University of Cincinnati

Date: 6/23/2023

I, Emily B. Kean, hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Nursing Research.

It is entitled:

Development and Testing of the Reliability and Validity of the IRMAT: Integrative Review Methodology Appraisal Tool

Student's name: **Emily B. Kean**

This work and its defense approved by:

Committee chair: Elaine Miller, Ph.D.

Committee member: Youn Seon Lim, Ph.D.

Committee member: Carolyn Smith, Ph.D.

UNIVERSITY OF Cincinnati

45577

Development and Testing of the Reliability and Validity of the IRMAT: Integrative Review Methodology Appraisal Tool

A dissertation submitted to the Graduate School of the University of Cincinnati

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in the College of Nursing

by

Emily B. Kean

June 23, 2023

Bachelor of Arts in English Literature, Northern Kentucky University

May 2004

Master of Science in Library Science, University of Kentucky

May 2010

Committee Chair: Elaine L. Miller, PhD, RN, CRRN, FAAN, FAHA

Committee Members:

Carolyn R. Smith PhD, RN, CNE

Youn Seon Lim, PhD, MS, MEd, MA

Abstract

Background: Expert literature reviews and evidence synthesis, such as systematic, scoping, or integrative reviews, represent some of the highest levels of evidence in the health sciences. There is a paucity of research concerning the methodology of integrative reviews, which are prevalent in the nursing literature. Because integrative reviews may incorporate the retrieval, appraisal, and synthesis of a variety of evidence types (empirical [quantitative, qualitative, and/or mixed methods], non-empirical, theoretical, and/or methodological), the integrative review is a somewhat unique type of evidence synthesis. While there is a growing body of literature concerning the methodological rigor of integrative reviews in nursing, advancements in methodological tools are lacking.

Purpose: The overarching aim of this dissertation research is to develop and test an appraisal tool to assess the methodological and reporting elements of nursing integrative reviews. The Integrative Review Methodology Appraisal Tool (IRMAT) appraises whether recommended methodological elements are present in published integrative reviews in the nursing literature.

Methods: A scoping review was conducted to identify methodological elements of nursing integrative reviews. The 210 methodological elements identified from the scoping review were coded and thematically analyzed to align to the Ellis Model of Information-Seeking Behavior, which resulted in 34 items being generated. Generated items were analyzed by a panel of five nursing integrative review experts for face and content validity. The subsequent version of the tool was used by 204 survey participants and two independent raters so that construct validity, internal consistency reliability, and inter-rater reliability could be analyzed statistically.

Results: The expert ratings were used to calculate content validity at the item level using the Content Validity Index (CVI). Twenty-eight items demonstrating sufficient content validity (> 0.80) were retained for further testing. A total of 204 respondents used the 28-item IRMAT to appraise a published nursing integrative review. The analysis of this psychometric testing resulted in ten items being removed. The remaining 18 items loaded onto two factors (*Design*

and *Analysis*) and were found to have acceptable internal consistency reliability (Cronbach's α of 0.741 and 0.719, respectively). After exploratory and confirmatory factor analyses, the 18-item tool demonstrated evidence of construct validity. The 18 items loading onto the two factors explained 25.2% of the cumulative variance. Inter-rater reliability (IRR) of the items was slight (0.00-0.20) to moderate (0.41-0.60) for 15 (83%) of the 18 items, with one item demonstrating substantial IRR (0.61-0.80) and two items demonstrating poor IRR (< 0.00). A 16-item version removing the two poor IRR items was analyzed for reliability and model fit; however, the Cronbach's α of factor two was reduced from adequate to questionable, thus the two items were retained in the final 18-item version of the tool.

Conclusion: The Integrative Review Methodology Appraisal Tool (IRMAT) demonstrates evidence of construct validity, internal consistency reliability, and inter-rater reliability. The IRMAT has implications for educating doctoral nursing students and nurses in clinical practice on appraising integrative reviews in the nursing literature. The IRMAT may also be used by editors and peer reviewers to appraise submitted nursing integrative reviews prior to publication.

Copyright

Dedication

For nurses. Throughout my career, I've always had an affinity for working with and for nurses. Thank you to all the nurses who have embraced and encouraged me over the decades. I am proud to contribute to nursing science through this research.

For my committee. Drs. Miller, Smith, and Lim – I am extremely grateful for your time and expertise. I learned the importance of teamwork as part of this process, and you were the best team I could have asked for.

For my PhD Fam. Christine, Jamilah, Kim, Shammah, Teenu, Ndate, Qutaibah, and Jerome. We talked early on about how our friends and family didn't quite understand the PhD journey that we were on. I am eternally grateful to have had my PhD fam who understood and supported me through it all.

Acknowledgements

This research project was supported by the Medical Library Association, Eugene Garfield Research Fellowship 2023 and used software resources provided through the National Center for Advancing Translational Sciences of the National Institutes of Health, Award Number 2UL1TR001425. The content in this project is the authors' responsibility and does not reflect the official views of the Medical Library Association nor the National Institutes of Health.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Expert literature reviews and evidence synthesis, such as systematic, scoping, or integrative reviews, represent some of the highest levels of evidence in the health sciences. Numerous scholars and agencies, such as the Joanna Briggs Institute, have conducted research on systematic and scoping reviews, often resulting in reporting guidelines and quality appraisal tools (Critical Appraisal Skills Programme, 2018; Joanna Briggs Institute, 2020; Moher et al., 2009; Page et al., 2021; Tricco et al., 2018). Currently, there is a paucity of research concerning the methodology of integrative reviews, which are prevalent in nursing and may incorporate the retrieval, appraisal, and synthesis of a variety of evidence types. Critical appraisal of published nursing integrative reviews reveals issues with adherence to methodology, resulting in low quality reviews (Crossetti, 2012; Hopia et al., 2016; Toronto et al., 2018). Several factors illustrate the significance of this lack of research: the Bureau of Labor Statistics predicts a need for more than 200,000 new registered nurses in the United States each year through 2026 (Torpey, 2018 April), and academic accrediting agencies are now emphasizing a need for nurses to pursue advanced practice and doctoral degrees (American Association of Colleges of Nursing, 2019). There is an urgent need to develop tools and resources for nurses authoring, publishing, and appraising integrative reviews, the findings from which are used in education, research, and practice settings.

In a 2019 methodological overview of expert literature review types, Sutton et al. define the integrative review as an “umbrella term for synthesis methods for integrating qualitative and quantitative data” (p. 206). As part of this definition, Sutton and colleagues reference the seminal 2005 article by Whittemore and Knafl which aligns integrative reviews as being situated “primarily within nursing research” (p. 206). As the methodological framework of integrative reviews has progressed over time, the integrative review has evolved so that it may incorporate a combination of empirical (quantitative, qualitative, and/or mixed methods), non-empirical, theoretical, and/or methodological sources (Toronto & Remington, 2020; Whittemore & Knafl,

2005). The inclusion and synthesis of empirical, non-empirical, theoretical, and/or methodological sources situates the integrative review as a somewhat unique type of evidence synthesis methodology in the larger landscape of expert literature reviews and synthesis. The option to include a variety of evidence types as part of the location, appraisal, and synthesis process is significant for a holistic science such as nursing, which routinely encompasses complex topics addressed by various types of evidence sources and research designs.

Background

Although currently identified as primarily a nursing research methodology, integrative reviews were adopted by nursing from the education and social science sectors (Cooper, 1989, 1998; Jackson, 1980). Integrative reviews are not exclusive to nursing, although a search of the general academic database EBSCOhost Academic Search Complete shows that in journals publishing ten or more integrative reviews since 1979, 70.588% (24 of 34) of those journals are classified as being solely about nursing subjects (Table 1.1).

Table 1.1

Academic Search Complete "Integrative Reviews" by Journal and Subject

Journal	Count	Subject
Journal of Clinical Nursing	99	Nursing
Journal of Advanced Nursing	63	Nursing
International Journal of Nursing Studies	49	Nursing
Brazilian Journal of Surgery & Clinical Research	33	Medical Sciences; Surgery
Enfermeria Global	32	Health & Medicine (General)
Western Journal of Nursing Research	30	Nursing
Journal of Nursing Management	28	Nursing
Journal of Advanced Nursing	25	Nursing
Oncology Nursing Forum	24	Nursing
Journal of Nursing Scholarship	22	Nursing
Revista Latino-Americana de Enfermagem (RLAE)	20	Nursing
Scandinavian Journal of Caring Sciences	20	Primary Health Care; Health & Medicine (General)
Cultura de los Cuidados	18	Health & Medicine (General)

Journal	Count	Subject
Nursing Research	18	Nursing
Aquichan	17	Nursing
International Journal of Mental Health Nursing	17	Nursing
International Nursing Review	17	Nursing
Revista de Enfermagem Referencia	17	Nursing
Revista CEFAC	16	Speech-Language Pathology; Audiology
Issues in Mental Health Nursing	15	Nursing
Revista Cientifica de Enfermagem – RECIEN	15	Nursing
Enfermagem Em Foco	14	Nursing
Revista Baiana de Enfermagem	14	Nursing
International Journal of Nursing Practice	13	Nursing
Journal of Psychiatric & Mental Health Nursing	13	Nursing
Palliative Medicine	13	Hospice & Palliative Care
Psychological Bulletin	13	Psychology
Neuroscience & Biobehavioral Reviews	12	Psychology; Neurology
Isokinetics & Exercise Science	11	Kinesiology & Exercise Science
Nursing Ethics	11	Nursing
Journal of the American Association of Nurse Practitioners	10	Nursing
Nursing & Health Sciences	10	Nursing
Nursing Forum	10	Nursing
Patient Education & Counseling	10	Mental Health Services; Health Education

This contextualization aims to not only distinguish integrative reviews from other types of evidence synthesis, but to demonstrate the need for additional methodological study of integrative reviews and their unique contribution in furthering nursing research and science.

Based on the number of published integrative reviews, there is no doubt that this is an established methodology for nursing, with Whittemore and Knafl (2005) being the most frequently cited framework at the time of this writing. Development of appraisal tools and reporting guidelines generally follow the adoption of an expert literature review or evidence synthesis methodology. This is demonstrated in the literature by the evolution over time of the systematic review methodology (Green & Higgins, 2008; Higgins & Cochrane Collaboration, 2019), and the subsequently published appraisal tools (Shea, Bouter, et al., 2007; Shea,

Grimshaw, et al., 2007) and reporting guidelines (Moher et al., 2009; Page et al., 2021). The overall aim of this dissertation is to develop and test the reliability and validity of the Integrative Review Methodology Appraisal Tool (IRMAT), which will allow for future development of integrative review reporting guidelines and further the methodological research for this type of evidence synthesis.

Key Concepts & Definitions

The central concept is "integrative reviews in nursing journals." The primary definition for "integrative review" comes from Whittemore and Knafl (2005): a literature review "… that allows for the inclusion of diverse methodologies (i.e. experimental and non-experimental research) …" (p. 547). This definition will be supplemented through the proposed work of this research, particularly the scoping review (Chapter 2). "Nursing journals" are defined by combining title lists from three reputable sources: (1) the nursing subset journals currently indexed by the National Library of Medicine (2020); (2) the Nursing and Allied Health Resources Section of the Medical Library Association List of Nursing Journals (2016); and (3) the Nursing Journal Directory compiled by the International Academy of Nursing Editors (INANE) and *Nurse Author & Editor* editor by Leslie Nicoll, PhD, MBA, RN, FAAN (2019). Cross referencing these three lists of nursing journals resulted in a list of 481 unique nursing journal titles.

It is necessary to distinguish between the definitions of appraisal tools and reporting guidelines. While there is certainly some potential overlap in educational use purposes, the intention of an appraisal tool is to appraise the methodological quality *and* reporting quality of an integrative review manuscript (Shea, Bouter, et al., 2007). Reporting guidelines are intended to be used in the planning and writing stages of an evidence synthesis manuscript and are defined by Moher et al. (2010) as "a checklist, flow diagram, or explicit text to guide authors in reporting a specific type of research, developed using explicit methodology" (p. 1).

For the purposes of this dissertation, differentiating between the research design approaches of appraisal tool and reporting guideline development is essential. Modeled after

the approaches used to develop the Measurement Tool to Assess Systematic Reviews (AMSTAR) (Shea, Bouter, et al., 2007; Shea, Grimshaw, et al., 2007) and the Mixed Methods Appraisal Tool (MMAT) (Pace et al., 2012; Pluye et al., 2009), appraisal tool design may be completed through factor analysis of methodological elements identified from the literature, using inter-rater agreement as a measure of reliability (DeVellis, 2017). Alternatively, the development of reporting guidelines, such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), has been extensively documented by Moher and colleagues and consists of a much more detailed process often completed with the use of a Delphi technique. The most significant commonality between the definition and development of appraisal tools and reporting guidelines is the identification of methodological elements.

Because reporting guidelines and appraisal tools for other types of expert literature reviews, such as systematic or scoping reviews, are well-defined and established, similar definitions of methodological elements will be used to frame this definition for integrative reviews. Integrative review methodological elements may consist of (but may not be limited to) question formulation, search strategies, quality appraisal of evidence, use of data matrices, use of theory or conceptual frameworks, summation of findings by subgroups, etc.

Problem Statement

Not unlike other expert literature review types, the evidence retrieval, appraisal, and synthesis for an integrative review should be guided by a comprehensive methodology. Although Whittemore and Knafl published an updated methodology for conducting and reporting an integrative review in 2005, it is not uncommon to see published examples of integrative reviews in the nursing literature that vary significantly in terms of adherence to the stated methodology, resulting in a lower level of rigor (Crossetti, 2012; Hopia et al., 2016; Toronto et al., 2018). The issues with rigor in nursing integrative reviews are not new; Whittemore and Knafl's 2005 methodology cites a 1987 article authored by Lawrence Ganong, which represents one of the first critical appraisals of the rigor of integrative reviews in the nursing literature.

While there is a somewhat significant amount of literature concerning the methodological rigor of integrative reviews in nursing, advancements in methodological tools are lacking. As recently as 2014, Whittemore et al. noted that while there are established reporting guidelines and quality appraisal tools for systematic and other types of expert literature reviews, there are no established reporting guidelines nor well-established criteria for quality appraisal of integrative reviews. To continue to advance the methodological rigor and clinical application of nursing integrative reviews, there is a need for validated methodological appraisal tools.

Significance

Development of an appraisal tool for integrative reviews in the nursing literature will advance knowledge in several ways. Because the integrative review is one of the highest levels of evidence and relied on to impact and shape evidence-based decision-making in nursing, it is crucial that these reviews are held to a high standard of methodological rigor. Toronto et al. (2018) discuss the rapid pace of clinical decision-making and place emphasis on the fact that in the healthcare arena, an integrative review is intended to provide a high-level, quickly consumable level of knowledge synthesis that "a nurse clinician or scholar may use … in the development of clinical practice guidelines" (p. 30). Toronto et al. (2018) proceed to detail several areas where the integrative reviews they examined fell short: lack of reported search strategies, inadequate methods for data extraction, and lacking/insufficient quality appraisal. Thus, it appears that if the integrative review was not produced with a strong adherence to a clearly outlined methodology, the information it imparts may not be of the highest quality. This issue is exacerbated by the very nature of the integrative review, which may have study findings of varied evidence types, research designs, and methodological rigor. Through the design and testing of the IRMAT, readers of integrative reviews in the clinical practice setting can be equipped with a checklist-style appraisal tool to quickly evaluate the methodological and reporting quality of nursing integrative reviews. The ability to appraise integrative reviews more

quickly and accurately may lead to an increased use of higher quality evidence, which may in turn lead to higher quality practice and potentially improve clinical outcomes.

An appraisal tool designed to assess the methodological and reporting quality of integrative reviews in the nursing literature will be useful in a broader educational and publishing context, too. The IRMAT will help future student authors and researchers evaluate the quality of published integrative reviews. On the publishing side, editors and peer reviewers could similarly use the IRMAT to appraise the methodological and reporting quality of integrative review submissions before publication. This potential use in a publishing context would likely be superseded by the development of reporting guidelines but would still represent progress towards standardization of the reporting of this type of evidence synthesis.

Specific Aims of the Dissertation

The overarching aim of this dissertation research is to develop and test the internal consistency reliability, inter-rater reliability, and construct validity of an appraisal tool to assess the methodological and reporting elements of nursing integrative reviews. The IRMAT will appraise whether the recommended methodological elements for nursing integrative reviews are present in published integrative reviews in the nursing literature. Due to the lack of any previously validated appraisal tools for integrative reviews, the research for this dissertation began with an exhaustive search of the literature to properly identify methodological elements of integrative reviews. Following this identification process, a thematic analysis of all identified methodological elements was conducted to generate the items for the pilot appraisal tool. Next, the research consisted of engaging a panel of expert participants to test the face and content validity of the pilot tool. The refined version of the pilot tool was then used by a larger sample of participants. The data generated from this larger sample was analyzed for construct validity using factor analysis. Finally, the inter-rater reliability of the tool was tested with a small sample of published nursing integrative reviews. Additions, deletions, and modifications were made to the tool as needed during all phases of statistical testing. The desired outcome is an appraisal

tool demonstrating sufficient reliability and validity to assess methodological and reporting quality of nursing integrative reviews.

