Lab; n.d.
[Reference Source](#)
- Shamseer L, Moher D, Clarke M, *et al.*: **Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation.** *BMJ: British Medical Journal.* 2015; **350**: 1–25.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Schmidt L, Olorisade BK, McGuinness LA, *et al.*: **Data extraction methods for systematic review (semi)automation: A living systematic review** [version 1; peer review: 3 approved]. *F1000Res.* 2021; **10**: 401.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Schmidt L, McGuinness LA, Olorisade BK, *et al.*: **Protocol.** 2020a, March 11.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Schmidt L, Olorisade BK, McGuinness LA, *et al.*: **Data extraction methods for systematic review (semi)automation: A living review protocol** [version 2; peer review: 2 approved]. *F1000Res.* 2020b; **9**: 210.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Scott A, Forbes C, Clark J, *et al.*: **Systematic review automation tools improve efficiency but lack of knowledge impedes their adoption: A survey.** *J. Clin. Epidemiol.* 2021; **138**: 80–94.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Short JC, McKenny AF, Reid SW: **More than words? Computer-aided text analysis in organizational behavior and psychology research.** *Annu. Rev. Organ. Psych. Organ. Behav.* 2018; **5**(1): 415–435.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Soderberg C: **Using OSF to share data: A step-by-step guide.** *Adv. Methods Pract. Psychol. Sci.* 2018; **1**(1): 115–120.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Tsafnat G, Glasziou P, Choong M, *et al.*: **Systematic review automation technologies.** *Syst. Rev.* 2014; **3**(1): 74.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wagner G, Lukyanenko R, Paré G: **Artificial intelligence and the conduct of literature reviews.** *J. Inf. Technol.* 2022; **37**(2): 209–226.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Wilkinson M, Dumontier M, Aalbersberg I, *et al.*: **The FAIR Guiding Principles for scientific data management and stewardship.** *Nature.* 2016; **53**(1): 160018.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Yarkoni T, Eckles D, Heathers JAJ, *et al.*: **Enhancing and accelerating social science via automation: Challenges and opportunities.** *Harv. Bus. Rev.* 2021.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Yu Z, Kraft N, Menzies T: **Finding better active learners for faster literature reviews.** *Empir. Softw. Eng.* 2018; **23**(6): 3161–3186.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Zhao X, Feng GC, Ao SH, *et al.*: **Interrater reliability estimators tested against true interrater reliabilities.** *BMC Med. Res. Methodol.* 2022; **22**: 232.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Peer Review Status:   

Version 2

Reviewer Report 10 February 2023

<https://doi.org/10.5256/f1000research.142586.r161616>

© 2023 Nuijten M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Michèle B. Nuijten 

Meta-Research Center, Tilburg University, Tilburg, The Netherlands

I thank the authors for the updated version of the manuscript. My concerns are all adequately addressed.

One minor point I noted in the authors' added section on the presentation of the results, is that they plan to use bar charts and pie charts. I would tentatively advise against this in favor of a different type of visualization, because bar charts and pie charts tend to obscure important information about the distribution of the data (a lot has been written about this, but see for instance this article: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002128>).

I'm looking forward to seeing the results of this project.

Best wishes,
Michèle Nuijten

References

1. Weissgerber TL, Milic NM, Winham SJ, Garovic VD: Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol.* 2015; **13** (4): e1002128 [PubMed Abstract](#) | [Publisher Full Text](#)

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research methods in psychology, meta-science, automated extraction of statistical results

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 27 January 2023

<https://doi.org/10.5256/f1000research.142586.r161615>

© 2023 Oswald F. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Frederick L. Oswald 

Department of Psychological Sciences, Rice University, Houston, TX, USA

No further comments - the authors were very thoughtful in response, and I think this paper will prove very useful to a very wide community of researchers.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 19 December 2022

<https://doi.org/10.5256/f1000research.137478.r155957>

© 2022 Rife S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Sean Rife 

Murray State University, Murray, KY, USA

This paper proposes a living review of papers that describe automated methods of data extraction in the social sciences that support meta analyses and other types of systematic reviews. This work is needed, as the number of tools of this type is increasing rapidly, and it is increasingly clear that they will play a prominent role in metascientific endeavors going forward. It is well written and clearly explicates the task at hand.

