

A conceptual framework for critical appraisal in systematic mixed studies reviews

Quan Nha Hong¹, Pierre Pluye¹

¹Department of Family Medicine, McGill University, Montréal, Canada

Corresponding Author:

Quan Nha Hong, Department of Family Medicine, McGill University, 5858 Côte-des-Neiges, Suite 300, Montréal, QC, Canada, H3S 1Z1; Tel: 1-514-398-8483, Fax: 1-514-398-4202.

Email address: quan.nha.hong@mail.mcgill.ca

Abstract

The past decade has been rich with methodological advancements in systematic reviews, several of which were inspired by the literature on mixed methods research. Systematic mixed studies reviews, i.e., reviews combining qualitative and quantitative evidence, are increasingly popular as they can provide a better understanding of complex phenomena and interventions. However, they raise new challenges, especially regarding how to perform critical appraisal of the included studies that vary regarding the methodologies used. To address this challenge, conceptually clarifying critical appraisal is necessary. To this end, this paper provides a framework for critical appraisal in systematic mixed studies reviews. This framework is an essential first step toward providing clear guidance on how to perform critical appraisal.

Keywords: critical appraisal, mixed methods research, quality assessment, systematic mixed studies review, systematic review

Introduction

Systematic reviews are considered among the best sources of research evidence, are used for decision-making, and are helpful for coping with the rapidly increasing volume of scientific literature (Bunn et al., 2015; Moynihan, 2004). There has been a call to broaden the scope of systematic reviews and integrate evidence from studies with diverse designs, especially to address the complexity of interventions, implementation, and context (Anderson et al., 2013; Pluye, Hong, Bush, & Vedel, 2016). Systematic reviews that include qualitative, quantitative and mixed methods studies (hereafter, *systematic mixed studies review* [SMSR]) respond to this need (Heyvaert, Hannes, & Onghena, 2016; Pluye & Hong, 2014).

Because of the heterogeneity in the designs of included studies, SMSRs raise several new challenges related to the syntheses of qualitative, quantitative and mixed methods studies and their integration, and the critical appraisal of the quality of included studies (Gough, 2015; Harden & Thomas, 2005). Previous work on SMSRs has focused on understanding how quantitative and qualitative evidence could be synthesized and integrated (Frantzen & Fetters, 2015; Heyvaert, Maes, & Onghena, 2013b; Hong, Pluye, Bujold, & Wassef, 2017; Sandelowski, Voils, Leeman, & Crandell, 2012). However, few papers have addressed the challenges of critical appraisal when the included studies have different designs. This paper focuses on this challenge.

Critical appraisal, the systematic and careful examination of study quality, is an important step in systematic reviews (Burls, 2009; Harden & Gough, 2012). Currently, there are over 500 critical appraisal tools for various study designs (Bai, Shukla, Bak, & Wells, 2012; Deeks et al., 2003; Katrak, Bialocerkowski, Massy-Westropp, Kumar, & Grimmer, 2004; West et al., 2002), but there is no clear guidance regarding which tool and approach to use, nor how or why to use them. This is may be due to a lack of conceptual clarity of what ‘critical appraisal’ means and what is appraised. This paper addresses this knowledge gap by providing a conceptual framework to better understand the components of critical appraisal in SMSRs. We have organized this paper into three main parts. The first provides a definition of SMSRs. The second presents a framework illustrating the different components involved in the critical appraisal process as well as some challenges and debates encountered in SMSRs. The third addresses the implications of the framework and suggests avenues for future research.

Systematic Mixed Studies Reviews

Definition

SMSR follows the principles of mixed methods research (Heyvaert et al., 2016; Pluye & Hong, 2014). In primary research, mixed methods research is often defined based on its core characteristics; that is, the combination of elements of qualitative and quantitative research approaches, namely research question, research design, data collection, data analysis, and results (Creswell & Plano Clark, 2011; Johnson, Onwuegbuzie, & Turner, 2007). Based on these core components, we propose the following definition: Mixed methods research is a research approach in which a researcher or team of researchers integrates (a) qualitative and quantitative research questions, (b) qualitative and quantitative research designs and methods, (c) techniques for collecting and analyzing qualitative and quantitative data, and (d) qualitative findings and quantitative results (Pluye & Hong, 2014).

Applied to secondary research (i.e., literature reviews), the same components of this mixed methods definition can be found with slight differences in the terminology: (a) qualitative and/or quantitative **review** questions, (b) qualitative and/or quantitative **synthesis designs**, and (c) techniques for **extracting** and **synthesizing** qualitative and quantitative data, and (d) qualitative findings and quantitative results **of the synthesis**. SMSR has been defined as a systematic literature review conducted by a team of researchers that includes qualitative, quantitative, and/or mixed methods studies, and uses qualitative and/or quantitative synthesis methods (Heyvaert et al., 2016; Pluye & Hong, 2014). The term “systematic” means that the review uses an explicit, transparent, and reproducible process with (a) specific review question(s) and precise study eligibility criteria; (b) a comprehensive set of information sources, and an exhaustive search strategy designed with specialized librarians; (c) a reliable or dependable (performed by at least two researchers) selection of relevant studies, data extraction, and critical appraisal; and (d) a rigorous synthesis (Pluye et al., 2016).

Two Levels of Integration

Integration can occur at two levels in SMSRs (Figure 1) (Heyvaert et al., 2013b). The first possible level of integration occurs during the selection of studies. A SMSR focuses on synthesizing quantitative and qualitative evidence and includes any combination of qualitative, quantitative, and/or mixed methods studies (see check marks in Figure 1).