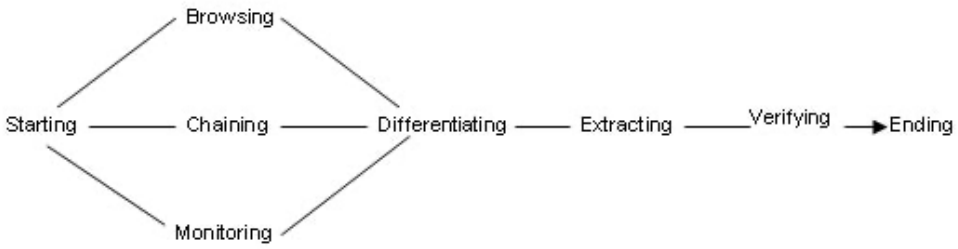
Conceptual Model

While some methodological work has been completed in nursing regarding integrative reviews (Evans, 2007; Russell, 2005; Toronto & Remington, 2020; Whittemore & Knafl, 2005), it is not common to see the formal adoption of a theoretical or conceptual framework to structure this work. Because of the nature of this research as a doctoral dissertation, the use of a conceptual model is paramount. Commonly referenced by the existing nursing literature, the five categories outlined by Cooper (1984, 1989, 1998) – Problem Formulation; Literature Search; Data Evaluation; Data Analysis; Interpretation and Presentation – were used a starting point to choose a theoretical model to serve as a framework for this research. At the essence of Cooper's five categories is the core concept of evidence synthesis and ultimately knowledge generation. Theories and models in the discipline of information science relating specially to knowledge generation were explored as options for the development of the IRMAT.

Ellis's Model of Information-Seeking Behavior was chosen as the conceptual model guiding this research. First developed by David Ellis in 1984 (Ellis, 1984a, 1984b), the Ellis Model has been tested and validated in numerous capacities (Ellis, 1993; Meho & Tibbo, 2003; Thivant, 2005). Although Ellis did not depict his model graphically (Case & Given, 2016), the graphical model developed by Tom Wilson (1999) is supplied as Figure 1.1.

Figure 1.1

Ellis's Model of Information-Seeking Behavior, Adapted by Wilson (1999)



Early versions of Ellis's model consisted of only six categories (Ellis, 2005), but the later version consisting of eight categories will be used for this research (Case & Given, 2016). The eight categories of the Ellis model being used to frame this research are: starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending. The most recent definitions of categories supplied by Case and Givens will be used for this research study:

a. Starting: the initial search for information in which potentially relevant sources are identified.

b. Chaining: following (backwards or forwards) chains of citations or other types of connections among materials.

c. Browsing: semi directed search in areas of potential interest, such as scanning tables of contents, indices, and subject headings.

d. Differentiating: assessing and filtering sources by examining differences in nature and quality.

e. Monitoring: maintaining awareness of developments in a given subject area by regular checking of key sources.

f. Extracting: systematic examination of a particular source to extract material of interest.

g. Verifying: checking that information (e.g., data, citations) is correct.

h. Ending: additional seeking at the end of a project, for example, a final search of literature (Case & Given, 2016, p. 151).

Not only are Case and Given's 2016 definitions of the Ellis model the most recent, these definitions align most appropriately with the five categories outlined by Cooper (1984, 1989; 1998) – Problem Formulation; Literature Search; Data Evaluation; Data Analysis; Interpretation and Presentation – which are pervasively adopted in the nursing literature surrounding integrative review methodology and reporting (Beyea & Nicoll, 1998; Broome, 1993, 2000; Crossetti, 2012; Evans, 2007; Ganong, 1987; Holly, 2019; Hopia et al., 2016; Kirkevold, 1997; Mendes et al., 2008; Russell, 2005; Toronto et al., 2018; Toronto & Remington, 2020; Whittemore, 2005; Whittemore & Knafl, 2005). As noted previously, Ellis's model was interpreted visually in later theoretical research by Wilson (1999). As depicted by Wilson's graphical interpretation (Figure 1.1), where applicable, browsing, chaining, and monitoring will be grouped together as a generalized *Searching* category for the purposes of this research.

The categories depicted by the Ellis Model of Information-Seeking Behavior were used to frame three areas of this research. First, as part of the scoping review conducted for Chapter 2, the integrative review methodological elements identified from the review of the literature were mapped to the starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending categories. Any sources identified as part of the scoping review as methodological in nature were scanned for elements and items that can be mapped to the categories identified by Ellis's model. Second, as part of the assessment of the pilot tool items by the expert participants, Ellis's model was used as constructs for the items generated from the literature. Experts were asked to assess face validity of the appropriateness of the items with the Ellis components, in addition to assessing the face validity of the items themselves. Finally, exploratory and confirmatory factor analyses were completed to test the construct validity of the pilot IRMAT. To provide continuity from the existing literature to the outcomes of the factor analysis, the component outcomes of the factor analysis were mapped to the Ellis model categories for one of the tested models.

Summary of Chapter 1

The goal of this research is to develop and test the IRMAT. As part of this appraisal tool development, the methodological elements recommended for integrative reviews will be defined from existing sources in the nursing literature. Using Ellis's Model of Information-Seeking Behavior will allow for the alignment of reported methodological elements with an established information science theoretical model. The conceptual model will also serve to guide this research and serve as a framework throughout the process. These identified methodological items will be validated by expert participants, refined as needed, and further tested statistically for reliability and validity, ultimately resulting in a tool demonstrating reliability and validity at the conclusion of the research process. The IRMAT will be the first tool of its kind specific to nursing integrative reviews and is intended to be used by clinicians and educators, as well as by editors and publishers, to appraise the methodological and reporting quality of integrative reviews in the nursing literature.

Chapter 2: Literature Review

Criticism of the methodological rigor of integrative reviews in the nursing literature has existed almost as long as the review type itself. Although the methodological aspects of published integrative reviews in nursing have evolved over time, there has never been a truly comprehensive review of how integrative reviews are defined nor which methodological elements have been reported in the literature. Before an appraisal tool for integrative reviews can be developed, an exhaustive search of existing literature must be completed.

In choosing a review methodology for this dissertation research, numerous review types were explored (systematic, integrative, methodological, and scoping) (Sutton et al., 2019). Because there are no existing appraisal tools for integrative reviews, the scoping review was ultimately chosen as the most suitable evidence synthesis methodology. A scoping review will allow for a scan of existing literature and the mapping of retrieved sources into categories. The resulting identified methodological elements of this scoping review will be mapped to the eight categories of Ellis's Model of Information-Seeking Behavior: starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending. Scoping reviews do not require quality appraisal nor formal evidence synthesis (Peters et al., 2020), which would not be appropriate for the questions guiding this research.

Aims

The overarching objective of this scoping review is to provide a list of recommended methodological elements that should be reported in integrative reviews in the nursing literature. There are two questions guiding this review:

1. How are nursing integrative reviews defined?
2. What are the recommended methodological elements that should be reported in an integrative review in the nursing literature?

Methods

This scoping review was conducted according to the methodology outlined by the Joanna Briggs Institute (Peters et al., 2020) and reported according to the PRISMA-ScR reporting guidelines (Tricco et al., 2018). As a preeminent international nursing organization, the methodological framework provided by the Joanna Briggs Institute is appropriate for this research. Similarly, the PRISMA-ScR reporting guidelines developed by Tricco et al. (2018) are commonly referenced in scoping reviews published in the nursing and health sciences literature. Deduplication of sources was conducted in EndNote citation management software (Bramer et al., 2016), and the open-source screening website Rayyan was used for screening (Ouzzani et al., 2016). Data extraction was performed using the matrix method (Garrard, 2017). Data were extracted to Ellis's Model of Information-Seeking Behavior (Case & Given, 2016) using the Best Fit Framework Synthesis (Booth & Carroll, 2015). The Best Fit Framework Synthesis was chosen for its application with the use of *a priori* themes, in this case, the eight Ellis model categories (starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending). Additionally, the Best Fit Framework Synthesis allows for an examination phase of the resulting findings that may not be accommodated by the guiding framework (Booth & Carroll, 2015).

Eligibility criteria

To be included in this scoping review, evidence sources must be describing methodological or reporting considerations for integrative reviews. Articles retrieved during literature searches that were examples of published integrative literature reviews were excluded. For the sources remaining after screening out published integrative review articles, emphasis for inclusion is placed on sources referencing nursing topics; sources exclusively referencing non-nursing disciplines were excluded. Articles broadly referencing health sciences or general literature review concepts were considered but were not ultimately included. No year nor language limits were applied.

Information sources

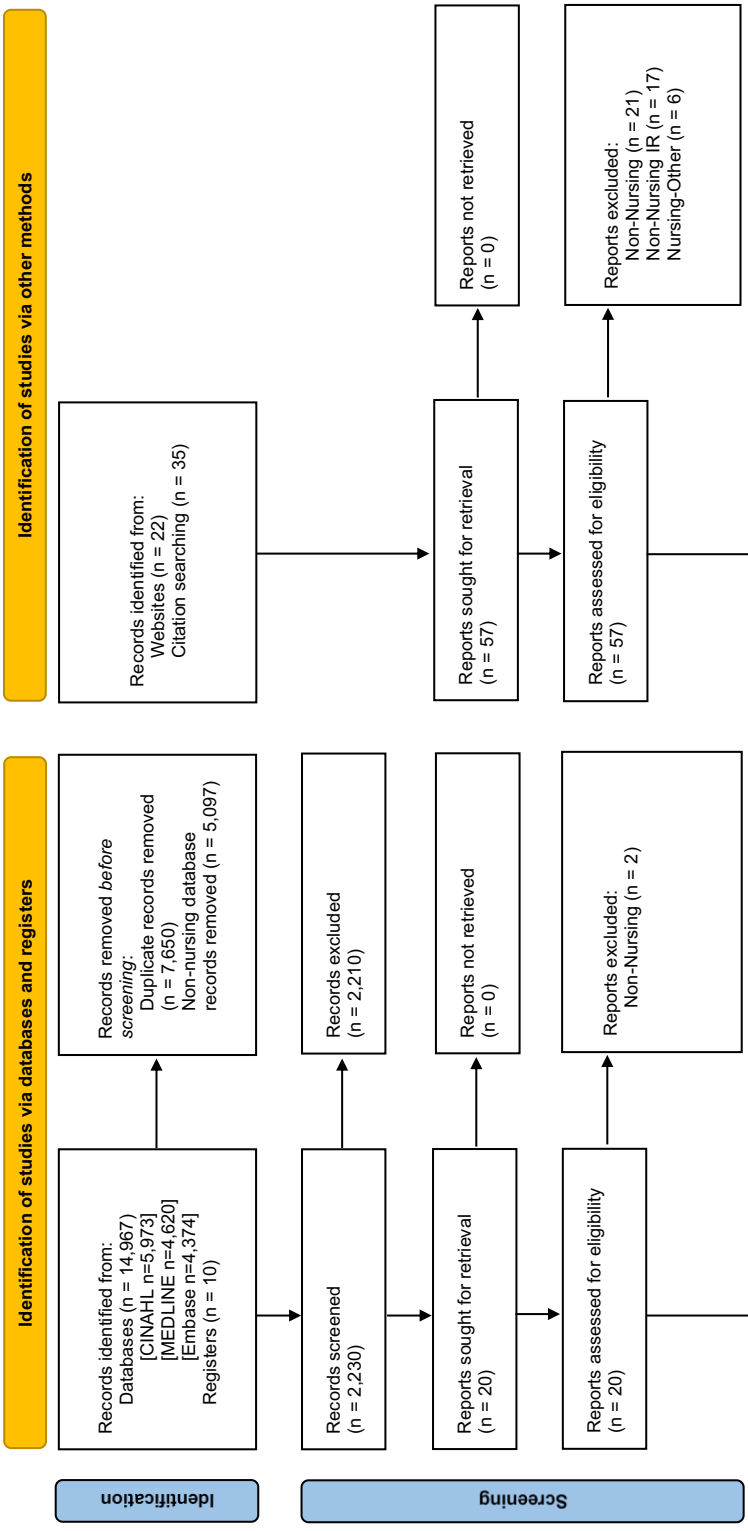
Due to a paucity of methodological experts and writing on this subject area, the types of evidence for this scoping review were not limited to encourage a diverse and representative collection of information. Evidence was primarily sought from peer-reviewed journal articles, book chapters, and books.

Search

Retrieval of sources was conducted using multiple search strategies. Searches for articles were conducted on March 20, 2020 in three databases: Elsevier Embase, EBSCOhost CINAHL Plus with Full Text and MEDLINE with Full Text. Variations of the phrase “integrative review” were used (integrated review, integrative literature review, integrated literature review, etc.). The University of Cincinnati online library catalog and Worldcat.org were searched for book sources. Updates of database and registry searches were conducted from March through July 2021. Google Scholar and Google searches for variations of the phrase “integrative review” were also conducted; results are reported as “Results from Websites” in the PRISMA flowchart (Figure 2.1). According to the eligibility criteria, no year nor language limits were applied to the searches.

Figure 2.1

Modified PRISMA 2020 flow diagram of results



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

Pilot searches conducted as part of the initial exploratory searches revealed several sources which were frequently cited as methodological frameworks for nursing integrative reviews (Broome, 1993; Cooper, 1998; Ganong, 1987; Jackson, 1980; Whittemore, 2005; Whittemore & Knafl, 2005). Forwards and backwards searching of the citations of these identified seminal works were conducted as part of the search process. A search of seminal papers on connectedpapers.com was also completed. Similar to the Google Scholar and Google searches, these results are reported as results from websites in the PRISMA flowchart (Figure 2.1).

Selection of sources of evidence

Sources were eligible for inclusion if they were providing guidance on the completion, writing, or dissemination of integrative reviews in the nursing literature. “Nursing literature” was defined for the purposes of this scoping review by combining journal title lists from three reputable sources: (1) the nursing subset journals currently indexed by the National Library of Medicine (2020); (2) the Nursing and Allied Health Resources Section of the Medical Library Association List of Nursing Journals (2016); and (3) the Nursing Journal Directory compiled by the International Academy of Nursing Editors (INANE) and *Nurse Author & Editor* editor by Leslie Nicoll, PhD, MBA, RN, FAAN (International Academy of Nursing Editors, 2019). Articles critical of the methodological adherence of published integrative reviews were included too. Due to the lack of overall literature on the topic, editorials and commentary sources were included; these sources provided valuable insight into answering the research questions contextually but were not used as sources for extraction of definitions nor methodological elements. Sources that were primary reports of original integrative literature reviews were excluded. Sources that were not directly relevant to nursing (education, broader health sciences, etc.) were retrieved and screened but were ultimately excluded.

Data charting process

Data were extracted and charted specifically to address the two research questions. Integrative review definitions were extracted from all data sources and charted in a data matrix.

Data from evidence sources identified as either methodological or commentary/criticism were extracted and mapped to the eight areas of Ellis's Model of Information-Seeking Behavior: starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending (Case & Given, 2016).

Data items

Definitions with corresponding page numbers and chapter numbers (if applicable), along with selected references pertinent to integrative reviews, were extracted from all data sources. Several sources contained multiple definitions throughout and were extracted accordingly. Data elements primarily consisted of methodological elements suggested for inclusion in the conduct and reporting of integrative reviews in nursing. Items that were not mappable to Ellis's Model were documented but not included. Although categories were decided *a priori* according to the Best Fit Framework Synthesis (Booth & Carroll, 2015), there were instances of identified elements not fitting into the predetermined categorization of the Ellis model. These elements fell into three categories: general methodological recommendations not aligning directly with the eight areas of Ellis's model; recommendations for work to be conducted prior to the *Starting* category; and recommendations specific to dissemination but not applicable to the *Ending* category.

Critical appraisal of individual sources of evidence

Based on the scoping review methodology, formal critical appraisal was not completed; however, due to the inclusion of diverse sources, sources such as editorials and commentary were weighted less prominently in the reported outcomes. As reported below in the *Mapped results* section, results deemed to be methodological in nature were given greater preference when documenting evidence for research question two on the methodological elements to be included in integrative reviews in the nursing literature. Resources deemed to be of historical importance or of a more educational nature were used to answer question one (definition) but not two (methodological elements).

Mapping of results

Retrieved sources were appraised and categorized according to source type: historical, educational, commentary/criticism/editorial, and methodological. Definitions and differentiations of review types were extracted from all sources. Results of the scoping review which were categorized as methodological sources were organized by the eight areas of the Ellis Model resulting in a comprehensive list of recommended methodological elements for inclusion in the conduct and reporting of integrative reviews in nursing.

Results

The initial database and registry searches retrieved 14,977 results. After deduplication and filtering to nursing-specific resources, 2,230 results remained. Title and abstract screening resulted in 20 items assessed for eligibility from the database and registry searches. Web searches and forward and backward citation searching of retrieved sources resulted in an additional 57 sources to be screened.