I have two concerns: first, the paper has a unique task at hand, as it is describing an ongoing review of tools that themselves are used to conduct reviews. For reasons I'm not entirely aware of, when first reading the paper, I assumed that the proposed living review would itself be using automated tools (it wasn't until I reached the "Search and updates" section that I realized my error - at which point I had to go back and review the Introduction and Background. I'm not sure if this warrants any changes in the manuscript, but I thought it was worth pointing out. (I am reminded of this cartoon: <https://xkcd.com/1447/>.)

Second, I worry that by focusing on published papers, this project will miss many programs of great interest. Tools for the automated analysis of scientific papers are often developed without any accompanying publication, at least initially (I am thinking here of projects such as Barzooka [<https://github.com/quest-bih/barzooka>] and Jetfighter [<https://github.com/smsaladi/jetfighter>]). Would it not be useful to also monitor code repositories such as GitHub and GitLab?

I also have a few minor notes:

- The last two items in the "Excluded papers" section seem redundant with the criteria laid out in the preceding "Included papers" section (that is, the inclusion criteria should be enough to cover those instances.)
- There are a number of references to "tangent disciplines" to social science. I've not heard this term previously (although perhaps I am just out of the loop?). Perhaps "related disciplines" is a better phrase?
- The list of extraction techniques provided in the OSF repository is extensive, but seems to be limited to relatively sophisticated techniques (e.g., machine learning, deep learning techniques and natural language processing). However, automated extraction tools (e.g., statcheck) perform quite well employing simpler techniques such as text search using regular expressions.
- The word clouds are neat, but I'm not sure they add much beyond what is conveyed in the main text of the manuscript.

Is the rationale for, and objectives of, the study clearly described?

Yes

Is the study design appropriate for the research question?

Yes

Are sufficient details of the methods provided to allow replication by others?

Yes

Are the datasets clearly presented in a useable and accessible format?

Not applicable

Competing Interests: I am a co-founder, employee, and shareholder of scite.ai, which uses technology related to that which is discussed in the manuscript to provide a related service.

Reviewer Expertise: Metascience; social psychology; moral and political psychology; quantitative methods

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Dec 2022

Amanda Legate

Comment

This paper proposes a living review of papers that describe automated methods of data extraction in the social sciences that support meta analyses and other types of systematic reviews. This work is needed, as the number of tools of this type is increasing rapidly, and it is increasingly clear that they will play a prominent role in metascientific endeavors going forward. It is well written and clearly explicates the task at hand.

- Thank you, Dr. Rife, for offering your time and expertise for the improvement of our study protocol. We hope that we have been able to effectively address your concerns in the revised version of the protocol manuscript.

Concern 1

I have two concerns: first, the paper has a unique task at hand, as it is describing an ongoing review of tools that themselves are used to conduct reviews. For reasons I'm not entirely aware of, when first reading the paper, I assumed that the proposed living review would itself be using automated tools (it wasn't until I reached the "Search and updates" section that I realized my error - at which point I had to go back and review the Introduction and Background. I'm not sure if this warrants any changes in the manuscript, but I thought it was worth pointing out. (I am reminded of this cartoon: <https://xkcd.com/1447/>.)

- Thank you for sharing this feedback. We hope that adjustments made to the revised version of protocol help to reduce confusion by more effectively describing the purpose, scope, and goals of the proposed study. We have also added a note explicitly stating that we do not intend to utilize automation tools for search, retrieval, or relevance decision tasks associated with this LSR. See section: "Objectives of this living review" – Footnote 1

Concern 2

Second, I worry that by focusing on published papers, this project will miss many programs of great interest. Tools for the automated analysis of scientific papers are often developed without any accompanying publication, at least initially (I am thinking here of projects such as Barzooka [<https://github.com/quest-bih/barzooka>] and Jetfighter [<https://github.com/smsaladi/jetfighter>]). Would it not be useful to also monitor code repositories such as GitHub and GitLab?