The second possible level of integration occurs during the synthesis, i.e., when the extracted data from the included studies are brought together using synthesis methods (Mays, Pope, & Popay, 2005). In SMSRs, there are multiple synthesis method options (Hong et al., 2017). As illustrated in Figure 1, the synthesis methods in SMSRs can be qualitative, quantitative, or mixed.

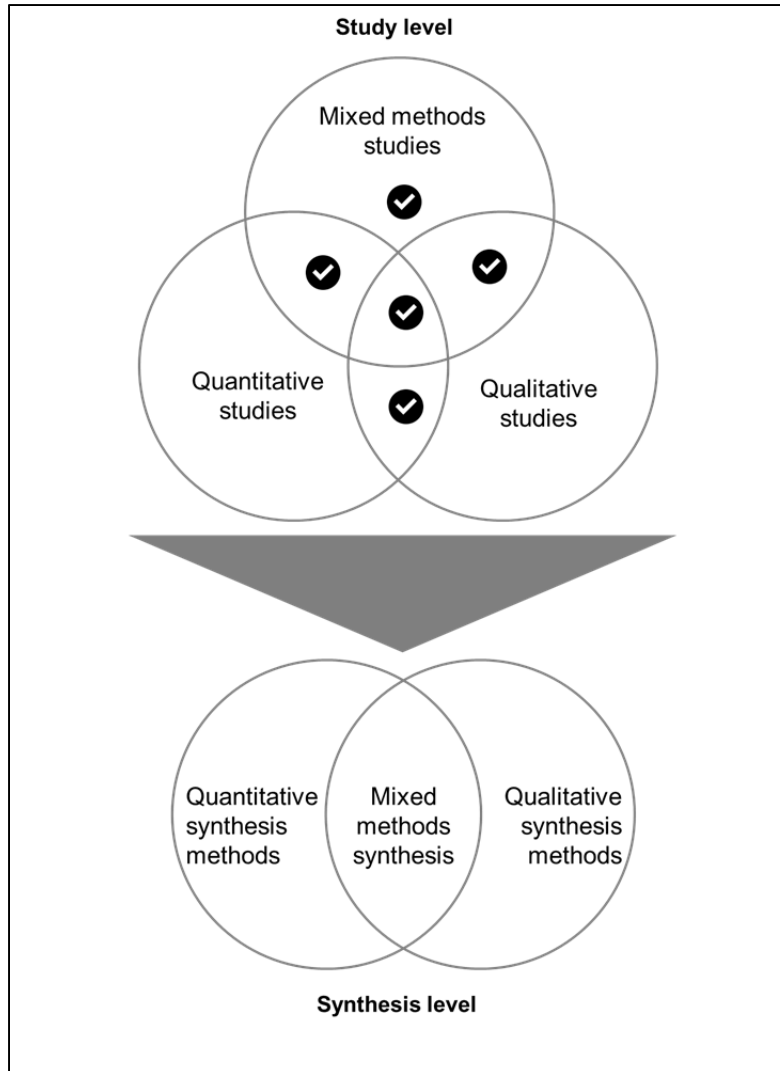


Figure 1. Integration of Studies and Integration of Synthesis Methods in Systematic Mixed Studies Reviews

SMSRs using qualitative synthesis methods will provide a summary or interpretation of data to generate outputs such as themes, concepts, or theories. Several qualitative synthesis methods have been developed such as thematic synthesis, framework synthesis, meta-narrative

synthesis, meta-ethnography, and critical interpretive synthesis (Barnett-Page & Thomas, 2009). Markoulakis and Kirsh (2013) provide an example of a SMSR using qualitative synthesis. They used critical interpretive synthesis (i.e., reciprocal translational analysis, lines of argument synthesis, and refutational synthesis) to develop a theory of difficulties faced by students with mental health issues in the university setting.

SMSRs using quantitative synthesis methods will provide numerical data and summaries of variables of interest of included studies. Basic and advanced meta-analysis methods (e.g., meta-regression and Bayesian synthesis) (Sutton & Higgins, 2008) are well known examples. Roberts, Dixon-Woods, Fitzpatrick, Abrams, and Jones (2002) provide an illustration of a Bayesian synthesis used in a SMSR of factors affecting uptake of childhood immunization. In this review, to establish prior probabilities, the authors transformed the data from the included qualitative studies into quantitative data using quantitative content analysis. Then, these prior probabilities were combined with the results of the included quantitative studies to calculate probabilities that factors might affect immunization uptake.

The synthesis is considered mixed in SMSRs when both quantitative and qualitative synthesis methods are used. For example, Thomas et al. (2004b) conducted a review on the consumption of fruits and vegetables intake among children in which they performed a meta-analysis of controlled trials of the effectiveness of interventions and a thematic synthesis of studies about children's views. Then, the findings of both syntheses were juxtaposed in a matrix to identify interventions that matched the children's views and to further explore if these interventions were more effective.

Critical Appraisal

Critical appraisal is usually performed in systematic reviews to identify the strengths and weaknesses of studies, to determine how much confidence to have in the findings, and to ensure that the recommendations and conclusions properly reflect the quality of evidence reviewed, using sensitivity analysis, for instance; i.e., the comparison of results of lower vs. higher quality studies (Booth, Papaioannou, & Sutton, 2012). Different terms have been used to designate this construct, such as quality appraisal, quality assessment, validity assessment, and assessment of risk of bias (Higgins & Green, 2008). Hereafter, we will use *critical appraisal* to encompass all of these terms. To better understand critical appraisal, we looked at how this construct has been

defined in literature on systematic reviews and how the critical appraisal process was performed in a sample of 459 SMSRs selected in a review of SMSRs (Hong et al., 2017). We compared the different definitions to highlight the commonalities and differences, and to identify the main components. We synthesized our findings into a conceptual framework. That is, we generated a representation of the interrelated constructs that provide a comprehensive understanding of a phenomenon (Jabareen, 2009).