Evidence sources

After final screening of the 77 retrieved sources, 31 resources were included in the review. From these 31 sources, 39 results made up the final sample. A modified PRISMA flowchart of results is available as Figure 2.1 (Page et al., 2021). Primary reasons for exclusion included sources which were not unique to nursing, did not sufficiently address integrative reviews, or were deemed to be historical to the nursing literature, but not significant to the specific objectives of this research.

Characteristics of sources of evidence

The 31 sources included in the review varied in characteristics. The sources have publication dates that span from 1987 to 2020. Twelve of the 31 sources are books, with seven of these books being updated editions by the same primary authors (Holly, 2014, 2019; LoBiondo-Wood & Haber, 2010c, 2014b, 2018b; Rodgers & Knafl, 1993, 2000). Eight results were categorized as being historically significant (Beyea & Nicoll, 1998; Broome, 1993, 2000;

Ganong, 1987; Kirkevold, 1997; Knafl & Deatrick, 2000; Roman & Friedlander, 1998; Stevens, 2001). Thirteen results were identified as being educational in nature (Conner, 2014; Coughlan et al., 2013; LoBiondo-Wood, 2014; Lobiondo-Wood, 2018; LoBiondo-Wood & Haber, 2010a, 2010b, 2010c, 2014a, 2014b, 2018a, 2018b; Soares et al., 2014; Whittemore et al., 2014). Five articles were criticism, commentary, or editorial about the methodological rigor of integrative reviews in the nursing literature (Crossetti, 2012; Flanagan, 2018; Hopia et al., 2016; Knafl & Whittemore, 2017; Toronto et al., 2018). The remaining 13 results contained methodological elements and guidance and were thus categorized as methodological sources (Brown, 2012; Christmals & Gross, 2017; de Souza et al., 2010; Evans, 2007; Holly, 2017, 2019; Mendes et al., 2008; Russell, 2005; Shuler, 2014; Toronto & Remington, 2020; Whittemore, 2005, 2007; Whittemore & Knafl, 2005). Additional details about the categorization as well as the results are provided below.

Critical appraisal within sources of evidence

While formal evidence appraisal is not required as part of a scoping review (nor is evidence appraisal entirely applicable to this subject matter), attention was paid to the type of evidence under inclusion. For example, peer-reviewed evidence sources and seminal nursing works were given greater credence than editorials and opinion statements.

Results of individual sources of evidence

The 39 results from the 31 sources were categorized as historical, educational, commentary/criticism/editorial, and methodological. Rationale for these categorizations and the results of the individual sources of evidence are discussed below.

Historical

There were several resources which defined and discussed integrative reviews in a historical context that would be more closely associated with the present-day definitions of systematic reviews and/or meta-analysis (Beyea & Nicoll, 1998; Broome, 1993, 2000; Ganong, 1987; Kirkevold, 1997; Knafl & Deatrick, 2000). Stevens (2001) makes a relevant distinction

between the scope of systematic and integrative reviews by providing a definition that differentiates integrative reviews as a less rigorous methodology.

Educational

Educational sources were generally broad in nature. Primarily from book sources, these 13 results provided broad overviews of what integrative reviews were but did not provide detailed guidance on conducting integrative reviews specifically (Conner, 2014; Coughlan et al., 2013; LoBiondo-Wood, 2014; Lobiondo-Wood, 2018; LoBiondo-Wood & Haber, 2010a, 2010b, 2010c, 2014a, 2014b, 2018a, 2018b; Soares et al., 2014; Whittemore et al., 2014). Conner (2014) and Whittemore et al. (2014) supply categorization and definitions in an educational context to assist the reader with choosing what type of review to conduct. The 2014 article by Soares et al. is a unique contribution to the educational research about integrative reviews.

Commentary/Criticism/Editorial

Five articles were identified as commentary or criticism of the methodological rigor of integrative reviews in the nursing literature (Crossetti, 2012; Flanagan, 2018; Hopia et al., 2016; Knafl & Whittemore, 2017; Toronto et al., 2018). Although two of these were editorials, there was relevant and unique information contained therein (Crossetti, 2012; Flanagan, 2018). Data were extracted from all five sources, with less credence being given to the two editorials. Of note is the recent publication dates for these five sources. While Ganong's seminal 1987 article could be considered criticism of the rigor of integrative reviews in the nursing literature, it is categorized as historical for the purposes of this research.

Methodological

The first appearances of discussion of methodology of integrative reviews in the nursing literature begin to appear in 2005 (Russell, 2005; Whittemore, 2005; Whittemore & Knafl, 2005). The most frequently cited and seminal of these early methodological works is the 2005 article by Whittemore and Knafl. Since 2007, additional sources advocating for enhanced methodological rigor have been published as well (Evans, 2007; Whittemore, 2007). Adding to the international

scope of nursing integrative reviews, several frequently cited sources have come out of Brazil (de Souza et al., 2010; Mendes et al., 2008). Numerous book sources were categorized as part of this review as methodological sources, although the quality was found to be not as high when compared side-by-side with the other methodological sources (Brown, 2012; Holly, 2017, 2019; Shuler, 2014). A specific use case of integrative reviews for postgraduate research was identified (Christmals & Gross, 2017). The specificity of this source led to it not being extracted. Finally, the text by Toronto and Remington was identified as one of the more recent and comprehensive methodological sources (2020).

Mapped results

As traditional synthesis is not recommended for a scoping review (Peters et al., 2020), the extracted data was mapped to answer the two questions guiding the review:

1. How are nursing integrative reviews defined?
2. What are the recommended methodological elements that should be reported in an integrative review in the nursing literature?

The results for each question are provided below.

Defining Integrative Reviews

Definitions of integrative reviews were extracted from all sources and are compiled in Appendix A. As denoted in the table provided as Appendix A, it was not uncommon for a single source to contain multiple varying but similar definitions. For example, book sources often contained definitions in multiple chapters as well as the book's supplemental glossary. Historically, integrative reviews were defined very generally and were often referred to as the "broadest" type of review (Coughlan et al., 2013; Evans, 2007; Flanagan, 2018; Lobiondo-Wood, 2018; LoBiondo-Wood & Haber, 2010b; Whittemore, 2005, 2007). As time has passed and greater attention has been paid to the rigor of integrative reviews, definitions have expanded to focus on the necessity for critical appraisal of sources included in the evidence

synthesis (Holly, 2017, 2019; Lobiondo-Wood, 2018; Shuler, 2014; Toronto & Remington, 2020). Definitions in recent years have evolved to highlight the diversity of the sampling frame of the integrative review. Where integrative reviews were sometimes defined as allowing for the combination of qualitative and quantitative sources, definitions of integrative review sampling frames now more consistently refer to the fact that they may encompass empirical (quantitative, qualitative, and/or mixed methods); non-empirical; theoretical, and/or methodological sources (Holly, 2017; Hopia et al., 2016; Knafl & Whittemore, 2017; LoBiondo-Wood, 2014; Lobiondo-Wood, 2018; LoBiondo-Wood & Haber, 2010b, 2014b; Mendes et al., 2008; Shuler, 2014; Toronto & Remington, 2020; Whittemore, 2005, 2007; Whittemore et al., 2014; Whittemore & Knafl, 2005). Definitions acknowledge the complexity inherent in the combination of these varied sources as well (Evans, 2007; Whittemore, 2005).

Extracted methodological elements for nursing integrative reviews

The extracted methodological elements for nursing integrative reviews identified as part of this review are listed in Appendix B. As part of the extraction, 210 elements were identified from the 13 sources describing methodological or reporting considerations for integrative reviews in nursing (Brown, 2012; Christmals & Gross, 2017; de Souza et al., 2010; Evans, 2007; Holly, 2017, 2019; Mendes et al., 2008; Russell, 2005; Shuler, 2014; Toronto & Remington, 2020; Whittemore, 2005, 2007; Whittemore & Knafl, 2005) and the five sources identifed as commentary/criticism (Crossetti, 2012; Flanagan, 2018; Hopia et al., 2016; Knafl & Whittemore, 2017; Toronto et al., 2018). During the extraction process, each methodological element was mapped to the corresponding Ellis model category of starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending. As depicted in the visual rendering of the Ellis model adapted by Wilson (see Chapter 1, Figure 1.1), where descriptions of search elements bridging the browsing, chaining, and monitoring categories were found, these were categorized as “browsing-chaining-monitoring.” Stand-alone examples of the browsing, chaining, or monitoring categories are delineated as such. Amongst the 210 identified

methodological elements, there is a somewhat even distribution of items amongst the Ellis categories: starting (n=22); browsing-chaining-monitoring (n=26); browsing (n=14); chaining (n=6); monitoring (n=7); differentiating (n=53); extracting (n=32); verifying (n=38); and ending (n=12). Within the methodological sources, it was commonly found that the methodological or reporting elements were presented in order within the source documents. For example, considerations of starting or browsing-chaining-monitoring (the grouped *Searching* functions) were often described earlier in the methodological sources than descriptions elements identified as belonging to the extracting, verifying, or ending categories.

Discussion

This scoping review reveals several strengths and gaps, as well as limitations, of the mapped evidence. The results of this scoping review overwhelmingly support the assumption that integrative reviews are a nursing research phenomenon. The resulting amount of evidence sources demonstrate that the discussion of integrative reviews is entrenched in the nursing literature. However, the results show that (except for the recent publication of the Toronto and Remington text) there is not much recent advancement in the methodological writings concerning integrative reviews in nursing. This is especially apparent when comparing these results to the broader literature and methodological writings/advancements for other review types, namely systematic and scoping reviews.

Overall, the specific objectives of this scoping review were met. The extracted definitions from the numerous identified sources clearly demonstrate not only a progression but a consistency over time of how integrative reviews are defined in a nursing context. The identification of recommended methodological elements for reporting were identified as part of this scoping review as well. The mapping of elements to the Ellis model of Information-Seeking Behavior provided particularly useful in organizing elements by the general categories of starting, chaining, browsing, differentiating, monitoring, extracting, verifying, and ending.

The use of the Best Fit Framework Synthesis and the *a priori* elements of the Ellis model did allow for an interesting trend to emerge. In mapping the identified elements, it became clear that there was significant attention paid to aspects of dissemination in the identified sources. Where Ellis refers to *Ending* as the final form of knowledge generation in the Information-Seeking Model, the nursing integrative review sources routinely discussed more tactical issues of dissemination: what elements should be reported and how as opposed to techniques and strategies for evidence synthesis leading to new generation of knowledge. Other types of expert literature reviews and evidence synthesis, such as systematic or scoping reviews, have the benefit of reporting guidelines to address issues of dissemination. This scoping review revealed that methodological sources on nursing integrative reviews are often incorporating recommendations for dissemination as part of their methodological writings.

Limitations

While this review is comprehensive in the breadth of approach in identifying and compiling methodological elements of integrative reviews in the nursing literature, it is not without its limitations. The search results were not limited to English; however, Google Translate was used for translation as opposed to formal translation services. While Google Translate has been found to be sufficient for the purposes of data extraction (Balk et al., 2013), formal translation services would have strengthened the accuracy of the data extracted from the non-English sources. The review is most limited by only having one reviewer serve as screener; although this was mitigated to an extent by having a second individual verify findings. As with any review of this type, there is always a potential limit of not having a truly exhaustive search and not having located all relevant sources.

Summary of Chapter 2

This scoping review represents one of the most comprehensive looks at the definitions and methodological elements of integrative reviews in the nursing literature completed to date. The mapped results of this scoping review have implications for future research. Generally, this

review demonstrates a need for additional methodological experts and advancements for integrative reviews in the nursing literature. Specifically, as it relates to this doctoral work, the methodological elements identified by this scoping review may be used to test methodological reporting of existing integrative reviews in the nursing literature. The elements identified by this review may also be used for future research involving the development of reporting guidelines and/or appraisal tools for nursing integrative reviews.

Chapter 3: Research Methods

The overarching objective of this research was to design and test the reliability and validity of an appraisal tool to assess methodological elements reported in integrative reviews in the nursing literature. Methodological elements were mapped to the Ellis Model of Information-Seeking Behavior at multiple points throughout this study. The methodological elements identified from the scoping review (Chapter 2) were used to develop a pilot version of the tool. Based on the mapping to the Ellis model completed as part of the scoping review, it was anticipated that the initial pilot version of the tool would consist of approximately 30-60 items. The pilot version of the tool was then reviewed by a panel of experts to test face and content validity. If applicable, the items were refined based on the feedback from the expert panel. After testing content validity, the refined tool was then used by a minimum of 180 nursing integrative review authors serving as reviewers to appraise one identical published nursing integrative review article. Exploratory and confirmatory factor analyses were then conducted on the data derived from the reviewers' sample to map the elements into the pre-determined components of the Ellis model with the goal of further reducing the number of items in the appraisal tool and demonstrating construct validity. To test inter-rater reliability, the refined tool was then used by two independent PhD nursing student raters to assess a sample of 50 nursing integrative reviews. Items not demonstrating sufficient inter-rater agreement were candidates for modification or removal.

Specific Aims/Research Questions

The overarching research question guiding this project is: Which methodological elements should be included in an integrative review appraisal tool? This question was addressed with three specific aims:

Specific Aim 1

Specific aim 1 is to confirm face and content validity of the IRMAT and is guided by the following research questions:

- According to a thematic analysis of the nursing literature, what are the recommended methodological elements that should be reported in an integrative review in the nursing literature?
- According to nursing integrative review experts, does the IRMAT demonstrate evidence of face and content validity?

To achieve aim 1, five nursing integrative review experts were recruited from a sample of authors who have written about methodological aspects of integrative reviews in the nursing literature and/or have edited, reviewed, or published a nursing integrative review in the past two years. Content validity was examined using the content validity index (CVI) scores at the item-level (Lynn, 1986; Polit & Beck, 2006). As part of this process, items could be removed, modified, or added based on experts' feedback.

Specific Aim 2

Specific aim 2 serves to demonstrate internal consistency reliability and construct validity of the IRMAT and is guided by the following research questions:

- Using exploratory factor analysis, does the IRMAT demonstrate evidence of internal consistency reliability and construct validity?
- Using the Ellis Model of Information-Seeking Behavior (2005) as a framework for confirmatory factor analysis, does the IRMAT demonstrate evidence of construct validity and good model fit?

Models derived from exploratory factor analysis and confirmatory factor analysis (using the areas of Ellis's Model of Information-Seeking Behavior) were compared for model fit. Internal consistency reliability was calculated using Cronbach's α.

Specific Aim 3

Specific aim 3 consists of determining the inter-rater reliability of the tool and is guided by the following research question:

- Based on a sample appraisal of two independent raters, does the IRMAT demonstrate satisfactory inter-rater reliability at the item level?

The pilot tool was used by two independent raters on a sample of 50 nursing integrative reviews. At the completion of this screening process, the resulting dataset was statistically analyzed for inter-rater reliability at the item level using Cohen's k.

Methods

It is important to differentiate not only between appraisal tools and reporting guidelines, but also to note the differences in how appraisal tools are developed compared to traditional instrument or scale development. Moher et al. (2010) have clearly established criteria for developing health research reporting guidelines. The guidance provided by this 2010 Moher et al. article has been used in numerous examples, most notably the development of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses, more commonly known as PRISMA (Moher et al., 2009; Tricco et al., 2018). However, as noted by Whittemore et al. (2014), integrative reviews currently have neither an established appraisal tool nor reporting guidelines. The process recommended for reporting guideline development by Moher et al. (2010) uses a Delphi method. The commitment of resources needed for a Delphi method exercise is outside the scope of expertise for a doctoral student researcher, hence why this research is pursuing the development of an appraisal tool. As noted by DeVellis (2017), for research endeavors such as appraisal tool development where raters or judges serve as indicators, Cohen's k coefficient is an appropriate statistical measurement of inter-rater agreement, in addition to traditional use of Cronbach's α as a measurement of internal consistency reliability.

Several examples of appraisal tool development in the health sciences literature were consulted as possible methodological models for this proposed research (Crowe & Sheppard, 2011; Pace et al., 2012; Pluye et al., 2009; Shea, Grimshaw, et al., 2007). Ultimately, the processes described in developing the mixed methods appraisal tool (MMAT) (Hong, Gonzalez-

Reyes, et al., 2018; Hong, Pluye, et al., 2018; Pace et al., 2012; Pluye et al., 2009) and the measurement tool for the assessment of multiple systematic reviews (AMSTAR) (Shea, Bouter, et al., 2007; Shea, Grimshaw, et al., 2007) were determined to be the most statistically rigorous and are used throughout to influence the methods for this dissertation research.

Tool Development

The initial list of items for the tool were generated from the outcome of the methodological elements identified from the scoping review (Appendix B). As part of the scoping review completed for this dissertation (Chapter 2), all sources identified as methodological in nature (Brown, 2012; Christmals & Gross, 2017; de Souza et al., 2010; Evans, 2007; Holly, 2017, 2019; Mendes et al., 2008; Russell, 2005; Shuler, 2014; Toronto & Remington, 2020; Whittemore, 2005, 2007; Whittemore & Knafl, 2005) were scanned for items that were considered to be identifiable methodological or reporting items. These items were then mapped to the Ellis Model of Information-Seeking Behavior as part of the scoping review process. To generate the actual items for the pilot version of the appraisal tool, additional analysis was required and is described below.