- We understand your concern. We first want to clarify that if tools/technologies are utilized for evidence synthesis, reviews, or otherwise applied in eligible studies, our review will capture this information as long as the author(s) reference the tool in their paper. There is no requirement that the research be conducted by or associated with the developer(s) of the tool/technology to be included in the review. We opted to employ a living review methodology to mitigate potential for overlooking newly developed tools or those currently under development. More specifically, if a tool does not appear in any published research (whether associated with the tool developers or otherwise reviewed, tested, and/or applied for evidence synthesis) during the base review, it can still be identified by later iterations of the search and retrieval stage. We have updated our methodological rationale to add clarity in this regard. See section: "Protocol" (Paragraph 1)
- Second, we agree with your assessment that monitoring code repositories would be

useful. We feel that we can, within the scope of this study, support future research in this task by offering evidence surrounding the current (and evolving) state of automation endeavors across social science domains. To better support this aim and assist in alleviating some concern, we have added code repositories among our key items of interest in the revised protocol. See section: "Primary items of interest" (Number 5)

Minor Notes

I also have a few minor notes:

Minor Notes - Bullet 1

The last two items in the "Excluded papers" section seem redundant with the criteria laid out in the preceding "Included papers" section (that is, the inclusion criteria should be enough to cover those instances.)

- Based on extant discourse related to SR screening behaviors, we incorporated these eligibility criteria to provide a coding framework for transparent documentation and reporting. We anticipate substantial variation in the level of detail included in article abstracts. Where a study may initially be considered for inclusion based on information available in the abstract, full-text review may ultimately result in a decision to exclude. In such cases, decisions to exclude articles based on full-text review must be coded accordingly. Thus, any one record may be coded based on both inclusion and exclusion criteria at different stages of relevance screening. To help address your concern related to redundancy, we have revised the protocol manuscript to better describe the purpose for the exclusion criteria above and beyond the inclusion criteria. See section: "Eligibility criteria" (Paragraph 3)

Minor Notes - Bullet 2

There are a number of references to "tangent disciplines" to social science. I've not heard this term previously (although perhaps I am just out of the loop?). Perhaps "related disciplines" is a better phrase?

- Thank you for this feedback. Both instances have been updated to "related disciplines" in the revised protocol.

Minor Notes - Bullet 3

The list of extraction techniques provided in the OSF repository is extensive but seems to be limited to relatively sophisticated techniques (e.g., machine learning, deep learning techniques and natural language processing). However, automated extraction tools (e.g., statcheck) perform quite well employing simpler techniques such as text search using regular expressions.

- We greatly appreciate you taking the time to review our supplementary files and provide feedback. We agree that the list of techniques that may surface over the course of the review could be more inclusive. A revised document has been added to the OSF repository. The table is amended to include a more comprehensive list of components or architectures that may be identified by the proposed LSR. See section: "Extended data" - Filename "Extraction Techniques Revised.docx"

Minor Notes - Bullet 4

The word clouds are neat, but I'm not sure they add much beyond what is conveyed in the main text of the manuscript.

- Thank you for bringing this to our attention. Figures 3 & 4 have been removed in the revised version of this protocol.

Competing Interests: No competing interests were disclosed.

Reviewer Report 24 November 2022

<https://doi.org/10.5256/f1000research.137478.r155953>

© 2022 Nuijten M. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Michèle B. Nuijten 

Meta-Research Center, Tilburg University, Tilburg, The Netherlands

I think it is useful to have a thorough overview of the available data extraction tools applicable to the social sciences. Meta-analysis and other synthesis methods are increasingly necessary to deal with the increasing scientific output, and automated data extraction can play an important role in this.

The authors present a thorough protocol for their LRS. I am happy to see that they preregistered their protocol and published their detailed supplemental materials online.

I do have some questions/concerns I would like the authors to address.