Figure 2 presents the conceptual framework illustrating the process of critical appraisal in SMSRs including three main components: studies, papers and review. Based on this framework, critical appraisal in SMSRs can be defined as: a process related to judging the quality of qualitative, quantitative and mixed methods studies reported in research papers. In this process, three main dimensions of quality can be appraised: methodological, conceptual, and reporting. The purposes and choice of dimensions of quality to judge will vary depending on the objectives and synthesis method(s) adopted in a given review.

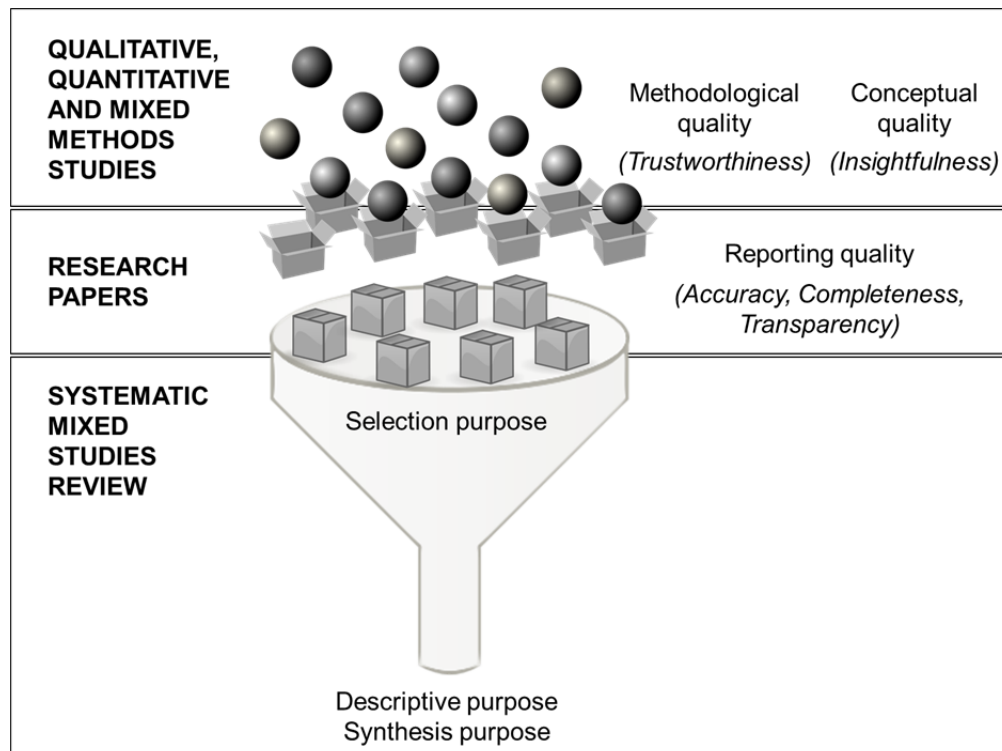


Figure 2. Framework of the Different Components Involved in the Critical Appraisal Process in Systematic Mixed Studies Reviews

The Quality of Quantitative, Qualitative and Mixed Methods Studies

The first component in the framework is *studies*, represented by spheres (Figure 2). Since the unit of synthesis in SMSRs is studies, the judgment made concerns two dimensions of quality: methodological (trustworthiness) and conceptual (insightfulness). Depending on the research designs of the included studies and the review objectives, the criteria used to appraise the methodological and/or conceptual quality will vary.

Methodological quality. Methodological quality is concerned with how a study is conducted. It is usually related to the construct of trustworthiness: is a study good enough for the results to be trustworthy? The judgment made about the trustworthiness of a study is typically related to the methodology and methods used and how biases were minimized (Higgins & Green, 2008; West et al., 2002).

There are two main approaches to appraising methodological quality of studies. In the first approach, studies are ranked based on their designs, with the assumption that some designs produce more credible inferences than others (Wells & Littell, 2009). In this approach, the methodological quality is conceived of as excellence. That is, quality studies meet the highest methodological standards that can yield results closer to the most plausible value. This approach is named the hierarchy of evidence or design hierarchy approach, in which systematic reviews with meta-analysis and randomized controlled trials are considered the best source of evidence (Wells & Littell, 2009). This approach is problematic in SMSRs since qualitative studies are excluded from the hierarchy of designs (Dixon-Woods et al., 2006).

A second approach used to appraise methodological quality of studies is associated with the absence of threats to validity (or risk of bias); the fewer threats or risks, the more trustworthy the results of the study. This threats-to-validity approach differs from the previous one by considering the specific features of a study design rather than contrasting these features with gold standards (Wells & Littell, 2009). One challenge when using this approach in SMSRs concerns the dimensions of trustworthiness that should be appraised. Table 1 presents different dimensions of trustworthiness that can be considered. In several critical appraisal tools, methodological quality refers to the internal validity of a study (Bai et al., 2012). For example, the Cochrane Risk of Bias Tool (Higgins et al., 2011) and the Newcastle-Ottawa Scale (Wells et al., 2000) include criteria that focus on how well a study was done to minimize bias. However, some tools suggest appraising other types of validity such as external validity (Dyrvig, Kidholm, Gerke, &

Vondeling, 2014). There are still diverging views on whether the appraisal should be limited to one or several types of validity and which types are the most important to appraise in SMSRs.