Item Generation

To generate the final list of items for the pilot appraisal tool, a thematic grouping analysis was conducted using the extracted elements from the methodological sources of the scoping review (Chapter 2). Data condensation was achieved by printing the identified elements on slips and paper and grouping the elements thematically (Miles et al., 2014). After data condensation, Nvivo (Release 1.6.2) was used for further organization and visualization to arrive at the themes and subthemes. The previously determined mapping to the Ellis Model of Information-Seeking Behavior was used as a framework for the categories of the generated items. Where possible, even distribution of items amongst the areas of the Ellis Model of Information-Seeking Behavior was sought. Due to the prominence of the framework provided by Whittemore and Knafl (2005) and the currency of the Toronto and Remington (2020) text, it was anticipated that a majority of

elements would be derived from these two sources, which proved true. Supplemental elements from the remaining eleven sources were also extracted and represented as items where unique themes emerged from these sources. In keeping with the Best Fit Framework Synthesis, items not generated from the thematic analysis may be also adapted from known evidence synthesis reporting guidelines and appraisal tools, such as the PRISMA and PRISMA-ScR reporting guidelines for systematic and scoping reviews (Moher et al., 2009; Page et al., 2021; Tricco et al., 2018). Following the model of other appraisal tool development (Hong, Pluye, et al., 2018; Pace et al., 2012; Pluye et al., 2009), a supplemental tutorial document explaining the objective of each item was generated to accompany the pilot appraisal tool. This thematic analysis to generate items was conducted primarily by the doctoral student researcher, with the secondary researchers serving to validate the findings of the thematic grouping analysis.

Face and Content Validity

Face and content validity was assessed by an expert panel. Integrative review experts are defined as having written about methodological aspects of integrative reviews in the nursing literature. The authors of the sources identified as having been of a methodological nature for the scoping review conducted for this study were approached via email for participation (Brown, 2012; Christmals & Gross, 2017; de Souza et al., 2010; Evans, 2007; Holly, 2017, 2019; Mendes et al., 2008; Russell, 2005; Shuler, 2014; Toronto & Remington, 2020; Whittemore, 2005, 2007; Whittemore & Knafl, 2005). Editors and peer-reviewers of journals that have been identified as having edited, reviewed, or published an integrative review in the past two years for a nursing journal publishing a high frequency of integrative reviews (Table 3.1) were also considered experts. Non-United States participants (particularly those in Brazil and Australia) were included. Ideally, the expert panel would consist of a maximum of 10 methodological experts and 10 editorial experts. Human subjects protection and institutional review board approval are discussed in detail in a subsequent section.

Table 3.1

Top 25 Nursing Journals Publishing Integrative Reviews by Frequency

Name	Frequency Count
Revista Brasileira De Enfermagem	426
Revista Gaucha De Enfermagem	190
Nurse Education Today	168
Journal of Clinical Nursing (John Wiley & Sons, Inc.)	134
International Journal of Nursing Studies	122
Journal of Pediatric Nursing	96
Journal of Advanced Nursing (Wiley-Blackwell)	95
Journal of Advanced Nursing (John Wiley & Sons, Inc.)	78
Western Journal of Nursing Research	78
Oncology Nursing Forum	77
Revista Da Escola De Enfermagem Da Usp	70
Journal of Nursing Scholarship	64
Revista Latino-Americana De Enfermagem (Rlae)	60
Scandinavian Journal of Caring Sciences	55
Issues in Mental Health Nursing	54
Nurse Education in Practice	50
International Nursing Review	46
International Journal of Mental Health Nursing	43
Journal of Gerontological Nursing	42
Journal of Nursing Education	42
Midwifery	42
AORN Journal	40
Cancer Nursing	39
Nursing Research	39
Journal of School Nursing	35
Nursing Forum	35

Expert participants were asked to rate the applicability of each item within the identified Ellis model categories using a 4-point Likert scale (Lynn, 1986; Polit & Beck, 2006). Data were collected via email survey using the REDCap (Research Electronic Data Capture) electronic data capture tools hosted by the Center for Clinical and Translational Science and Training at the University of Cincinnati and Cincinnati Children's (Harris et al., 2019; Harris et al., 2009). Items were assessed as 1 = not relevant; 2 = somewhat relevant; 3 = quite relevant; and 4 =

highly relevant (Polit & Beck, 2006). Any 1 (not relevant) or 2 (somewhat relevant) ratings received open-text follow-up prompts in REDCap asking the experts to provide feedback on ways to improve, re-word, or modify the items. These follow-up prompts also asked the experts if the items were recommended for deletion. The data generated from the Likert ratings of each item were analyzed to generate content validity index (CVI) at the item-level (I-CVIs) (Polit & Beck, 2006). Items with I-CVIs greater than 0.780 were included in the revised version of the IRMAT (Lynn, 1986; Polit & Beck, 2006). Open-text feedback received for low-rated items was examined using content analysis to improve, modify, or delete items based on expert feedback.

Pilot Tool Validation

After determining the content validity and reliability of the pilot tool, the subsequently refined version of the IRMAT was assessed for construct validity.

Data Collection

Eligible nursing integrative review author participants received the refined pilot version of the IRMAT and the accompanying tutorial document via an emailed REDCap survey link. Screening was conducted using the REDCap (Research Electronic Data Capture) electronic data capture tools hosted by the Center for Clinical and Translational Science and Training at the University of Cincinnati and Cincinnati Children's (Harris et al., 2019; Harris et al., 2009). Each participant was provided with the full text of one identical sample article as part of the REDCap survey tool. All items from the pilot IRMAT were used to appraise the sample article by each participant scoring a "4" for "Yes," meaning the methodological item is present; "3" for "Partial Yes," meaning the item is only partially present; "2" for "Partial No," meaning the rater cannot determine whether or not the item is present; and "1" for "No," meaning the item is not present (Shea et al., 2017). Demographic data were also collected from participants. Optional questions pertaining to the participants' age, gender, education level, and primary employment responsibility (clinical, education, research, etc.) were asked as part at the conclusion of the IRMAT survey. The data collection process was estimated to take each participant 20-30

minutes to complete the appraisal. Human subjects protection and institutional review board approval are discussed in detail in a subsequent section.

Participants

The target sample size for recruitment was 180 participants, which according to Fabrigar and Wegener (2012) should be a sufficient minimal sample size allowing for communalities ranging from 0.40 to 0.70 with at least three strong loadings (as cited in Pituch & Stevens, 2016, p. 347).

Sampling Plan

Purposive sampling targeted individuals who have authored an integrative review published in a nursing journal; editors of nursing journals; and health sciences nursing librarians, if needed. A search of the Elsevier Scopus database for integrative reviews published in the subset of 481 nursing journals in the past 20 years was conducted. From the Scopus search, the email addresses of over 3,500 nursing integrative review authors meeting the inclusion criteria for this aim were extracted. Five percent of this pool of potential participants would need to be recruited for this study, which is feasible based on reported rates of nurse survey participation (VanGeest & Johnson, 2011). Targeted email distribution to these 3,500 authors was the first phase of purposive sampling.

Snowball Sampling

In addition to purposive sampling, recruitment may have also occurred with snowball sampling. However, because of successful response rates of the nursing integrative review author participants, it was not necessary to employ the second (targeted email distribution to the approximately 100 editors of nursing journals identified by the International Academy of Nursing Editors) nor third (health sciences nursing librarians via the Nursing and Allied Health Resources Section of the Medical Library Association email distribution list) approaches to participant recruitment via snowball sampling.

Data Analysis

Descriptive statistics including mean, range, and standard deviation (where appropriate) were determined. The cut-off for statistical significance is a p-value less than 0.05. P-values ≥ 0.001 will be reported to 3 decimal places; p-values less than 0.001 will be reported as "< 0.001". Descriptive statistics for the demographic data collected from the participant sample were analyzed statistically. Where necessary, responses with missing values were excluded from analysis. Data analysis was conducted in R (version 4.2.2) and RStudio (version 2021.09.0, build 351) using the following packages: psych (Revelle, 2023), nFactors (Raiche, 2022), performance (Lüdecke et al., 2021), ltm (Rizopoulos, 2006), and lavaan (Rosseel, 2011). All data analysis was conducted independently by the primary researcher (Kean) and was verified by co-researchers (Lim, Miller, and Smith). The R syntax is included as Appendix C.

Exploratory factor analysis was first conducted using the psych package in R (Revelle, 2023) to test which items load onto which components (Pituch & Stevens, 2016). Covariance and correlation matrices were examined. Principal axis factoring using promax rotation was used to condense the correlation matrix, as well as to establish which items load onto which components (Kaiser, 1960; Pituch & Stevens, 2016; Shea, Grimshaw, et al., 2007). A heat map of the correlation matrix was produced. Floor and ceiling effects were assessed for all items. Internal consistency reliability was examined using Cronbach's α. Items loading highly onto one component (>0.30) were retained as a means of reducing the number of items in the appraisal tool (Shea, Grimshaw, et al., 2007). The exploratory factor analysis included computation for Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and the Bartlett's test of sphericity (Field, 2018). KMO values are reported according to the following scale: 0.50-0.70 – mediocre; 0.70-0.80 – good; 0.80-0.90 – great; and 0.90 and above – superb (Kaiser, 1970).

Confirmatory factor analysis (CFA) was then conducted in R using the lavaan package (Rosseel, 2012). Ellis's Model of Information-Seeking Behavior was used as a guiding

framework to identify, map, and name the components derived from the factor analysis. The pilot tool was tested with several models to test overall goodness of fit:

1. Single Factor Model consisting of all items
2. Multiple Factor Model consisting of all items
3. Adjusted Multiple Factor Models consisting of pruned item models from exploratory factor analysis outcome.

Robust maximum likelihood estimation with a promax rotation was used for all models (Maydeu-Olivares, 2017). The following goodness of fit indices at the global level were examined using: model chi-square (χ^2), goodness of fit (GFI), adjusted goodness of fit (AGFI), comparative fit index (CFI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA) (Kline, 2016). Local fit was examined at the item level standardized loadings, as well as the residual matrices. Factor scores were estimated, and reliability of factor scores was computed. Items with low covariances or residual value greater than 0.10 were candidates for removal. Goodness of fit indices are reported with the following cut-offs: χ^2 p-value > 0.05; GFI ≥ 0.95; AGFI ≥ 0.90; CFI ≥ 0.90; SRMR < 0.08; and RMSEA < 0.08 (Hooper et al., 2007; Kline, 2016). Standardized loadings of 0.30 or higher were considered as the cut-off for local fit. The cut-off for residuals at the local level was 0.10.

Pilot Tool Inter-Rater Reliability

After determining the construct validity of the pilot tool, the subsequently refined version of the IRMAT was assessed for inter-rater reliability. A sample of 50 published nursing integrative reviews was screened anonymously by two independent raters, with the lead researcher serving as third rater to settle any disagreements as needed (Pace et al., 2012). Prior data analysis demonstrated positivity rates of 65-69% (Kean, 2021). A sample size calculation with a minimum acceptable kappa of 0.40 aiming to detect a kappa of 0.80 with 90% power and a 0.05 alpha (Bujang & Baharum, 2017; Cantor, 1996; Sim & Wright, 2005) was

computed in the irr package of R (Gamer et al., 2019) to arrive at the sample size of 50 articles. Two independent raters were recruited from the pool of University of Cincinnati (UC) College of Nursing (CoN) PhD students. Raters received training from the doctoral student researcher on use of the IRMAT and clarification on how ratings categories are defined. The two independent raters were provided with the pilot version of the IRMAT and the accompanying tutorial document. The full text of the 50 articles in the sample were obtained by the lead researcher and were given to the two PhD student raters both in print and online format for independent screening using the IRMAT. All items from the pilot IRMAT were used to appraise each article by each rater independently scoring a "4" for "Yes," meaning the methodological item is present; "3" for "Partial Yes," meaning the item is only partially present; "2" for "Partial No," meaning the rater cannot determine whether or not the item is present; and "1" for "No," meaning the item is not present (Shea et al., 2017). Screening was conducted using the REDCap (Research Electronic Data Capture) electronic data capture tools hosted by the Center for Clinical and Translational Science and Training at the University of Cincinnati and Cincinnati Children's (Harris et al., 2019; Harris et al., 2009). After the independent screenings were completed, the results were shared between the two raters. Any disagreements between the two raters were discussed to achieve agreement. If agreement between the two raters could not be reached after discussion, the lead reviewer served as the deciding vote.

Inter-rater reliability (IRR) was calculated for all items using Cohen's k (Cohen, 1960; DeVellis, 2017). Ratings of items were grouped to reflect the presence ("Yes" and "Partial Yes") and the absence ("Partial No" and "No") of items (Pace et al., 2010; Shea, Bouter, et al., 2007). IRR is reported according to the following scale: < 0.00 – poor agreement; 0.00-0.20 – slight agreement; 0.21-0.40 – fair agreement; 0.41-0.60 – moderate agreement; 0.61-0.80 – substantial agreement; and 0.81-1.00 – almost perfect agreement (Landis & Koch, 1977). Items demonstrating substantial (0.61-0.80) and almost perfect (0.81-1.00) inter-rater agreement are included in the modified version of the appraisal tool (McHugh, 2012). Items demonstrating poor

(> 0.00), slight (0.00-0.20), fair (0.21-0.40), or moderate (0.41-0.60) agreement were candidates for modification or removal.

Quality

Quality is defined in terms of the threats to validity established by Polit and Beck (2017). To address issues of quality in terms of the participant sample for Aim 2, the inclusion criteria was designed to create as homogenous sample as possible. According to Polit and Beck (2017), although a homogenous sample may limit generalizability, it could potentially enhance the interpretability of the results, which is desirable for this study. Attention was paid to adequately sizing the sample of participants as a way of addressing potential threats to statistical conclusion validity. There is a possible threat to construct validity with the researcher's expectations for the study outcome (Polit & Beck, 2017); however, this threat is mitigated with the use of the expert panel, two independent raters to calculate inter-rater reliability, and the dissertation committee members serving as external controls.

Procedures for Human Subjects

Based on the nature of the study, there is anticipated minimal risk to participants. There is a potential risk to confidentiality of the study participants. The researcher made every attempt to ensure a safe online environment for participation. It was essential that the participants be made to feel safe and that their responses are confidential. All participant data was collected online via REDCap survey collection tools (Harris et al., 2019; Harris et al., 2009). The researcher was mindful of environment and any consequences of participants' involvement in the process and plan for ways to mitigate these barriers before the participation commences. Institutional review board (IRB) authorization for this study was sought and approved through the University of Cincinnati's IRB. All participant data was stored on secure cloud-based research computer servers at the University of Cincinnati.

Implications

This study would result in the first tool designed to appraise the methodological and reporting quality of integrative reviews in the nursing literature. The expectation is that this tool could be used by clinicians to evaluate the methodological rigor of integrative reviews, which implies that a published integrative review may or may not be used in clinical practice based on the outcome of the appraisal tool. Similarly, this tool may be used in educational settings by nursing students conducting evidence appraisal of integrative reviews in the nursing literature. While it is not anticipated that the IRMAT should be used as a scoring device, using the tool to evaluate evidence levels may be appropriate in an educational setting. At the conclusion of this study, implications for future research would include additional research to test construct validity of the tool with more diverse samples of articles.

Summary of Chapter 3

In summary, the methodological elements identified from the scoping review were coded and thematically analyzed to align to the Ellis Model of Information-Seeking Behavior, resulting in the generation of initial IRMAT items. The generated items were analyzed by a panel of nursing integrative review experts for face and content validity. Any items not demonstrating content validity were eligible for deletion or modification. After expert review, the subsequent version of the tool was used by at minimum 180 survey participants and two independent raters so that construct validity, internal consistency reliability, and inter-rater reliability could be analyzed statistically.

Chapter 4: Results

The 210 methodological elements which were identified and grouped as part of the scoping review (Chapter 2; Appendix B) were used to generate the items for the IRMAT. All 210 elements were printed on slips of paper and grouped initially according to the Ellis category alignment that had been determined as part of the scoping review classification: starting (n=22); searching (53); differentiating (n=53); extracting (n=32); verifying (n=38); and ending (n=12). (Items identified as browsing, chaining, or monitoring are now grouped together as the *Searching* category.) Using data condensation, the slips of paper were rearranged and regrouped into the categories of the Ellis model. Due to deficiencies in the use of the paper slips, it is difficult to ascertain an exact number of items per category; however, estimated percentages alongside comparisons of the scoping review categorization percentages are presented in Table 4.1.

Table 4.1

Items by Ellis Category from Scoping Review and Data Condensation

Ellis Category	Scoping Review Coverage	Data Condensation Coverage
Starting	11%	16%
Searching	25%	21%
Differentiating	25%	12%
Extracting	15%	13%
Verifying	18%	17%
Ending	6%	21%

The grouping for the Verifying category proved most difficult to categorize. Elements relating to analysis and comparison of similarities and differences were included in the Verifying category.

Item Generation

After categorization using the paper slips, items were entered into Nvivo (Release 1.6.2) for additional analysis of subthemes to assist with item generation. Wherever possible, items were generated using exact wording from the extracted methodological elements (Appendix B).