First, the authors plan to exclude data extraction tools published/applied in the medical sciences. They correctly argue that data extraction tools for the medical sciences often rely on the systematic reporting in medical papers, which means they can often not be applied in the social sciences. However, by excluding ALL tools from the medical sciences, it seems possible to overlook tools that WOULD be applicable to the social sciences as well. This would make the overview of tools in the LRS incomplete. Did the authors consider this?

Second, related to the point above, the authors intend to exclude tools from clinical research, but what does this mean for tools applied in fields that are on the overlap of the social sciences and clinical sciences, such as clinical psychology?

Third, the authors plan to extract a lot of information from the included articles. I think it is important to mention in the protocol how this information will be presented. Will the authors use tables to present everything? Visualizations? A database? Summary statistics? Something else?

Some small additional comments/questions:

- P.8 (of the pdf) Search and Updates: what exactly will the reliability be calculated about? Reliability of what? Also: please define what qualifies as “questionable” inter-rater reliability.
- P.9: why is psychometric assessment considered an evidence-synthesis method?

Best wishes and good luck with this project,

Is the rationale for, and objectives of, the study clearly described?

Yes

Is the study design appropriate for the research question?

Partly

Are sufficient details of the methods provided to allow replication by others?

Yes

Are the datasets clearly presented in a useable and accessible format?

Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research methods in psychology, meta-science, automated extraction of statistical results

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 24 Dec 2022

Amanda Legate

Concern 1

First, the authors plan to exclude data extraction tools published/applied in the medical sciences. They correctly argue that data extraction tools for the medical sciences often rely on the systematic reporting in medical papers, which means they can often not be applied in the social sciences. However, by excluding ALL tools from the medical sciences, it seems possible to overlook tools that WOULD be applicable to the social sciences as well. This would make the overview of tools in the LRS incomplete. Did the authors consider this?

- You highlight a relevant concern. First, we have given this thoughtful consideration. The ongoing LSR conducted by Schmidt et al., (2021) focuses exclusively on semi-automated extraction techniques in medical domains. Additionally, several recent studies (including user surveys of SR automation tools) address application of (semi)automation tools for clinical, medical, and healthcare research. We aim to complement the growing body of automated data extraction research by uncovering evidence related to how, where, and for what purpose(s) technologies are being applied across social science domains. This strategy is intended to reduce redundancy while allowing for comparison against a similar LSR that is currently in progress. That said, we understand that additional clarity surrounding the purpose and scope of our proposed LSR is needed. To address this concern, we have updated the protocol to add clarity surrounding the purpose and scope of the review. See section: "Objectives of this living review" (paragraph 1) - added clarity surrounding

aims of the present study considering similar ongoing efforts targeting medical/clinical research domains.

- Second, we will not exclude tools that have been published and/or applied in the medical sciences. Our goal is to identify tools that have been or are being applied in social science research (regardless of domain of origin). Given the discourse surrounding the lack of tools developed for use outside of the medical research community, we anticipate that existing tools, even those developed for the medical community (or for specific methodologies - e.g., RCTs), have been and/or are being applied/tested outside of those areas. It is possible that a substantial amount these activities are already occurring, however, existing systematic efforts to synthesize and present evidence focus on evidence-based medicine. We hope that by systematically and transparently aggregating what is known about how these tools/technologies are applied in social science research, our study will offer sufficient evidence for side-by-side comparison with existing research and support future research and/or opportunities for application of existing technologies supporting evidence synthesis in social sciences. While our search strategy does not contain exclusion criteria related to the tools/technologies themselves, we have updated the protocol to enhance clarity surrounding our search strategy goals and eligibility criteria. (A) See section: "Eligibility criteria" (Paragraph 2) – added paragraph to better describe search strategy goals and exclusion criteria. (B) See section: "Excluded papers" (Bullet 1) - added "exclusively" to enhance specificity in exclusion criteria.

Concern 2

Second, related to the point above, the authors intend to exclude tools from clinical research, but what does this mean for tools applied in fields that are on the overlap of the social sciences and clinical sciences, such as clinical psychology?