Table 1. Dimensions of Trustworthiness and Comparison of Criteria in Quantitative, Mixed Methods, and Qualitative Research

Dimensions of trustworthiness	Quantitative research	Mixed methods research	Qualitative research
Truthfulness	Internal validity	← + →	Credibility
Applicability	External validity	← + →	Transferability
Consistency	Reliability	← + →	Dependability
Neutrality	Objectivity	← + →	Confirmability

(adapted from Heyvaert et al., 2016)

Another challenge in SMSRs is evaluating and comparing the quality of studies from different epistemological and methodological traditions. As presented in Table 1, the dimensions of trustworthiness differ for quantitative and qualitative research. For mixed methods studies, the qualitative and quantitative components are combined to produce an integration that is greater than the sum of each component (Fetters & Freshwater, 2015). This might preclude the use of a single critical appraisal instrument for all included studies in SMSRs. There remains a lack of consensus on how critical appraisal should be performed and what criteria should be used, especially for qualitative and mixed methods studies (Carroll & Booth, 2015). Reviews on the quality in mixed methods have identified up to 13 different checklists for appraising mixed methods studies (Heyvaert, Hannes, Maes, & Onghena, 2013a) and 19 quality criteria (Fàbregues & Molina-Azorín, 2017). From our review on SMSRs, we identified four main approaches that were used for appraising the quality of mixed methods studies. One approach is to use specific criteria for the quantitative and qualitative components of the studies. To exemplify, several SMSRs used different tools such as the CASP tool for qualitative studies (Critical Appraisal

Skills Programme (CASP), 2017) and the Effective Public Health Practice Project (EPHPP) tool for quantitative studies (Thomas, Ciliska, Dobbins, & Micucci, 2004a). A second approach is to use generic criteria that could be applied to all studies, such as the assessment form suggested by Hawker, Payne, Kerr, Hardey, and Powell (2002). A third approach consists of using specific criteria for mixed methods studies. For example, some used the Mixed Methods Appraisal Tool (MMAT) that includes qualitative, quantitative, and mixed methods criteria (Pluye et al., 2011). A final approach is to appraise only the dominant component (qualitative or quantitative) of a mixed methods study.

Conceptual quality. Conceptual quality is defined as how clearly a concept is articulated to facilitate theoretical insight (Toye et al., 2013). This dimension of quality is related to insightfulness: does the study provide a clear, rich, and deep understanding of a phenomenon? This dimension has been explored in a study on the process of critical appraisal of qualitative studies (Toye et al., 2013). The authors found that conceptual clarity was an important dimension of quality used by reviewers to determine the inclusion of qualitative studies in a review using meta-ethnography. This dimension is linked with clarity but also with depth of description providing rich insight into a concept (Toye et al., 2013). Some reviews have also used this dimension of quality to appraise quantitative studies. For example, Beaugard, Marchand, and Blanc (2011) were interested in clarifying a construct in their review and appraised the conceptual quality of observational longitudinal studies using two criteria (i.e., analytical breadth and depth).

Conceptual quality is usually mentioned in systematic reviews interested in generating new understanding of a phenomenon. In these reviews, authors argue that too much emphasis on methodological rigor can limit the insight that could be gained from included studies (Sandelowski, 2000). Campbell et al. (2011) observed an inverse correlation between methodological and conceptual quality (i.e., papers providing good conceptual insight are generally of low methodological quality) that they explained, in part, due to the inadequate reporting of qualitative research methods. They suggest limiting methodological quality appraisal to a few screening criteria that allow identifying and excluding fatally flawed papers and focusing on papers that are conceptually useful for the synthesis (Campbell et al., 2011; Dixon-Woods et al., 2006).

The Quality of Research Papers

The second component of the framework is *research papers*, represented by boxes (Figure 2). Research papers can take several forms such as a journal article, dissertation, or report. The quality of how a research paper reports a study (reporting quality) varies widely depending on the authors and the structure of each form (e.g., journal articles are more concise than dissertations). This influences reviewers' judgment of the methodological and conceptual quality of a study.

Reporting quality. Reporting quality is related to the extent to which a paper “provides information about the design, conduct, and analysis of a study” (Huwiler-Müntener, Jüni, Junker, & Egger, 2002, p. 2801). This quality dimension is linked with the constructs of transparency, accuracy, and completeness (Simera et al., 2010). These constructs can be defined as the extent to which a paper provides clear, detailed, and easy to understand information about a study (transparency), provides correct and true information (accuracy), and includes sufficient information (completeness) to allow readers to understand a study (Hornby, 2000; Simera et al., 2010). Over the past decade, more than 90 guidelines have been developed to provide standards for reporting research (Simera et al., 2010). These guidelines focus on issues that might introduce bias into a given study, and thus need to be reported such that readers can judge the quality of that study (Simera et al., 2010). In our review of SMSRs, several reporting guidelines have been used for critical appraisal such as the CONSORT statement for randomized controlled trials (Moher et al., 2010), the STROBE statement for observational studies (von Elm et al., 2007), the COREQ for qualitative studies (Tong, Sainsbury, & Craig, 2007), and TREND statement for non-randomized designs (Armstrong et al., 2008).