As part of the item generation process, Nvivo was used to group subthemes by the Ellis model categories and to generate percentages of coverage by subtheme (Appendix D). After the coding exercise, 34 items were generated (Table 4.2).

Table 4.2

34 Generated Items by Ellis Theme

No.	Theme	IRMAT Item
1	Starting	Was the protocol described?
2	Starting	Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?
3	Starting	Were the variables of interest clearly stated?
4	Starting	Were operational and conceptual definitions of variables provided?
5	Starting	Was the problem stated unambiguously and was it easy to identify?
6	Starting	Was there an explicit statement of the purpose or aim of the review?
7	Starting	Was there an explicit statement of the research question(s) the review addresses?
8	Starting	Was the sampling frame clearly stated?
9	Starting	Is a theoretical or conceptual framework used to guide the review?
10	Searching	Is a description of a comprehensive search provided?
11	Searching	Were detailed descriptions of the database search processes provided?
12	Searching	Was a reproducible line-by-line search strategy (or a sequence of terms for simpler interfaces) provided for at least one database?
13	Searching	Were other recommended approaches to searching the literature used?
14	Searching	If applicable to the purpose and type of literature included, was publication bias addressed?
15	Differentiation	Were the inclusion/exclusion criteria for the eligible sources clearly stated?
16	Differentiation	Was purposive sampling of the search results used?
17	Differentiation	Were the search results screened for relevance using a pre-specified set of eligibility criteria?
18	Differentiation	Is a flowchart of search results provided?
19	Differentiation	Were the retrieved sources divided into subgroups?
20	Extracting	Were the relevant data extracted from all sources using a pre-determined extraction instrument (for all subgroups, if applicable)?

No.	Theme	IRMAT Item
21	Extracting	Is the data extraction process explicit, unbiased, and reproducible?
22	Extracting	Were data analyzed using a review matrix?
23	Extracting	Is a data display assembling the data from retrieved sources provided?
24	Verifying	Was a systematic analytic method explicitly identified?
25	Verifying	Were commonalities and differences identified?
26	Verifying	Was conflicting evidence addressed?
27	Verifying	Was the quality of retrieved sources addressed in a meaningful way?
28	Verifying	Was quality verified by two independent reviewers?
29	Verifying	Were any articles excluded based on quality appraisal?
30	Ending	Were patterns, themes, relationships, or conclusions verified with sources?
31	Ending	Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?
32	Ending	Were the conclusions reported in table or diagrammatic form?
33	Ending	Were implications discussed for: research, practice, education, or policy?
34	Ending	Were methodological or other limitations of the review explicitly stated?

After the 34 items were finalized, a description and rating scale was written for all items. This process was completed by the primary researcher (Kean) and was reviewed by two co-researchers (Miller and Smith) for accuracy in coding, as well as clarity in the descriptions of the items, descriptions, and ratings scales. All notes provided by the secondary reviewers were incorporated into the IRMAT User Guide listing the 34 items, descriptions, and rating scales. Version 1 of the IRMAT User Guide which includes the subthemes, extracted methodological elements, and coding percentage of each methodological component by IRMAT item is presented as Appendix E.

Content and Face Validity

To test face and content validity, a panel of expert reviewers was recruited. Nursing integrative review experts were identified as having written about methodological aspects of integrative reviews in the nursing literature (Crossetti, 2012; Holly, 2019; Toronto & Remington,

2020; Whittemore & Knafl, 2005) and/or have edited, reviewed, or published a nursing integrative review in the past two years (Table 3.1). The University of Cincinnati Institutional Review Board determined that querying the expert panel about the face and content validity of the IRMAT items was non-human subjects research (Appendix F). In February of 2022, 12 nursing integrative review experts were identified and emailed invitations to participate as expert raters. Five completed responses were received and were analyzed for face and content validity of the 34 items.

Content validity was examined using the content validity index (CVI) scores at the item-level (Lynn, 1986; Polit & Beck, 2006). Thirty of the 34 items (88%) had a content validity equal to or greater than 0.80; however, according to the expert review, two of these 30 items demonstrated insufficient face validity ("*Was the protocol described?"* and "*Were patterns, themes, relationships, or conclusions verified with sources?*"). Using a combination of face validity, content validity, and expert feedback, six of the 34 items were deleted, leaving 28 items (Table 4.3). The full list of items with all expert comments and revisions is available as Appendix G.

Table 4.3

Items Remaining, Modified, or Deleted after Expert Review

No.	Theme	Content Validity	Original	Notes	Final
1	Starting	0.80	Was the protocol described?	Deleted (Face Validity)	
2	Starting	1.00	Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?	Retained with Modifications	Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?

No.	Theme	Content Validity	Original	Notes	Final
3	Starting	0.80	Were the variables of interest clearly stated?	Retained with Modifications	Were the factors of interest (concepts, variables, etc.) clearly stated?
4	Starting	0.80	Were operational and conceptual definitions of variables provided?	Retained with Modifications	Were conceptual and/or operational definitions of factors provided?
5	Starting	1.00	Was the problem stated unambiguously and was it easy to identify?	Retained with Modifications	Was the problem stated unambiguously and was it easy to identify?
6	Starting	1.00	Was there an explicit statement of the purpose or aim of the review?	Retained As Is	Was there an explicit statement of the purpose or aim of the review?
7	Starting	0.80	Was there an explicit statement of the research question(s) the review addresses?	Retained As Is	Was there an explicit statement of the research question(s) the review addresses?
8	Starting	1.00	Was the sampling frame clearly stated?	Retained with Modifications	If applicable to the purpose and/or type of literature included, was the sampling frame clearly stated?
9	Starting	0.60	Is a theoretical or conceptual framework used to guide the review?	Deleted (Face and Content Validity)	
10	Searching	1.00	Is a description of a comprehensive search provided?	Retained with Modifications	Is a description of a comprehensive search provided?
11	Searching	1.00	Were detailed descriptions of the database search processes provided?	Retained with Modifications	Were detailed descriptions of the database search processes provided?
12	Searching	1.00	Was a reproducible line-by-line search strategy (or a sequence of terms	Retained with Modifications	Are reproducible line-by-line search strategies

No.	Theme	Content Validity	Original	Notes	Final
			for simpler interfaces) provided for at least one database?		(or a sequence of terms for simpler interfaces) provided?
13	Searching	1.00	Were other recommended approaches to searching the literature used?	Retained with Modifications	Were other recommended approaches to searching the literature used?
14	Searching	0.80	If applicable to the purpose and type of literature included, was publication bias addressed?	Retained with Modifications	If applicable to the purpose and type of literature included, was the potential for publication bias acknowledged?
15	Differentiating	1.00	Were the inclusion/exclusion criteria for the eligible sources clearly stated?	Retained As Is	Were the inclusion/exclusion criteria for the eligible sources clearly stated?
16	Differentiating	0.60	Was purposive sampling of the search results used?	Deleted (Face and Content Validity)	
17	Differentiating	1.00	Were the search results screened for relevance using a pre-specified set of eligibility criteria?	Retained As Is	Were the search results screened for relevance using a pre-specified set of eligibility criteria?
18	Differentiating	1.00	Is a flowchart of search results provided?	Retained with Modifications	Is a flowchart of search results provided?
19	Differentiating	0.80	Were the retrieved sources divided into subgroups?	Retained with Modifications	If applicable to the purpose and/or type of literature included, were the retrieved sources divided into subgroups?
20	Extracting	0.80	Were the relevant data extracted from all sources using a pre-determined extraction instrument (for all subgroups, if applicable)?	Retained As Is	Were the relevant data extracted from all sources using a pre-determined extraction

No.	Theme	Content Validity	Original	Notes	Final
					instrument (for all subgroups, if applicable)?
21	Extracting	1.00	Is the data extraction process explicit, unbiased, and reproducible?	Retained with Modifications	Is the data extraction process explicit, unbiased, and reproducible?
22	Extracting	1.00	Were data analyzed using a review matrix?	Retained As Is	Were data analyzed using a review matrix?
23	Extracting	1.00	Is a data display assembling the data from retrieved sources provided?	Retained As Is	Is a data display assembling the data from retrieved sources provided?
24	Verifying	0.80	Was a systematic analytic method explicitly identified?	Retained As Is	Was a systematic analytic method explicitly identified?
25	Verifying	0.80	Were commonalities and differences identified?	Retained As Is	Were commonalities and differences identified?
26	Verifying	1.00	Was conflicting evidence addressed?	Retained As Is	Was conflicting evidence addressed?
27	Verifying	1.00	Was the quality of retrieved sources addressed in a meaningful way?	Retained with Modifications	Was the quality of retrieved sources addressed in a meaningful way?
28	Verifying	1.00	Was quality verified by two independent reviewers?	Retained As Is	Was quality verified by two independent reviewers?
29	Verifying	0.60	Were any articles excluded based on quality appraisal?	Deleted (Content Validity)	
30	Ending	0.80	Were patterns, themes, relationships, or conclusions verified with sources?	Deleted (Face Validity)	
31	Ending	1.00	Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?	Retained As Is	Were important elements or conclusions (for all subgroups, if applicable) synthesized into a

No.	Theme	Content Validity	Original	Notes	Final
					summation of the topic or phenomenon?
32	Ending	0.60	Were the conclusions reported in table or diagrammatic form?	Deleted (Content Validity)	
33	Ending	0.80	Were implications discussed for: research, practice, education, or policy?	Retained As Is	Were implications discussed for: research, practice, education, or policy?
34	Ending	1.00	Were methodological or other limitations of the review explicitly stated?	Retained As Is	Were methodological or other limitations of the review explicitly stated?

Based on expert feedback, 14 of the remaining items were modified for clarity. Details about the modifications to the items are provided in Table 4.3, with additional commentary and expert comments provided in detail in Appendix G. The experts' recommendations for modifications more commonly centered around the item rating scales and, to some extent, the item descriptions rather than the wording of the actual items. Underlining is added to the text of Appendix G to further illustrate additions, with strikethrough of text being used to highlight deletions. All modifications to items, descriptions, and rating scales based on expert feedback were compiled by the primary researcher (Kean) and were reviewed and approved by two secondary reviewers (Miller and Smith) independently. After review by the expert panel, 28 items remained.

Sample Demographic Characteristics

Prior to distributing the survey to the purposive sample, approval was obtained from the University of Cincinnati Institutional Review Board and is attached as Appendix H. Participants identified from the sampling plan (Chapter 3, *Sampling Plan*) received an invitation to participate in the study via the REDCap (Research Electronic Data Capture) electronic data capture tools

hosted by the Center for Clinical and Translational Science and Training at the University of Cincinnati and Cincinnati Children's (Harris et al., 2019; Harris et al., 2009). Those who elected to participate were provided with the full text of one identical sample article as part of the survey tool and used the 28 items of the refined pilot version of the IRMAT to assess the accompanying sample article. At the completion of the survey, optional demographic data (age range, gender, education level, and primary employment responsibility) was collected from participants.

Demographics for the 204 participants are provided in Table 4.4. Most survey respondents identified as women (*n* = 162, 79.4%). The age range of 27% of participants (*n* = 55) was between 40 and 49 years of age, with 93% of respondents (*n* = 190) between the ages of 30 and 69 years. Almost all respondents (*n* = 203, 99.5%) reported that their highest level of education was a Master's degree or higher. Eighty-seven percent of participants (*n* = 179) reported having a Ph.D., DNP, or higher degree. The most prominently reported employment area was education for 96 respondents (47.1%), followed by research employment (*n* = 74, 36.3%), and employment in a clinical area (*n* = 22, 10.8%).

Table 4.4

Survey Participant Demographics

Demographic Variable	Number of Participants	Valid Percentage
	Age	
20-29 years	5	2.5%
30-39 years	53	26%
40-49 years	55	27%
50-59 years	51	25%
60-69 years	31	15.2%
70-79 years	6	2.9%
80 years or older	1	0.5%
No Response	2	2.5%
	Gender Identity	
Man	36	17.6%
Woman	162	79.4%
Non-binary/Non-conforming	2	1%
Prefer not to respond	2	1%
No Response	2	1%

Demographic Variable	Number of Participants	Valid Percentage
	Highest Level of Education	
Bachelor's Degree	1	0.5%
Master's Degree	24	11.8%
Ph.D., DNP, or higher	179	87.7%
	Primary Employment Area	
Clinical	22	10.8%
Education	96	47.1%
Research	74	36.3%
Other[a]	11	5.4%
No Response	1	0.5%

[a]Other responses for primary employment area (*n* = 1, unless otherwise noted) included: academic; clinical research/clinical trials management/CRO; consulting; doctoral student; full time editor; intellectual property; medical product company president; medical science liaison; publishing; retired (*n* = 2).

Sample Descriptives

Descriptives for all items are provided in Table 4.5. A total of 204 completed responses were received. Items 8, 14, and 19 allow for "N/A" responses. Excluding these three missing items results in 160 completed responses used for initial statistical analysis.

Table 4.5

Descriptive Statistics for All 28 Items (N = 204)

No.	n	Mis.	Mean	SD	Med.	Min	Max	Skew.	SE	Kurt.	SE
2	204	0	3.56	0.776	4.00	1	4	-1.986	0.170	3.584	0.339
3	204	0	3.57	0.702	4.00	1	4	-1.929	0.170	4.053	0.339
4	204	0	3.03	1.096	3.00	1	4	-0.784	0.170	-0.755	0.339
5	204	0	3.54	0.758	4.00	1	4	-1.831	0.170	3.055	0.339
6	204	0	3.92	0.355	4.00	1	4	-5.169	0.170	30.955	0.339
7	204	0	3.72	0.741	4.00	1	4	-2.789	0.170	6.928	0.339
8	198	6	3.70	0.719	4.00	1	4	-2.538	0.173	5.795	0.344
10	204	0	3.84	0.453	4.00	1	4	-3.206	0.170	11.597	0.339
11	204	0	3.79	0.421	4.00	2	4	-1.627	0.170	1.242	0.339
12	204	0	3.70	0.624	4.00	1	4	-2.372	0.170	6.007	0.339
13	204	0	2.69	1.251	3.00	1	4	-0.282	0.170	-1.567	0.339
14	187	17	2.29	1.242	2.00	1	4	0.276	0.178	-1.558	0.354
15	204	0	3.88	0.404	4.00	1	4	-4.515	0.170	25.258	0.339
17	204	0	3.75	0.631	4.00	1	4	-2.595	0.170	6.198	0.339

No.	n	Mis.	Mean	SD	Med.	Min	Max	Skew.	SE	Kurt.	SE
18	204	0	3.96	0.260	4.00	1	4	-8.633	0.170	87.492	0.339
19	177	27	3.36	1.046	4.00	1	4	-1.389	0.183	0.442	0.363
20	204	0	3.59	0.753	4.00	1	4	-1.961	0.170	3.294	0.339
21	204	0	3.32	0.974	4.00	1	4	-1.269	0.170	0.403	0.339
22	204	0	3.35	0.998	4.00	1	4	-1.314	0.170	0.383	0.339
23	204	0	3.54	0.879	4.00	1	4	-1.947	0.170	2.729	0.339
24	204	0	3.70	0.646	4.00	1	4	-2.388	0.170	5.599	0.339
25	204	0	3.42	0.887	4.00	1	4	-1.479	0.170	1.257	0.339
26	204	0	2.46	1.084	2.00	1	4	0.125	0.170	-1.263	0.339
27	204	0	3.58	0.742	4.00	1	4	-1.698	0.170	2.035	0.339
28	204	0	3.59	0.846	4.00	1	4	-1.973	0.170	2.651	0.339
31	204	0	3.72	0.632	4.00	1	4	-2.653	0.170	7.369	0.339
33	204	0	3.77	0.474	4.00	1	4	-2.278	0.170	6.350	0.339
34	204	0	3.75	0.569	4.00	1	4	-2.717	0.170	8.124	0.339

Note. Missing Completed Reponses (Mis.); Standard Deviation (SD); Median (Med.); Minimum (Min); Maximum (Max); Skewness (Skew.); Standard Error (SE); Kurtosis (Kurt.)

Skewness and kurtosis for all items was outside of acceptable ranges. It was determined that robust maximum likelihood estimation would be used for confirmatory factor analysis since the normality assumption was violated (Maydeu-Olivares, 2017), and all items were included in further analysis. Histograms for all items are shown in Figure 4.1.

Construct Validity

Examination of the covariance (Figure 4.2) and correlation (Figure 4.3) matrices show low co-variance among items, which is to be expected due to the means of the items being "lopsided" (DeVellis, 2017, p. 143). Visual examination of the heat maps of the matrices showed that items 18 ("*Is a flowchart of search results provided?*"), 10 ("*Is a description of a comprehensive search provided?*"), and 6 ("*Was there an explicit statement of the purpose or aim of the review?*") showed low item covariance with other items, were poorly correlated, and were candidates for deletion.