- Our search strategy applies subject category filters with the goal of identifying studies that are not already included in extant projects addressing clinical research (e.g., clinical trials). However, a variety of potentially overlapping subject categories are targeted for inclusion in our search strategy (e.g., applied psychology, developmental psychology, experimental psychology, social psychology, among many other subjects). We limited the scope of our study not to exclude potentially transferable tools and technologies, but rather, to mitigate potential for research redundancy and focus on an underrepresented area in research automation reviews. In reflecting on your comments, we noticed that our search strategy document was mislabeled in the protocol manuscript such that readers could reasonably assume the document only contains search syntax. We have corrected the filename in our revised protocol. See section: "Extended data" (Bullet 5) - Corrected filename: "Search Strategy.docx"

Concern 3

Third, the authors plan to extract a lot of information from the included articles. I think it is important to mention in the protocol how this information will be presented. Will the authors use tables to present everything? Visualizations? A database? Summary statistics? Something else?

- This is an excellent suggestion, thank you. We have added a section (immediately following key items of interest section) in the revised protocol to describe our plan for presenting results of the LSR. See section: "Presentation of results"

Some small additional comments/questions**Bullet 1**

P.8 (of the pdf) Search and Updates: what exactly will the reliability be calculated about? Reliability of what? Also: please define what qualifies as “questionable” inter-rater reliability.

- Thank you for bringing this omission to our attention. Our revised protocol includes additional detail related to a structured plan for assessing and reporting the reliability of screening and coding activities. See section: “Search and updates” (Paragraphs 2 & 3)

Bullet 2

P.9: why is psychometric assessment considered an evidence-synthesis method?

- Our intent was to reference assessment of psychometric properties (e.g., reliability/validity of instrumentation) which often involves extracting/synthesizing a substantial amount of evidence from published research. This content has been adjusted to reflect more accurate terminology. See section: “Included papers” (Bullet 1)

Closing Comments

Best wishes and good luck with this project.

- Thank you, Dr. Nuijten. We appreciate your thoughtful review and value-enhancing feedback. We hope that we have been able to effectively address your concerns in the revised version of this protocol.

Competing Interests: No competing interests were disclosed.

Reviewer Report 24 October 2022

<https://doi.org/10.5256/f1000research.137478.r152219>

© 2022 Oswald F. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Frederick L. Oswald

Department of Psychological Sciences, Rice University, Houston, TX, USA

1. Authors propose to review the available-and-growing literature, technologies, and practices pertaining to conducting systematic reviews and meta-analysis by automated or semi-automated means. This review will be continuously updated via the living systematic review (LSR).
2. The importance of this work is evident - because meta-analyses are extremely popular, useful, and necessary in many research domains, any technologies that stand to improve the conduct and quality of meta-analyses are welcome.
3. Automated methods could apply not only to the extraction of data, but also to literature

searching itself (e.g., see the Elicit AI assistant at <https://elicit.org/>).

4. Authors might dive deeper on what benchmark standards might be used (overall and within domains); maybe the comparison across the technologies reviewed are relative benchmarks across a variety of dimensions (e.g., errors of omission vs. commission across JARS elements, extent of human involvement, domain-specific effectiveness).
5. From this birds-eye view, the LSR might make summary recommendations of things such as (a) the conditions under which some tools are more useful than others; (b) how specific tools might be improved or technologies advanced more generally; (c) best reporting practices for authors (and reporting errors to avoid); (d) how humans tend to be most helpful (e.g., automation might err on the side of inclusion, for the human to winnow down).
6. Side notes:
 - p. 4 refers to 'social research' but I think this is 'social sciences research'
 - top of Fig 2 looks like a header, but it only refers to the content of Table 1 (in other words, the reader wouldn't want to think that 'Bayesian techniques' are 'recommended for inclusion in manuscripts regardless of design')
 - Instead of the word clouds in Figs 3 and 4, I'd suggest tabling the similarities and differences, or perhaps not including them at all (e.g., the comparison wouldn't be as helpful in domains that do not conduct clinical trials)

Good luck to the authors in pursuing this LSR; I think many scholars (including myself) will find it very useful.