Opposing views regarding the use of reporting quality in systematic reviews remain. On the one hand, some are against the use of reporting quality in systematic reviews, especially when used as a surrogate for appraising methodological quality (Higgins & Green, 2008; Wells & Littell, 2009). It was found that using reporting quality as a proxy measure for methodological quality could lead to the misinterpretation of study quality (Huwiler-Müntener et al., 2002). Thus, the results, recommendations and conclusions of a review should be consistent with what was appraised.

On the other hand, reviewers have argued that reporting quality and methodological quality are related since a poorly reported paper will hinder the proper assessment of the

trustworthiness of a study (Carroll, Booth, & Lloyd-Jones, 2012). Reporting quality criteria are said to be easier to judge and less prone to subjectivity (Carroll et al., 2012). Carroll et al. (2012) tested the effect of excluding papers solely based on the adequacy of their reporting using four criteria (pertaining to information provided on the question and study design, selection of participants, methods of data collection, and methods of analysis). They found that excluding inadequately reported papers had no meaningful impact on the results of a review. They suggest appraising reporting quality in a first step to exclude inadequately reported papers and then appraising the methodological quality of the remaining studies.

The Purposes of Critical Appraisal in SMSRs

A third component of the framework is the *review process*, represented by a funnel (Figure 2). In the review process, studies/research papers are identified through databases and other sources, selected using clear eligibility criteria, appraised, and synthesized. In our review of SMSRs, several reasons were provided for performing critical appraisal such as to describe the quality of the papers retained, to exclude papers of low quality or fatally flawed, to do a sensitivity analysis, to guide and strengthen the interpretation of study findings, and to explain differences in study results. These results can be grouped into three main purposes for performing critical appraisal in SMSRs.

The first purpose is for the selection of papers. A threshold approach has been suggested in which only studies meeting a predefined cut-off value of quality are retained (Gough, Thomas, & Oliver, 2012). Other approaches focus on conceptual quality in order to judge the utility, relevance, worth or value of each study, and include only studies meeting minimum criteria of scientific rigor (Pawson, Greenhalgh, Harvey, & Walshe, 2005).

The second purpose is descriptive. That is, the results of the critical appraisal are used to describe the quality of the studies included in a review. This can contribute to understanding how much readers may trust the results, identifying knowledge gaps, and making recommendations for future research (Booth et al., 2012). For this purpose, the appraisal focuses mainly on methodological quality.

The third purpose is related to the synthesis and interpretation of papers. Different strategies have been suggested as alternatives to excluding low methodological quality papers. One strategy is to use a weighting approach in which less weight is given to papers of low

quality during the synthesis and interpretation of results (Gough et al., 2012). Another suggested strategy is to perform a sensitivity analysis based on the results of the critical appraisal. Sensitivity analysis consists of repeating an analysis by removing the studies that failed to achieve a pre-defined quality threshold (Carroll & Booth, 2015). If results differ based on the quality of studies, the conclusions of the review should be nuanced, with more importance placed on the results from higher-quality studies.

In summary, in our analysis of the literature on critical appraisal and SMSRs, we identified three main dimensions of quality (methodological, conceptual, and reporting) that are summarized in Table 2. These dimensions are related to different components: studies and research papers. Although these two components are closely linked, we found it necessary to present them separately since they address different dimensions of quality. That is, methodological and conceptual qualities are associated with studies, whereas, reporting quality is related to research papers. Distinguishing these dimensions in a review process is important since it will influence the review results and recommendations.

Table 2. Three Dimensions of Quality in Critical Appraisal

Features	Quality dimension		
	Methodological	Conceptual	Reporting
Definition	Extent to which a study's design, conduct, and analysis have minimized selection, measurement, and confounding biases	Extent to which a concept is clearly articulated to facilitate theoretical insight	Extent to which a paper provides information about the design, conduct, and analysis of a study
Constructs	Trustworthiness	Insightfulness	Accuracy Completeness Transparency
Component	Study	Study	Research paper
Example of criteria*	Were the statistical tests used to assess the main outcomes appropriate?	Are there clear translatable concepts?	Is the hypothesis/aim/objective of the study clearly described?

* Examples from: Downs and Black (1998) and Toye et al. (2013).

Different Quality Dimensions Used Based on the Objectives of SMSRs

A variety of synthesis methods have been developed for SMSRs to address different objectives (Tricco et al., 2016). We used the components of our framework to understand the differences in critical appraisal approaches used in various SMSRs. The following presents three main review objectives and the dimension of quality addressed for each (Table 3).

One objective of SMSRs can be to test hypotheses by using aggregative synthesis approach such as meta-analysis. The questions may concern, for example, understanding the magnitude of a problem, testing the effectiveness of an intervention, or highlighting the association between factors. One or several predetermined critical appraisal tools are generally used to estimate bias that could lead to drawing misleading conclusions (Gough et al., 2012). Thus, the appraisal will mainly focus on the methodological quality of studies.

A second objective consists of interpreting and arranging the results of studies to generate new ways of understanding a subject and articulate new concepts or theories. To achieve this objective, configurative synthesis (or interpretive) approaches are usually used (Gough et al., 2012). In this type of review, there is no consensus regarding how the critical appraisal should be performed. Appraisal processes range from using the tools employed in the aggregative synthesis approach to focusing on study relevance and contribution to generating new understanding, rather than the methodological quality (Barnett-Page & Thomas, 2009). For example, in critical interpretive synthesis, Dixon-Woods et al. (2006) propose excluding papers that are deemed fatally flawed according to five quality criteria that cover reporting (e.g., are the aims and objectives of the research clearly stated?) and methodological (e.g., is the method of analysis appropriate and adequately explicated?) quality. On the other hand, in meta-narrative synthesis, Greenhalgh et al. (2005) suggest appraising the validity and relevance of primary studies using criteria within their respective research traditions.