Figure 4.1

Histograms for All Items

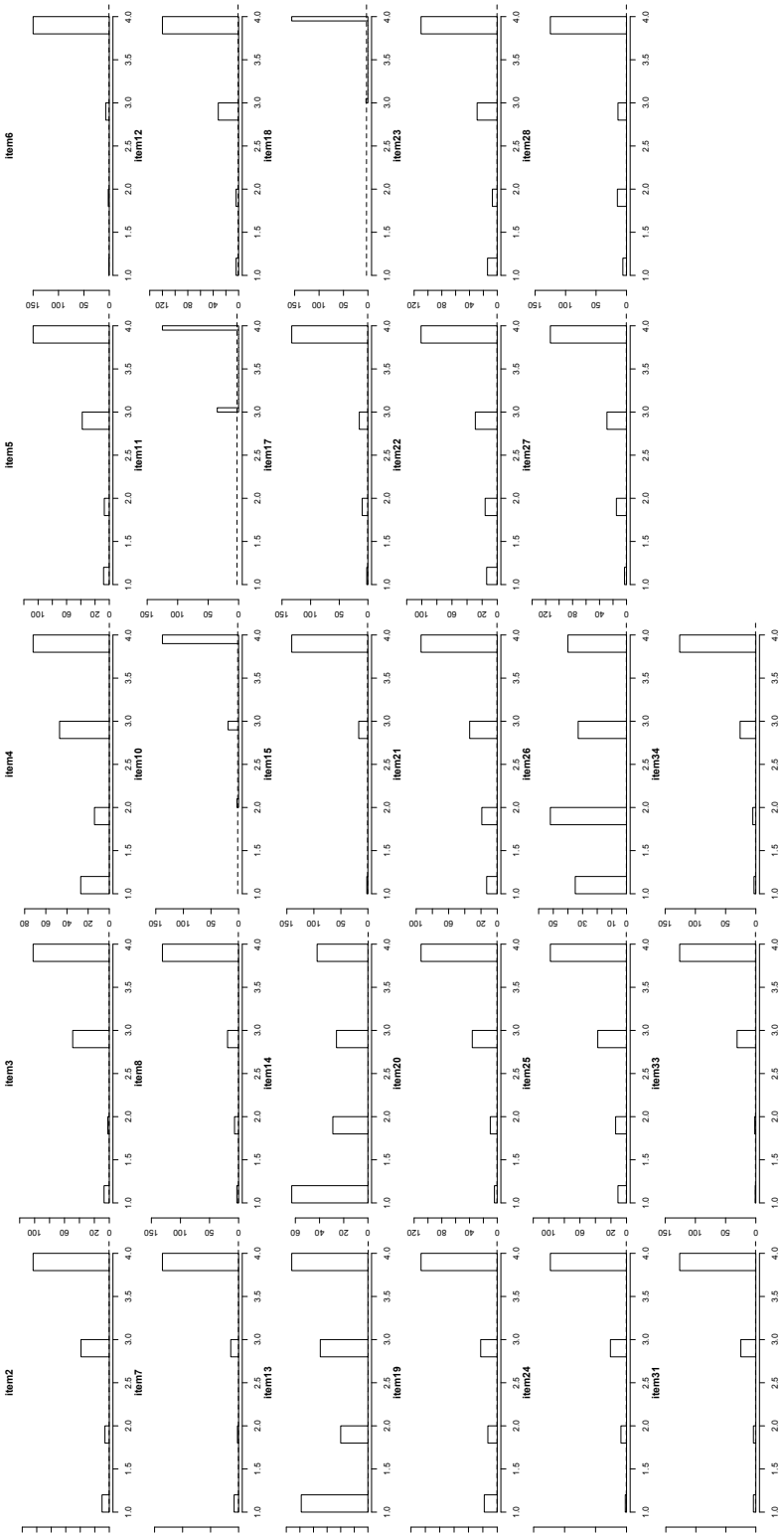
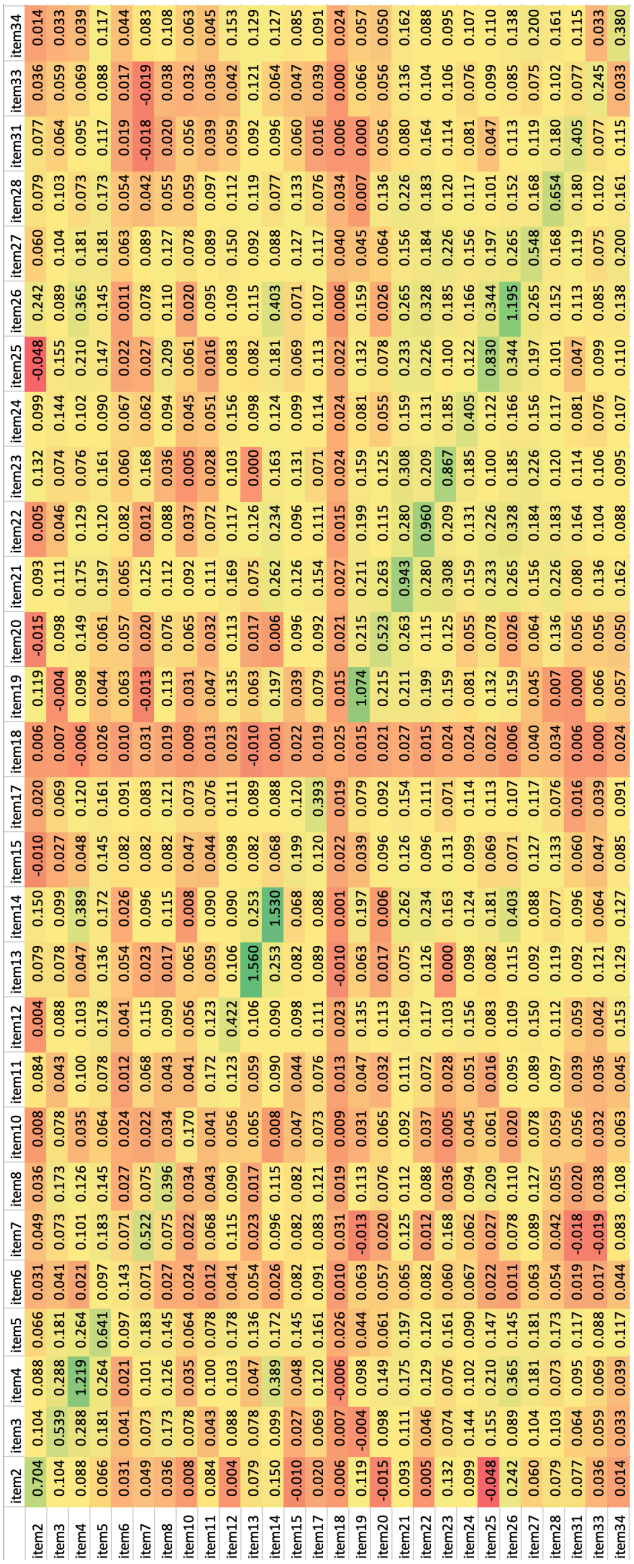


Figure 4.2

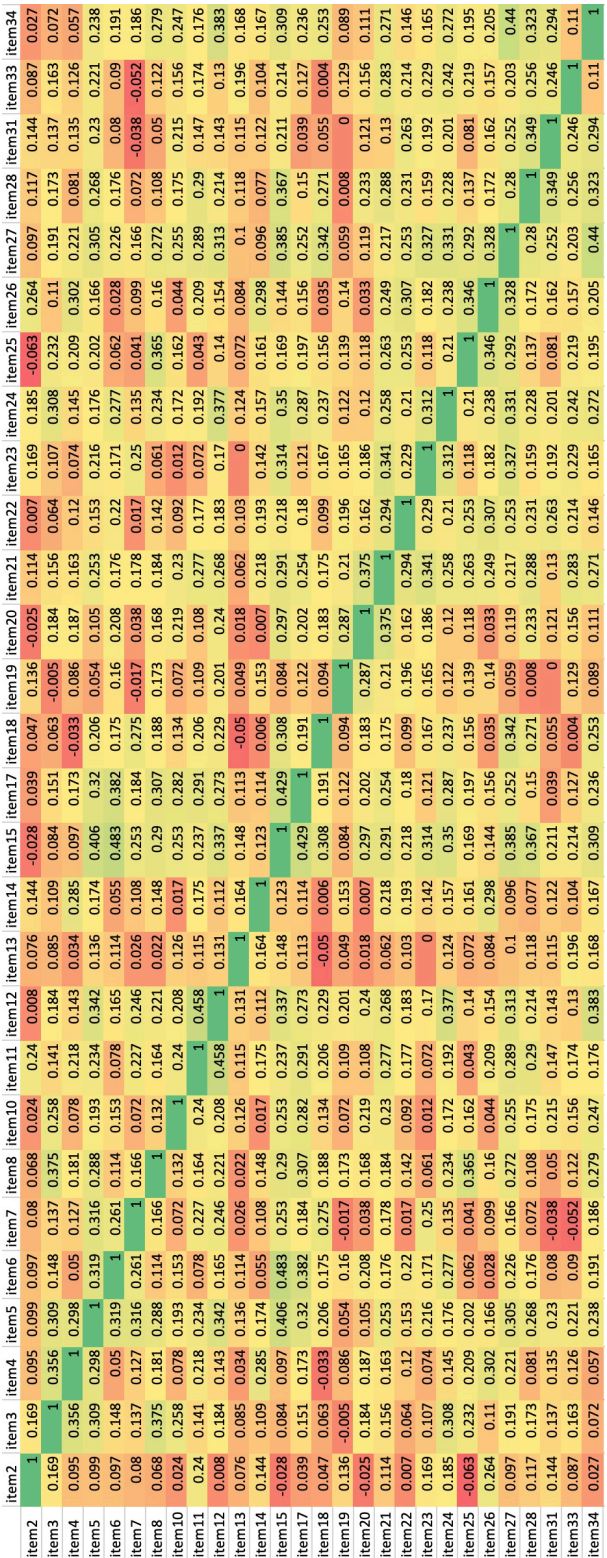
Covariance Matrix Heat Map



	item2	item3	item4	item5	item6	item7	item8	item10	item11	item12	item13	item14	item15	item17	item18	item19	item20	item21	item22	item23	item24	item25	item26	item27	item28	item31	item33	item34
item2	0.704	0.104	0.088	0.066	0.031	0.049	0.036	0.008	0.084	0.004	0.079	0.150	-0.010	0.020	0.006	0.119	-0.015	0.093	0.005	0.132	0.099	-0.048	0.242	0.060	0.079	0.077	0.036	0.014
item3	0.104	0.539	0.288	0.181	0.041	0.073	0.173	0.078	0.043	0.088	0.078	0.099	0.027	0.069	0.007	-0.004	0.098	0.111	0.046	0.074	0.144	0.155	0.089	0.104	0.103	0.064	0.059	0.033
item4	0.088	0.288	1.219	0.264	0.021	0.101	0.126	0.035	0.100	0.103	0.047	0.389	0.048	0.120	-0.006	0.098	0.149	0.175	0.129	0.076	0.102	0.210	0.365	0.181	0.073	0.095	0.069	0.039
item5	0.066	0.181	0.264	0.641	0.097	0.183	0.145	0.064	0.078	0.178	0.136	0.172	0.145	0.161	0.026	0.044	0.061	0.197	0.120	0.161	0.090	0.147	0.145	0.181	0.173	0.117	0.088	0.117
item6	0.031	0.041	0.021	0.097	0.143	0.071	0.027	0.024	0.012	0.041	0.054	0.026	0.082	0.091	0.010	0.063	0.057	0.065	0.082	0.060	0.067	0.022	0.011	0.063	0.054	0.019	0.017	0.044
item7	0.049	0.073	0.101	0.183	0.071	0.522	0.075	0.022	0.068	0.115	0.023	0.096	0.082	0.083	0.031	-0.013	0.020	0.125	0.012	0.168	0.062	0.027	0.078	0.089	0.042	-0.018	-0.019	0.083
item8	0.036	0.173	0.126	0.145	0.027	0.075	0.396	0.034	0.043	0.090	0.017	0.115	0.082	0.121	0.019	0.113	0.076	0.112	0.088	0.036	0.094	0.209	0.110	0.127	0.055	0.020	0.038	0.108
item10	0.008	0.078	0.035	0.064	0.024	0.022	0.034	0.170	0.041	0.056	0.065	0.008	0.047	0.073	0.009	0.031	0.065	0.092	0.037	0.005	0.045	0.061	0.020	0.078	0.059	0.056	0.032	0.063
item11	0.084	0.043	0.100	0.078	0.012	0.068	0.043	0.041	0.172	0.123	0.059	0.090	0.044	0.076	0.013	0.047	0.032	0.111	0.072	0.028	0.051	0.016	0.095	0.089	0.097	0.039	0.036	0.045
item12	0.004	0.088	0.103	0.178	0.041	0.115	0.090	0.056	0.123	0.422	0.106	0.090	0.098	0.111	0.023	0.135	0.113	0.169	0.117	0.103	0.156	0.083	0.109	0.150	0.112	0.059	0.042	0.153
item13	0.079	0.078	0.047	0.136	0.054	0.023	0.017	0.065	0.059	0.106	1.560	0.253	0.082	0.089	-0.010	0.063	0.017	0.075	0.126	0.000	0.098	0.082	0.115	0.092	0.119	0.092	0.121	0.129
item14	0.150	0.099	0.389	0.172	0.026	0.096	0.115	0.008	0.090	0.090	0.253	1.530	0.068	0.088	0.001	0.197	0.006	0.262	0.234	0.163	0.124	0.181	0.403	0.088	0.077	0.096	0.064	0.127
item15	-0.010	0.027	0.048	0.145	0.082	0.082	0.082	0.047	0.044	0.098	0.082	0.068	0.199	0.120	0.022	0.039	0.096	0.126	0.096	0.131	0.099	0.069	0.071	0.127	0.133	0.060	0.047	0.085
item17	0.020	0.069	0.120	0.161	0.091	0.083	0.121	0.073	0.076	0.111	0.089	0.088	0.120	0.393	0.019	0.079	0.092	0.154	0.111	0.071	0.114	0.113	0.107	0.117	0.076	0.016	0.039	0.091
item18	0.006	0.007	0.007	0.010	0.010	0.031	0.019	0.009	0.013	0.023	-0.010	0.001	0.022	0.019	0.025	0.015	0.021	0.027	0.015	0.024	0.024	0.022	0.006	0.040	0.034	0.006	0.000	0.024
item19	0.119	-0.004	0.098	0.044	0.063	-0.013	0.113	0.031	0.047	0.135	0.063	0.197	0.039	0.079	0.015	1.074	0.215	0.211	0.199	0.159	0.081	0.132	0.159	0.045	0.007	0.000	0.066	0.057
item20	-0.015	0.098	0.149	0.061	0.057	0.020	0.076	0.065	0.032	0.113	0.017	0.006	0.096	0.092	0.021	0.215	0.523	0.263	0.115	0.125	0.055	0.078	0.026	0.064	0.136	0.056	0.056	0.050
item21	0.093	0.111	0.099	0.197	0.065	0.125	0.112	0.092	0.111	0.169	0.098	0.262	0.126	0.154	0.027	0.211	0.263	0.943	0.280	0.308	0.159	0.233	0.265	0.156	0.226	0.080	0.136	0.162
item22	0.005	0.046	0.129	0.120	0.082	0.012	0.088	0.037	0.072	0.117	0.126	0.234	0.096	0.111	0.015	0.199	0.115	0.280	0.960	0.209	0.131	0.226	0.328	0.184	0.183	0.164	0.104	0.088
item23	0.132	0.074	0.076	0.161	0.060	0.168	0.036	0.005	0.028	0.103	0.000	0.163	0.131	0.071	0.024	0.159	0.125	0.308	0.209	0.867	0.185	0.100	0.185	0.226	0.120	0.114	0.106	0.095
item24	0.099	0.144	0.102	0.090	0.067	0.062	0.094	0.045	0.051	0.156	0.098	0.124	0.099	0.114	0.024	0.081	0.055	0.159	0.131	0.185	0.405	0.122	0.166	0.156	0.117	0.081	0.076	0.107
item25	-0.048	0.155	0.210	0.147	0.022	0.027	0.209	0.061	0.016	0.083	0.082	0.181	0.069	0.113	0.022	0.132	0.078	0.233	0.226	0.100	0.122	0.830	0.344	0.197	0.101	0.047	0.099	0.110
item26	0.242	0.089	0.365	0.145	0.011	0.078	0.110	0.020	0.095	0.109	0.115	0.403	0.071	0.107	0.006	0.159	0.026	0.265	0.328	0.185	0.166	0.344	1.195	0.265	0.152	0.113	0.085	0.138
item27	0.060	0.104	0.181	0.181	0.063	0.089	0.127	0.078	0.089	0.150	0.092	0.088	0.127	0.117	0.040	0.045	0.064	0.156	0.184	0.226	0.156	0.197	0.265	0.548	0.168	0.119	0.075	0.200
item28	0.079	0.103	0.073	0.173	0.054	0.042	0.055	0.059	0.097	0.112	0.119	0.077	0.133	0.076	0.034	0.007	0.136	0.226	0.183	0.120	0.117	0.101	0.152	0.168	0.654	0.180	0.102	0.161
item31	0.077	0.064	0.095	0.117	0.019	-0.018	0.020	0.056	0.039	0.059	0.092	0.096	0.060	0.016	0.006	0.000	0.056	0.080	0.164	0.114	0.081	0.047	0.113	0.119	0.180	0.405	0.077	0.115
item33	0.036	0.059	0.069	0.088	0.017	-0.019	0.038	0.032	0.036	0.042	0.121	0.064	0.047	0.039	0.000	0.066	0.056	0.136	0.104	0.106	0.076	0.099	0.085	0.075	0.102	0.077	0.245	0.033
item34	0.014	0.033	0.039	0.117	0.044	0.083	0.108	0.063	0.045	0.153	0.129	0.127	0.085	0.091	0.024	0.057	0.050	0.162	0.088	0.095	0.107	0.110	0.138	0.200	0.161	0.115	0.033	0.380