Is the rationale for, and objectives of, the study clearly described?

Yes

Is the study design appropriate for the research question?

Partly

Are sufficient details of the methods provided to allow replication by others?

Partly

Are the datasets clearly presented in a useable and accessible format?

Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Meta-analysis, psychometrics, individual differences, employment testing, open science

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 31 Oct 2022

Amanda Legate

Dear Dr. Oswald,
Thank you for serving as a reviewer for our protocol manuscript. We greatly appreciate your thoughtful feedback. When all reviewer reports have been received and considered, we will incorporate suggestions where appropriate and respond directly to each point outlined in your report.
Kind regards,
Amanda Legate
Dr. Kim Nimon

Competing Interests: No competing interests were disclosed.

Author Response 24 Dec 2022

Amanda Legate**Comment 3**

Automated methods could apply not only to the extraction of data, but also to literature searching itself (e.g., see the Elicit AI assistant at <https://elicit.org/>).

- Thank you for this feedback. We could have better described our goals as they relate to SR stages that are considered relevant within the scope of this review. We have incorporated a statement to enhance clarity related to potential for application across various stages of the review process. See section: "Key items of interest" (Paragraph 1)

Comment 4

Authors might dive deeper on what benchmark standards might be used (overall and within domains); maybe the comparison across the technologies reviewed are relative benchmarks across a variety of dimensions (e.g., errors of omission vs. commission across JARS elements, extent of human involvement, domain-specific effectiveness).

- We agree and believe that incorporating your suggestions will offer readers more context for understanding our research goals and anticipated outcomes related to benchmarking and performance assessment. In the revised protocol, we have included a summary of anticipated LSR outcomes related to benchmarking standards and performance assessment measures. See section: "Key items of interest" (Paragraph 2)

Comment 5

From this birds-eye view, the LSR might make summary recommendations of things such as (a) the conditions under which some tools are more useful than others; (b) how specific tools might be improved or technologies advanced more generally; (c) best reporting practices for authors (and reporting errors to avoid); (d) how humans tend to be most helpful (e.g., automation might err on the side of inclusion, for the human to winnow down).

- This is wonderful feedback, thank you. We have incorporated your suggestions under the item of interest pertaining to challenges identified by researchers that may offer meaningful summary recommendations for the social science research community. See section: "Secondary items of interest" (Number 3)

Comment 6, Bullet 1

p. 4 refers to 'social research' but I think this is 'social sciences research'

- Thank you for bringing this to our attention, we have corrected the sentence. See section: "The need for this review in social sciences" (Paragraph 3)

Comment 6, Bullet 2

top of Fig 2 looks like a header, but it only refers to the content of Table 1 (in other words, the reader wouldn't want to think that 'Bayesian techniques' are 'recommended for inclusion in manuscripts regardless of design')

- We appreciate this feedback; upon thoughtful consideration, we agree that the figure is potentially misleading, especially for audiences less familiar with APA reporting guidelines. In the revised protocol, we have updated the figure and added additional figure notes to enhance clarity surrounding APA JARS guidance for applying tables/modules based on research design. See section: "The need for this review in social sciences" (Figure 2)

Comment 6, Bullet 3

Instead of the word clouds in Figs 3 and 4, I'd suggest tabling the similarities and differences, or perhaps not including them at all (e.g., the comparison wouldn't be as helpful in domains that do not conduct clinical trials)

- Figures 3 and 4 have been removed in the revised version of the protocol. Due to size, a comparative table is inappropriate for inclusion within the body of the manuscript. However, we have added a file containing the tabled data elements for Cochrane reviews, APA Module C (clinical trials), and APA (all study designs) to the OSF project. See section: "Extended data" – Filename "APA & Cochrane Data Elements.xlsx".

Closing Comments

Good luck to the authors in pursuing this LSR; I think many scholars (including myself) will find it very useful.

- Thank you for lending your time and expertise for the improvement of our protocol manuscript. We hope that we have been able to effectively address your concerns in the revised version of this protocol.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research