A third objective is found in realist synthesis that is interested in understanding narrative causation using middle range theories (Jagosh et al., 2014). This synthesis approach seeks to explore and contextualize a complex intervention in multiple social settings and to answer the following question: “what is it about this kind of intervention that works, for whom, in what circumstances, in what respects and why?” (Pawson et al., 2005). This synthesis approach involves an ongoing iterative interpretive process that uses abductive reasoning; i.e., hunches

about conditions and outcomes can be incorporated in the synthesis (Jagosh et al., 2014). The appraisal in realist synthesis is more interested in the merit of each paper for the purpose of identifying/testing the middle range theory. Papers are appraised based on the minimum criteria of relevance (i.e., whether the study contributes to theory building and/or testing) and rigor (i.e., whether the method used is credible) (Pawson et al., 2005).

Table 3. Comparison of Critical Appraisal Based on the Objectives of Reviews

Objective of reviews	Example of synthesis methods	Type of reasoning	Purpose of the appraisal	Dimension of quality appraised
Test hypothesis	Meta-analysis	Deduction	Determine if studies are affected by significant bias	Methodological
Provide causal pathway or causal explanation	Realist synthesis	Abduction	Determine if studies are fit for purpose for theory development and/or testing	Conceptual and methodological
Develop conceptual understanding	Critical interpretive synthesis Meta-narrative synthesis	Induction	Determine the relevance, credibility, and contribution of studies	No consensus Conceptual and methodological

Discussion

The critical appraisal process in SMSRs is complex due to the heterogeneity of studies designs included. We found the literature on critical appraisal to be disparate, lacking consensus, and subject to multiple debates. There are various definitions of research quality but no agreement regarding what quality is. Also, it is not always clear why critical appraisal is performed, nor is consensus on how to perform it. There exists a wide variety of critical appraisal tools and approaches as well as debate regarding the appropriate expertise required for appraising studies.

To help reviewers deal with this complexity, this paper provides a conceptual framework of critical appraisal in SMSRs in which three components and dimensions of quality are

described (Figure 2). The three dimensions of quality (methodological, conceptual, and reporting) are intertwined in the critical appraisal process. For example, inadequate reporting will preclude a proper appraisal of the methodological and conceptual qualities of studies, selective reporting can be a source of methodological bias, and clear and concise definitions of constructs are fundamental for empirical testing (Higgins & Green, 2008; Simera et al., 2010; Suddaby, 2010). This suggests that multidimensional approaches to critical appraisal could be considered when performing a SMSR. How such approaches can be used needs to be explored further.

The critical appraisal process in SMSRs can be illustrated by the analogy of a courtroom trial where the three components described in Figure 2 can be found: evidence (studies), lawyer (papers), and judge and jury (review). First, in a courtroom trial, evidence will come from various sources such as witnesses and experts. Several questions can be posed: Are the witnesses relevant to the case? Are they credible? Are they making truthful claims? Analogously, questions posed during critical appraisal of studies in a review can be likened to those listed above: Is this study relevant for the review? Are adequate methods used in a rigorous manner? Are the results of the study trustworthy? Second, in a courtroom trial, lawyers are responsible for conveying pertinent evidence of what happened, and convincing the judge and jury of their case. Similarly, studies generally become accessible to the reviewers when they are published. Researchers 'package' their work, communicating it in a way that will convince reviewers it is worthy of being published and also convince others to read and cite it. Third, once the jury and judge have heard all the evidence, they will need to reach a decision. They might have different questions: Which if the diverging accounts presented is true? Should the evidence provided by less credible witnesses be excluded or weighted? Similarly, in a review, when different studies present contradictory results, how can we explain the differences? Which studies are credible and valid? What recommendations should be made based on all the evidence gathered? This analogy illustrates the intermediate position of lawyers (research papers) to convey the evidence (studies) to the judge and jury (review). The way the evidence is 'packaged' can greatly influence the judgment made.

Several future research avenues may be pursued on the critical appraisal process in SMSRs. First, the framework needs to be validated with a group of experts to determine if other dimensions of quality are addressed in SMSRs, and to refine the dimensions. Second, there is a need to explore the interdependencies between the methodological, conceptual, and reporting

dimensions of quality and how they influence the appraisal. Third, there is much debate around appraising methodological quality. It is necessary to test which criteria (for qualitative, quantitative, and mixed methods studies) have significant impact on review recommendations and conclusions. Fourth, an analysis of how the available critical appraisal tools and approaches fit within this framework could be performed. This analysis could lead to proposing a typology of tools and approaches, which will provide guidance for reviewers in selecting the most appropriate one for their reviews. Finally, improving our understanding of how criteria differ among studies from different epistemological and methodological traditions is needed.

Conclusion

The lack of conceptual underpinnings of critical appraisal in SMSRs is a source of multiple debates and inconsistency in the terminology and approaches used. We focused on what critical appraisal is, why the definition of quality varies, and how the findings of critical appraisal can be used. Another important question needs to be addressed: How should the critical appraisal of quantitative, qualitative and mixed methods studies be performed? The proposed framework is an essential first step to help answer this question.