Figure 4.3

Correlation Plot

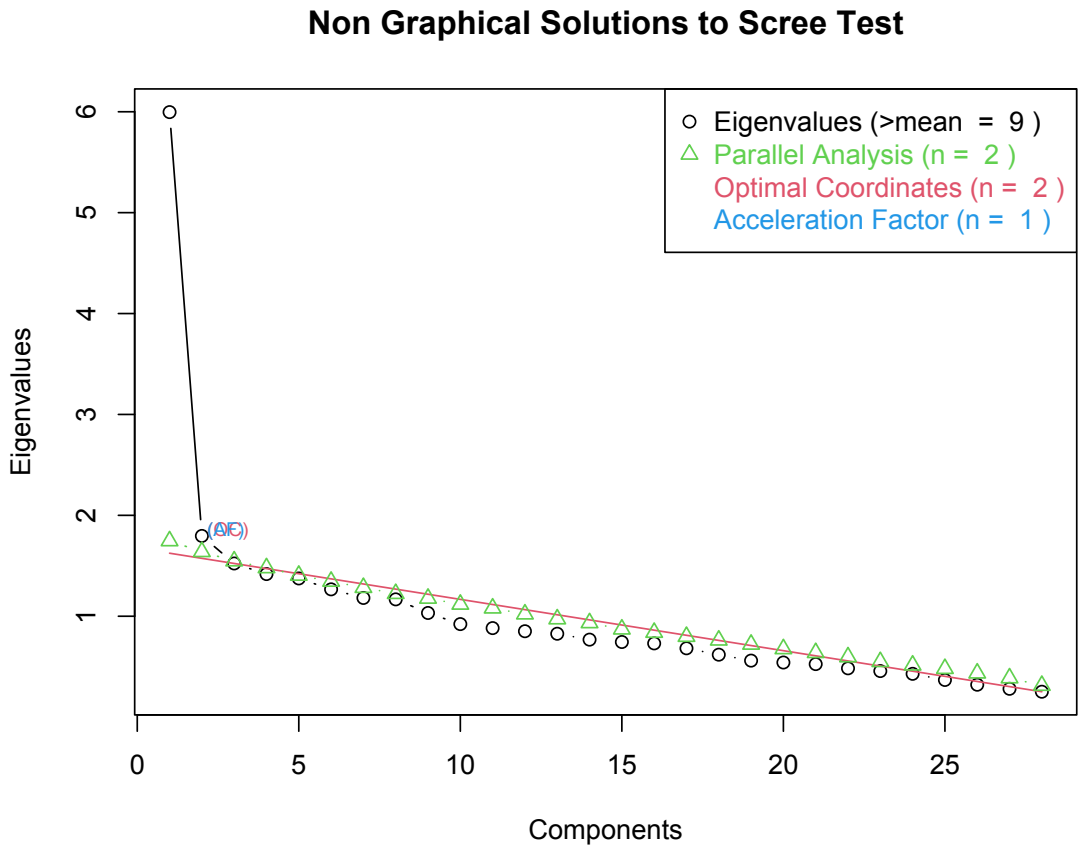


Exploratory Factor Analysis

Baseline models consisting of a single-factor model of all 28 items (Model A) and a 6-factor model of all 28 items based on the Ellis model (Model B) were created to serve as comparisons to the pruned model derived from the exploratory factor analysis. The initial exploratory factor analysis (EFA) model consisted of all 28 items. Based on the eigenvalues over 1 and a visual assessment of the scree plot (Figure 4.4) generated in R using the nFactors package (Raiche, 2022), a model was fit consisting of 2 factors using a promax rotation with a cut-off of 0.30.

Figure 4.4

Scree Plot



Using the Ellis model as an underlying conceptual framework to make pruning decisions, the pattern and structure matrices of several iterations of exploratory factor analysis were analyzed until all items loaded onto one of the two factors with loadings > 0.30. During separate rounds of the iterative exploration, items 19 and 14 (which allowed for "N/A" responses) were removed, which resulted in datasets of 183 and 198 completed responses, respectively. To achieve an equal number of items loading onto each of the two factors, item 7 ("*Was there an explicit statement of the research question(s) the review addresses?")* was removed. Although item 7 did load onto factor one with a 0.356 loading, there was conceptual basis for removing the item, as this aspect of the Ellis model was addressed by a remaining item – "*Was the problem stated unambiguously and was it easy to identify?"* (item 5) – with a higher loading (0.585). During expert review, it was noted that there was redundancy between items 5, 6, and 7; thus, retaining only one of these three similar items in the final tool also improved clarity. At the completion of this iterative EFA process, ten items were removed from the model. Items removed, along with reasons for removal, are detailed in Table 4.6.

Table 4.6

Pruned Items with Reason for Removal

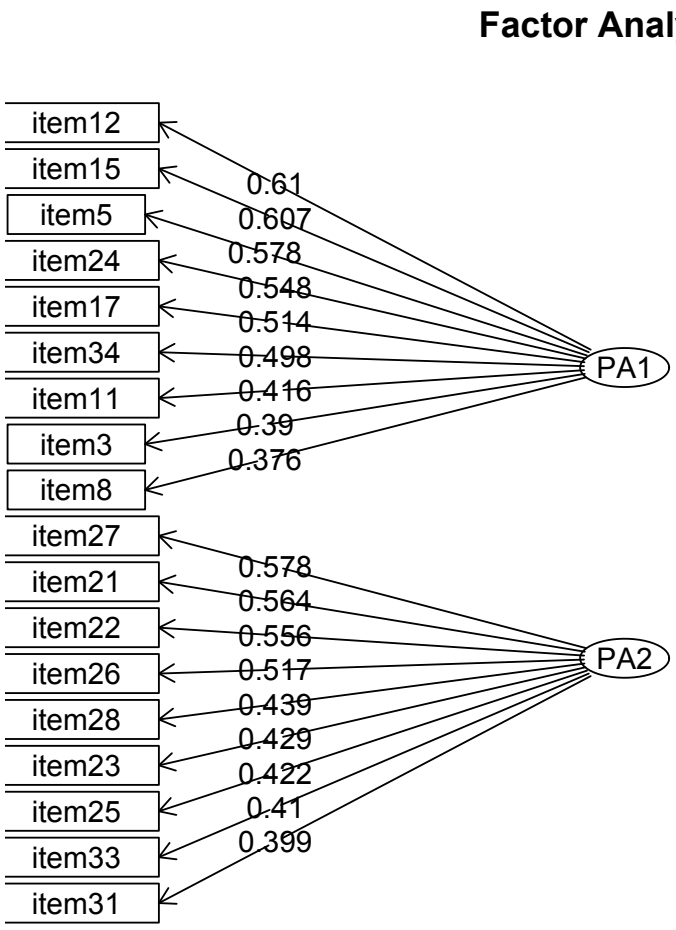
No.	Item	Reason for Removal
2	Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?	Loading below cutoff (< 0.30)
4	Were conceptual and/or operational definitions of factors provided?	Loading below cutoff (< 0.30)
6	Was there an explicit statement of the purpose or aim of the review?	Removed due to low covariance/correlation
7	Was there an explicit statement of the research question(s) the review addresses?	Loaded onto F1 (0.356) but removed to balance factor loadings, supported by expert feedback and conceptual model
10	Is a description of a comprehensive search provided?	Removed due to low covariance/correlation
13	Were other recommended approaches to searching the literature used?	Loading below cutoff (< 0.30)

No.	Item	Reason for Removal
14	If applicable to the purpose and type of literature included, was the potential for publication bias acknowledged?	Loading below cutoff (< 0.30)
18	Is a flowchart of search results provided?	Removed due to low covariance/correlation
19	If applicable to the purpose and/or type of literature included, were the retrieved sources divided into subgroups?	Loading below cutoff (< 0.30)
20	Were the relevant data extracted from all sources using a pre-determined extraction instrument (for all subgroups, if applicable)?	Loading below cutoff (< 0.30)

Model C consisted of a total of 18 items (3, 5, 8, 11, 12, 15, 17, 21, 22, 23, 24, 25, 26, 27, 28, 31, 33, 34) loading on two factors (Figure 4.5). As noted previously, item 7 ("*Was there an explicit statement of the research question(s) the review addresses?")* initially loaded onto factor one but was removed, resulting in nine items loading onto each factor. The cumulative variance for all factors was 25.2%. The two factors for Model C can be represented as: Factor 1 = *Design* and Factor 2 = *Analysis*. The nine items loading onto factor one (3, 5, 8, 11, 12, 15, 17, 24, and 34) all pertain in some regard to decisions that are intended to be made *a priori*. These items represent functions surrounding problem identification, search parameters, inclusion/exclusion criteria of identified sources, screening eligibility criteria, and use of a systematic analytic method – all of which should be determined prior to analysis of the retrieved sources. Factor two pertains to how the retrieved sources were analyzed, as well as outcomes and implications of the actual analysis. The nine items loading onto factor two (21, 22, 23, 25, 26, 27, 28, 31, and 33) all represent aspects of *Verifying*, *Extracting*, or *Ending* – later-stage Ellis model categories.

Figure 4.5

Pruned Model Loadings



Confirmatory Factor Analysis

In comparing the overall global fit of the baseline Models A (single factor, 28 items) and B (6 factors of Ellis framework, 28 items) with the modified Model C (pruned model derived from EFA: two factors, 18 items), Model C was found to have the best overall fit and was an improvement over both Models A and B. The outputs of the confirmatory factor analyses fit indices for all models are shown in Table 4.7.

Table 4.7

Confirmatory Factor Analysis Fit for Models A, B, and C

No.	KMO	Bartlett's			Model χ2	*p* (χ2)	GFI	AGFI	CFI	RMSEA	SRMR
		χ2	*df*	*p*							
A	0.78	1213.7	27	< 2.2e-16	940.877	<.001	0.995	0.994	0.045	0.123	1.501
B	0.78	1213.7	27	< 2.2e-16	475.972	<.001	0.997	0.996	0.818	0.051	0.066
C	0.83	520.9	17	< 2.2e-16	191.573	<.001	0.996	0.994	0.903	0.047	0.056

Note. Kaiser-Meyer-Olkin (KMO); goodness of fit index (GFI); adjusted goodness of fit index (AGFI); comparative fit index (CFI); root mean square error of approximation (RMSEA); standardized root mean square residual (SRMR)

Except for model chi-square, goodness of fit for Model C was found to be very good or good for the following reported indices: GFI = 0.996; AGFI = 0.994; CFI = 0.903; RMSEA = 0.047, and SRMR = 0.056. As noted by Babyak and Green (2010) and Kline (2016), the significance of model chi-square is very sensitive to sample size. As such, the assessment of best overall model fit was conducted holistically, examining numerous indices of fit. Model C was equivalent to or showed improvement in global model fit over Model A in GFI, AGFI, CFI, RMSEA, and SRMR (Table 4.7). Model B did have slightly better GFI and AGFI over Model C, but CFI, RMSEA, and SRMR all showed improved fit with Model C. There were slight issues with local fit at the item level for all models; however, Model C again demonstrated improvement with local fit with the least number of residuals above 0.1. Standardized and unstandardized loadings for both factors and all 18 items of Model C were statistically significant ($p < 0.05$) and are reported in detail in Table 4.8.

Table 4.8

Standardized and Unstandardized Loadings by Factor and Item for Model C

No.	Estimate	Std. Err	z-value	P-value	Std. lv	Std. all	R^2
		Factor 1: Design					
3	0.268	0.069	3.874	0.001	0.268	0.384	0.147
5	0.428	0.076	5.626	0.001	0.428	0.561	0.314
8	0.272	0.085	3.716	0.001	0.272	0.379	0.143
11	0.171	0.031	5.452	0.001	0.171	0.422	0.178
12	0.373	0.075	4.958	0.001	0.373	0.605	0.365
15	0.245	0.082	2.992	0.003	0.245	0.600	0.360
17	0.297	0.083	3.574	0.001	0.297	0.502	0.252
24	0.350	0.075	4.648	0.001	0.350	0.564	0.318
34	0.289	0.076	3.805	0.001	0.289	0.507	0.257
		Factor 2: Analysis					
21	0.545	0.081	6.757	0.001	0.545	0.556	0.309
22	0.486	0.088	5.490	0.001	0.486	0.491	0.241
23	0.397	0.080	4.983	0.001	0.397	0.458	0.210
25	0.361	0.073	4.922	0.001	0.361	0.404	0.163
26	0.474	0.073	6.451	0.001	0.474	0.438	0.192
27	0.464	0.081	5.751	0.001	0.464	0.629	0.395
28	0.412	0.083	4.990	0.001	0.412	0.487	0.238
31	0.248	0.046	5.395	0.001	0.248	0.411	0.169
33	0.189	0.045	4.232	0.001	0.189	0.399	0.159

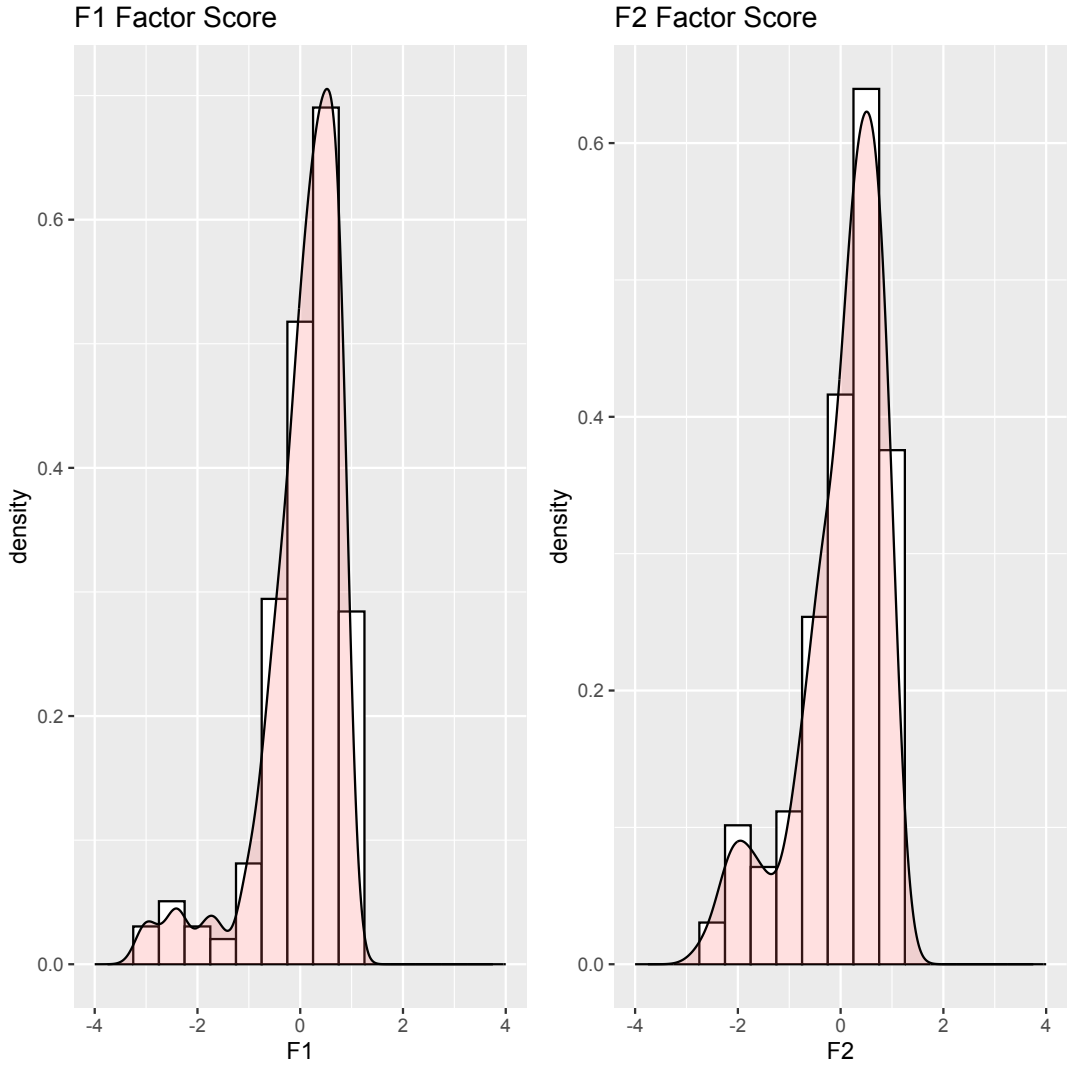
Note. Std. Err = Standard Error; Std. lv = Standardized latent variables ; Std. all = Standardized all (both latent and observed variables)

Internal Consistency Reliability

Cronbach's α values for both factors were found to be respectable (0.70-0.80) (DeVellis, 2017). Cronbach's α values for both factors were estimated using the ltm package in R (Rizopoulos, 2006) and were found to be: factor one = 0.741 and factor two = 0.719. Factor score reliability was estimated using the lavaan package (Rosseel, 2011) and was found to be: factor one = 0.838 and factor two = 0.826. Histograms with density plot overlays for each factor are included as Figure 4.5.