Acknowledgements:

Quan Nha Hong, OT, MSc, PhD candidate, holds a Doctoral Fellowship Award from the Canadian Institutes of Health Research (CIHR). Pierre Pluye, MD, PhD, Full Professor, holds a Senior Investigator Award from the Quebec Health Research Funds (FRQS) and is the Director for Method Development at the Quebec SPOR-SUPPORT Unit, which is funded by the CIHR, the FRQS, and the Quebec Ministry of Health. The authors would like to thank Dr. Suzanne Rivard, Full Professor at HEC Montréal, and Dr. Paula Bush, Academic Associate at McGill University, for constructive feedback on previous versions of this manuscript. We are also grateful to reviewers and editors who provided helpful comments and suggestions on behalf of the journal.

References

- Anderson, L. M., Oliver, S. R., Michie, S., Rehfuss, E., Noyes, J., & Shemilt, I. (2013). Investigating complexity in systematic reviews of interventions by using a spectrum of methods. *Journal of Clinical Epidemiology*, *66*(11), 1223-1229.
- Armstrong, R., Waters, E., Moore, L., Riggs, E., Cuervo, L. G., & Lumbiganon, P. (2008). Improving the reporting of public health intervention research: Advancing TREND and CONSORT. *Journal of Public Health*, *30*(1), 103-109.
- Bai, A., Shukla, V. K., Bak, G., & Wells, G. (2012). *Quality Assessment Tools Project Report*. Ottawa: Canadian Agency for Drugs and Technologies in Health.
- Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: A critical review. *BMC Medical Research Methodology*, *9*(59).
- Beauregard, N., Marchand, A., & Blanc, M.-E. (2011). What do we know about the non-work determinants of workers' mental health? A systematic review of longitudinal studies. *BMC Public Health*, *11*(1).
- Booth, A., Papaioannou, D., & Sutton, A. (2012). *Systematic Approaches to a Successful Literature Review*. London, UK: Sage Publications.
- Bunn, F., Trivedi, D., Alderson, P., Hamilton, L., Martin, A., Pinkney, E., et al. (2015). The impact of Cochrane Reviews: A mixed-methods evaluation of outputs from Cochrane Review Groups supported by the National Institute for Health Research. *Health Technology Assessment*, *19*(28), 1-100.
- Burls, A. (2009). *What is Critical Appraisal?* (2nd Ed.). Newmarket, UK: Hayward Medical Communications.
- Campbell, R., Pound, P., Morgan, M., Daker-White, G., Britten, N., Pill, R., et al. (2011). Evaluating meta-ethnography: Systematic analysis and synthesis of qualitative research. *Health Technology Assessment*, *15*(43), i-164.
- Carroll, C., & Booth, A. (2015). Quality assessment of qualitative evidence for systematic review and synthesis: Is it meaningful, and if so, how should it be performed? *Research Synthesis Methods*, *6*(2), 149-154.

- Carroll, C., Booth, A., & Lloyd-Jones, M. (2012). Should we exclude inadequately reported studies from qualitative systematic reviews? An evaluation of sensitivity analyses in two case study reviews. *Qualitative Health Research*, *C22*(10), 1425-1434.
- Creswell, J. W., & Plano Clark, V. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Critical Appraisal Skills Programme (CASP). (2017). 10 Questions to Help You Make Sense of Qualitative Research. Retrieved November 3, 2017, from http://docs.wixstatic.com/ugd/dded87_25658615020e427da194a325e7773d42.pdf
- Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovich, C., Song, F., et al. (2003). Evaluating non-randomised intervention studies. *Health Technology Assessment*, *7*(27).
- Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., et al. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology*, *6*(35).
- Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *52*(6), 377-384.
- Dyrvig, A.-K., Kidholm, K., Gerke, O., & Vondeling, H. (2014). Checklists for external validity: A systematic review. *Journal of Evaluation in Clinical Practice*, *20*(6), 857-864.
- Fàbregues, S., & Molina-Azorín, J. F. (2017). Addressing quality in mixed methods research: A review and recommendations for a future agenda. *Quality & Quantity*, *51*(6), 2847-2863.
- Fetters, M. D., & Freshwater, D. (2015). The 1+ 1= 3 Integration Challenge. *Journal of Mixed Methods Research*, *9*(2), 115-117.
- Frantzen, K. K., & Fetters, M. D. (2015). Meta-integration for synthesizing data in a systematic mixed studies review: Insights from research on autism spectrum disorder. *Quality & Quantity*, *50*(5), 2251-2277.
- Gough, D. (2015). Qualitative and mixed methods in systematic reviews. *Systematic Reviews*, *4*(181).
- Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews*, *1*(28).

- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science and Medicine*, *61*(2), 417-430.
- Harden, A., & Gough, D. (2012). Quality and relevance appraisal. In D. Gough, S. Oliver & J. Thomas (Eds.), *An Introduction to Systematic Reviews* (pp. 153-178). London, UK: Sage Publications.
- Harden, A., & Thomas, J. (2005). Methodological issues in combining diverse study types in systematic reviews. *International Journal of Social Research Methodology*, *8*(3), 257-271.
- Hawker, S., Payne, S., Kerr, C., Hardey, M., & Powell, J. (2002). Appraising the evidence: Reviewing disparate data systematically. *Qualitative Health Research*, *12*(9), 1284-1299.
- Heyvaert, M., Hannes, K., Maes, B., & Onghena, P. (2013a). Critical appraisal of mixed methods studies. *Journal of Mixed Methods Research*, *7*(4), 302-327.
- Heyvaert, M., Hannes, K., & Onghena, P. (2016). *Using Mixed Methods Research Synthesis for Literature Reviews: The Mixed Methods Research Synthesis Approach* (Vol. 4). Thousand Oaks, CA: Sage Publications.
- Heyvaert, M., Maes, B., & Onghena, P. (2013b). Mixed methods research synthesis: Definition, framework, and potential. *Quality & Quantity*, *47*(2), 659-676.
- Higgins, J. P., & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. 5). Chichester, UK: Wiley Online Library.
- Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*, *343*(d5928).
- Hong, Q. N., Pluye, P., Bujold, M., & Wassef, M. (2017). Convergent and sequential synthesis designs: Implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Systematic Reviews*, *6*(61).
- Hornby, A. S. (2000). *Oxford Advanced Learner's Dictionary of Current English* (6th ed.). Oxford, UK: Oxford University Press.
- Huwiler-Müntener, K., Jüni, P., Junker, C., & Egger, M. (2002). Quality of reporting of randomized trials as a measure of methodologic quality. *Journal of the American Medical Association*, *287*(21), 2801-2804.

- Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods*, 8(4), 49-62.
- Jagosh, J., Pluye, P., Wong, G., Cargo, M., Salsberg, J., Bush, P. L., et al. (2014). Critical reflections on realist review: Insights from customizing the methodology to the needs of participatory research assessment. *Research Synthesis Methods*, 5(2), 131-141.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133.
- Katrak, P., Bialocerkowski, A. E., Massy-Westropp, N., Kumar, S., & Grimmer, K. A. (2004). A systematic review of the content of critical appraisal tools. *BMC Medical Research Methodology*, 4(22).
- Markoulakis, R., & Kirsh, B. (2013). Difficulties for university students with mental health problems: A critical interpretive synthesis. *Review of Higher Education: Journal of the Association for the Study of Higher Education*, 37(1), 77-100.
- Mays, N., Pope, C., & Popay, J. (2005). Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *Journal of Health Services Research and Policy*, 10(Suppl 1), 6 - 20.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., et al. (2010). CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, 63(8), e1-e37.
- Moynihan, R. (2004). *Evaluating Health Services: A Reporter Covers the Science of Research Synthesis*. New York: Milbank Memorial Fund.
- Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review - A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, 10 (Suppl 1), 21-34.
- Pluye, P., & Hong, Q. N. (2014). Combining the power of stories and the power of numbers: Mixed methods research and mixed studies reviews. *Annual Review of Public Health*, 35, 29-45.
- Pluye, P., Hong, Q. N., Bush, P. L., & Vedel, I. (2016). Opening-up the definition of systematic literature review: The plurality of worldviews, methodologies and methods for reviews and syntheses. *Journal of Clinical Epidemiology*, 73(5), 2-5.

- Pluye, P., Robert, E., Cargo, M., Bartlett, G., O’Cathain, A., Griffiths, F., et al. (2011). Proposal: A Mixed Methods Appraisal Tool for systematic mixed studies reviews. Retrieved November 15, 2013, from <http://mixedmethodsappraisaltoolpublic.pbworks.com>
- Roberts, K. A., Dixon-Woods, M., Fitzpatrick, R., Abrams, K. R., & Jones, D. R. (2002). Factors affecting uptake of childhood immunisation: A Bayesian synthesis of qualitative and quantitative evidence. *Lancet*, 360(9345), 1596-1599.
- Sandelowski, M. (2000). Focus on research methods-whatever happened to qualitative description? *Research in Nursing and Health*, 23(4), 334-340.
- Sandelowski, M., Voils, C. I., Leeman, J., & Crandell, J. L. (2012). Mapping the mixed methods–mixed research synthesis terrain. *Journal of Mixed Methods Research*, 6(4), 317-331.
- Simera, I., Moher, D., Hirst, A., Hoey, J., Schulz, K. F., & Altman, D. G. (2010). Transparent and accurate reporting increases reliability, utility, and impact of your research: Reporting guidelines and the EQUATOR Network. *BMC Medicine*, 8(24).
- Suddaby, R. (2010). Editor’s comments: Construct clarity in theories of management and organization. *Academy of Management Review*, 35(3), 346-357.
- Sutton, A. J., & Higgins, J. (2008). Recent developments in meta-analysis. *Statistics in Medicine*, 27(5), 625-650.
- Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004a). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing*, 1(3), 176-184.
- Thomas, J., Harden, A., Oakley, A., Oliver, S., Sutcliffe, K., Rees, R., et al. (2004b). Integrating qualitative research with trials in systematic reviews. *British Medical Journal*, 328(7446), 1010-1012.
- Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, 19(6), 349-357.
- Toye, F., Seers, K., Allcock, N., Briggs, M., Carr, E., Andrews, J., et al. (2013). ‘Trying to pin down jelly’ - Exploring intuitive processes in quality assessment for meta-ethnography. *BMC Medical Research Methodology*, 13(46).

- Tricco, A. C., Antony, J., Soobiah, C., Kastner, M., MacDonald, H., Cogo, E., et al. (2016). Knowledge synthesis methods for integrating qualitative and quantitative data: A scoping review reveals poor operationalization of the methodological steps. *Journal of Clinical Epidemiology*, 73(5), 29-35.
- von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLOS Medicine*, 4(10), e296.
- Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., et al. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Retrieved April 16, 2016, from http://www.ohri.ca/programs/clinical_epidemiology/nosgen.pdf
- Wells, K., & Littell, J. H. (2009). Study quality assessment in systematic reviews of research on intervention effects. *Research on Social Work Practice*, 19(1), 52-62.
- West, S. L., King, V., Carey, T. S., Lohr, K. N., McKoy, N., Sutton, S. F., et al. (2002). *Systems to Rate the Strength of Scientific Evidence*. Rockville, MD: Agency for Healthcare Research and Quality.