Figure 4.6

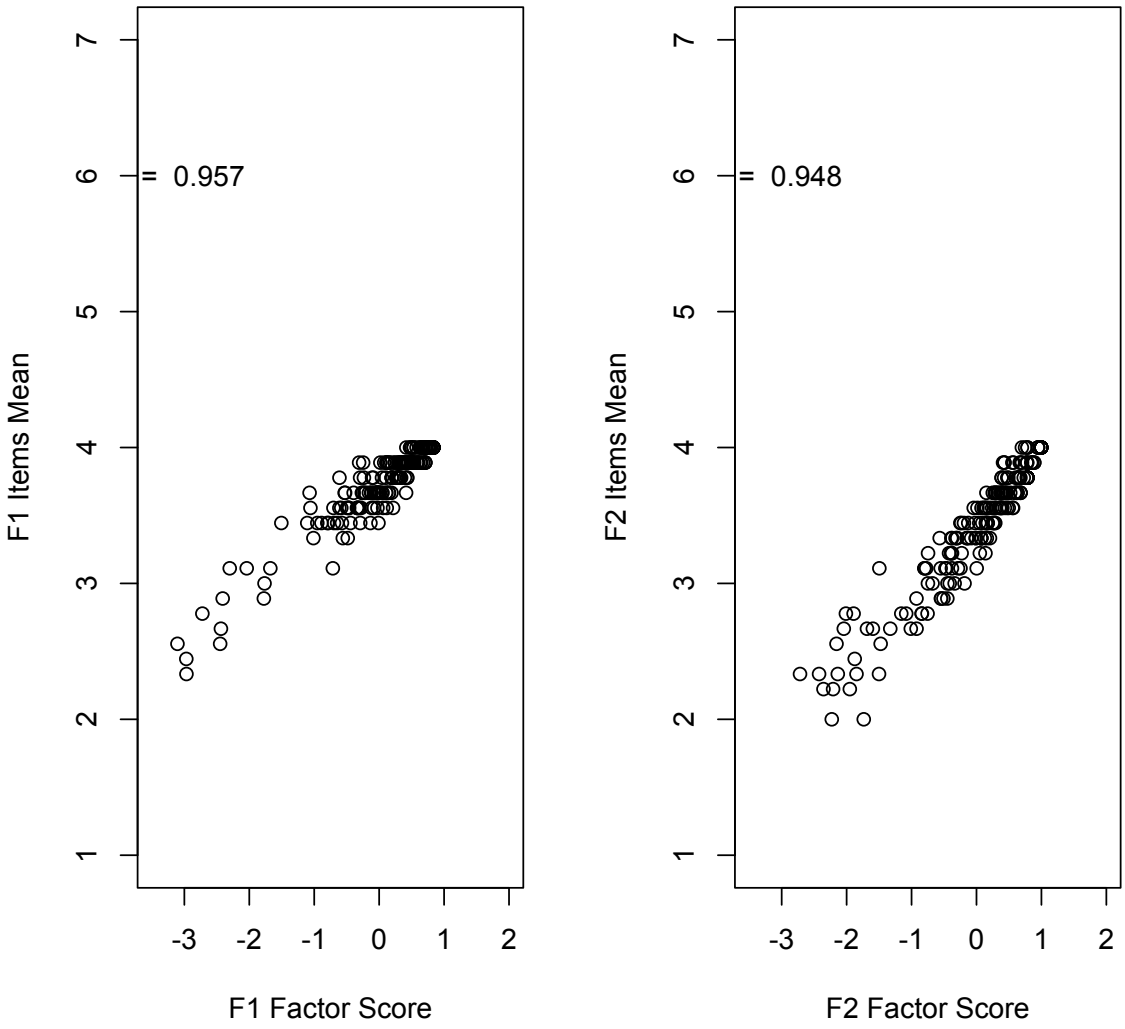
Histograms with Density Plot Overlays for Factors One and Two



Scatterplots for observed versus mean factor scores for Model C are provided as Figure 4.6.

Figure 4.7

Scatterplot for Observed Versus Mean Factor Scores for Factors One and Two



Inter-Rater Reliability

Inter-rater reliability (IRR) was assessed for the refined version of the IRMAT version 1.2 (Appendix G). The recruitment of two nursing research PhD students to serve as raters for IRR testing of the IRMAT was approved by the University of Cincinnati Institutional Review Board (Appendix H). The two raters used the REDCap (Research Electronic Data Capture) electronic

data capture tools hosted by the Center for Clinical and Translational Science and Training at the University of Cincinnati and Cincinnati Children's (Harris et al., 2019; Harris et al., 2009) to use the 18-item version of the IRMAT to appraise the full text of 50 nursing integrative review articles on the subject of nursing leadership/management. This appraisal process was completed independently by the two raters, and any disagreements amongst the two raters were settled by the primary researcher. The results of the IRR for the 18 items are presented as Table 4.9.

Table 4.9

Cohen's kappa (κ) with Confidence Intervals (CI) for 18 IRMAT Items

Factor	No.	Item	κ (95% CI)
1	3	Were the factors of interest (concepts, variables, etc.) clearly stated?	κ = -0.13 [-0.37, 0.11]
1	5	Was the problem stated unambiguously and was it easy to identify?	κ = 0.02 [-0.22, 0.26]
1	8	If applicable to the purpose and/or type of literature included, was the sampling frame clearly stated?	κ = 0.018 [-0.23, 0.26]
1	11	Were detailed descriptions of the database search processes provided?	κ = 0.16 [-0.078, 0.4]
1	12	Are reproducible line-by-line search strategies (or a sequence of terms for simpler interfaces) provided?	κ = 0.15 [-0.12, 0.42]
1	15	Were the inclusion/exclusion criteria for the eligible sources clearly stated?	κ = 0.59 [0.3, 0.89]
1	17	Were the search results screened for relevance using a pre-specified set of eligibility criteria?	κ = 0.15 [-0.0023, 0.31]
1	24	Was a systematic analytic method explicitly identified?	κ = 0.29 [0.061, 0.51]
1	34	Were methodological or other limitations of the review explicitly stated?	κ = 0.66 [0.45, 0.87]
2	21	Is the data extraction process explicit, unbiased, and reproducible?	κ = 0.33 [0.12, 0.55]
2	22	Were data analyzed using a review matrix?	κ = 0.086 [-0.0016, 0.17]
2	23	Is a data display assembling the data from retrieved sources provided?	κ = 0.55 [0.31, 0.79]
2	25	Were commonalities and differences identified?	κ = 0.0072 [-0.005, 0.019]
2	26	Was conflicting evidence addressed?	κ = -0.0065 [-0.095, 0.082]
2	27	Was the quality of retrieved sources addressed in a meaningful way?	κ = 0.53 [0.31, 0.75]

Factor	No.	Item	κ (95% CI)
2	28	Was quality verified by two independent reviewers?	κ = 0.44 [0.19, 0.69]
2	31	Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?	κ = 0.001 [-8.4, 8.4]
2	33	Were implications discussed for: research, practice, education, or policy?	κ = 0.46 [0.017, 0.9]

Eight of the 18 items (44%; items 5, 8, 11, 12, 17, 22, 25, and 31) fell within the range of slight agreement (0.00-0.20), with an additional five items (28%) – items 15, 23, 27, 28, and 33 – having moderate IRR (0.41-0.60). Two items (21 and 24) had fair agreement (0.21-0.40). There was only one item ("Were methodological or other limitations of the review explicitly stated?") with substantial agreement (κ = 0.66, 95% CI [0.45, 0.87]).

Items three ("*Were the factors of interest (concepts, variables, etc.) clearly stated?")* and 26 ("*Was conflicting evidence addressed?*") demonstrated poor IRR (< 0.00). Because of the insufficient IRR for these two items, an additional factor analysis was run excluding these two items to ascertain model fit and internal consistency reliability. While overall model fit of the refined 16-item model was comparable to the model fit of the 18-item version of the tool, Cronbach's α was decreased for factor two (*Analysis*) from *Respectable* (α = 0.719) under the 18-item model to *Minimally acceptable* (α = 0.693) under the 16-item model. Because the internal consistency reliability of factor two of the pruned 16-item model fell below acceptable standards (< 0.70) (DeVellis, 2017), it was determined that items three and 26 would remain in the final pilot version of the IRMAT even though the IRR of these two items was considered poor. The final items for the pilot version of the IRMAT are listed in Table 4.10. The final IRMAT items in the format of a user guide with instructions for using the tool, as well as descriptions and rating scales for all items, are provided as Appendix I.

Table 4.10

Final Version (1.3) of Pilot IRMAT

IRMAT Items
Were the factors of interest (concepts, variables, etc.) clearly stated?
Was the problem stated unambiguously and was it easy to identify?
Were the inclusion/exclusion criteria for the eligible sources clearly stated?
If applicable to the purpose and/or type of literature included, was the sampling frame clearly stated?
Were detailed descriptions of the database search processes provided?
Are reproducible line-by-line search strategies (or a sequence of terms for simpler interfaces) provided?
Were the search results screened for relevance using a pre-specified set of eligibility criteria?
Is the data extraction process explicit, unbiased, and reproducible?
Were data analyzed using a review matrix?
Is a data display assembling the data from retrieved sources provided?
Was a systematic analytic method explicitly identified?
Were commonalities and differences identified?
Was conflicting evidence addressed?
Was the quality of retrieved sources addressed in a meaningful way?
Was quality verified by two independent reviewers?
Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?
Were methodological or other limitations of the review explicitly stated?
Were implications discussed for: research, practice, education, or policy?

Summary of Chapter 4

The results of this study demonstrate that the pilot version of the Integrative Review Methodology Tool (IRMAT) demonstrates sufficient content (CVI > 0.80) and construct validity (overall goodness of model fit) and acceptable internal consistency reliability (F1 $\alpha = 0.741$; F2 $\alpha = 0.719$). The inter-rater reliability for 15 of the 18 IRMAT items (83%) was generally found to be between the ranges of slight (0.00-0.20) to moderate (0.41-0.60). The final version of the IRMAT

consists of 18 items, which were psychometrically tested and demonstrated to be valid and reliable in appraising reported methodological elements of nursing integrative reviews.

Chapter 5: Discussion

The outcome of this study is the development of the IRMAT to be used to appraise methodological and reporting quality of integrative reviews in the nursing literature. The final version of the tool consists of 18 items and demonstrates sufficient content (CVI = > 0.80) and construct validity (overall goodness of model fit), acceptable internal consistency reliability (F1 α = 0.741; F2 α = 0.719), and slight (0.00-0.20) to moderate (0.41-0.60) inter-rater reliability. General discussion of the study follows. The strengths and limitations of this study are discussed in detail. Implications for education, practice, and publishing are presented.

The statistical analysis of the overall model fit showed that the IRMAT version using the 6 factors of the Ellis model was not the best overall model fit, although the two-factor version of the tool demonstrating construct validity is still representative of overall thematic areas. The exploratory factor analysis (EFA) conducted as part of specific aim 2 revealed that the model was best fit to two factors, as opposed to the 6-factor model using the Ellis themes. The confirmatory factor analysis of this same 2-factor model derived from the EFA demonstrated better overall model fit when compared to the 6-factor model using Ellis (Model B, Table 4.7). However, it is worth noting that the final version of the 18-item IRMAT is representative of items which were originally aligned with all Ellis themes: starting, searching, differentiation, extracting, verifying, and ending.

It is important to note one item from the final pilot version of the IRMAT which is unique to the integrative review methodology. The item regarding the use of a sampling frame – *If applicable to the purpose and/or type of literature included, was the sampling frame clearly stated?* – is unique to the integrative review methodology. According to the methodological elements identified as part of the scoping literature review for this study (Chapter 2), a sampling frame for an integrative review may include (but is not limited to): the type of empirical studies, specific research design(s), or the inclusion of methodological or theoretical literature/framework, etc. Integrative review authors are encouraged to provide rationale for the

choice of sampling frame, which will strengthen the choice of using a sampling frame. Integrative review authors are also encouraged to be deliberate in the alignment of the choice of sampling frame with the integrative review methodology. For example, the choice to use a sampling frame which limits the retrieved sources to only quantitative results does not align with the choice to use an integrative review methodology, which implies inclusion of varying result types (empirical [quantitative, qualitative, mixed methods], methodological, or theoretical). Where other evidence synthesis methodologies may be more prescriptive in their guidance on what type of sources may be included in the synthesis, the integrative review methodology allows for broader inclusion of sources at the authors' discretion. In this respect, the optional use of a sampling frame to narrow this decision-making process is somewhat unique to the integrative review methodology and is thus reflected in this IRMAT item.

There were also items of note that were deleted as part of the development and testing of the IRMAT. The item addressing the use of a theoretical or conceptual framework – *Is a theoretical or conceptual framework used to guide the review?* – was removed due to a lack of face and content validity based on expert review. However, as one expert noted "This is an important element of an [integrative review] but often not addressed" (Appendix G). Whether or not the use of a conceptual or theoretical framework is an identifying characteristic of integrative reviews likely warrants further examination. Early in the testing of construct validity, the item *Is a flowchart of search results provided?* was removed for lack of covariance and correlation with other items. If future research is undertaken for the development of integrative review reporting guidelines, this item may make a good candidate for inclusion.

Two items that reflected the inclusion of gray or non-traditional, non-academic sources were excluded from the final version of the tool. The following items were removed during the pruning conducted during the exploratory factor analysis for not sufficiently loading onto any themes: *Were other recommended approaches to searching the literature used?* and, *If applicable to the purpose and type of literature included, was the potential for publication bias*

acknowledged?. During the review of content validity, several experts expressed concern over the difficulties with quality appraisal of gray literature sources, so the removal of these items also aligns with those concerns. However, the inclusion of gray literature is addressed (where applicable) in other evidence synthesis appraisal tools, such as the AMSTAR (Shea et al., 2017). Future versions of the IRMAT may benefit from modifying the description and rating scale for items addressing search components of the integrative review process to include references to gray or non-traditional, non-academic sources.

Strengths

There are several strengths of this study. First, the development and testing of an appraisal tool for nursing integrative reviews fills an important gap in the health sciences evidence synthesis literature. As noted in Chapter 1, there is a wealth of research surrounding evidence synthesis methodologies commonly used in medicine, such as systematic or scoping reviews. The research around integrative review methodologies commonly used in nursing is not as robust, and this study helps to fill that gap. Where tools like the AMSTAR for systematic reviews are already on their second versions, no appraisal tools for published integrative reviews existed prior to this research.

An additional strength of this study is the use of sound methodology in developing and testing the tool. This study mirrors documented methodologies that have been used successfully to develop appraisal tools in other evidence synthesis areas of the health sciences. To reflect the overall purpose of this tool as an appraisal tool, the incorporation of calculation of the inter-rater reliability of the tool strengthens this study.

Finally, perhaps the greatest strength of this study is that it may be used to further the integrative review methodology itself. The comprehensive scoping literature review (Chapter 2) that was completed as part of this study is the first of its kind. This research makes an important contribution to the integrative review methodology commonly used in nursing. With the recent publication of additional guidance on how to conduct integrative reviews, future methodologists

in this area may use this research to advance the methodology of integrative reviews in nursing and other disciplines.

Limitations

There are several limitations to this study. First, the study is somewhat limited by the use of only five experts in the initial item generation for the tool. While it would have been ideal to have approximately ten experts, the repeated invitations for expert participation only generated five completed responses. Although the survey responses for the expert review were blinded and anonymous, the primary researcher did have subsequent email correspondence with all five experts who completed the review of face and content validity. Four of the five expert participants are seminal experts in the field of nursing integrative reviews, with the fifth being the editor-in-chief of one of the most prominent nursing research scholarly journals. The area of integrative review methodology in the health sciences is not as robust as other evidence synthesis methodologies, such as systematic or scoping reviews. Although the total number of completed responses from expert reviewers was somewhat low for traditional tool development, the five expert responses received for this study is representative of all current known experts in nursing integrative reviews.

An additional limitation of this study are the portions completed individually by the primary researcher. Due to the nature of this study being completed as doctoral research, several portions that would have been completed by a team of researchers were completed solely by the primary researcher. To mitigate against this limitation, secondary reviewers were employed wherever possible to review the research output of items completed by the primary reviewer in isolation. As one example, the screening of methodological items for the scoping literature review (Chapter 2) was completed independently by the primary researcher. This was subsequently reviewed by two co-researchers (Miller and Smith), but the scoping review methodology recommends that this process is completed by two reviewers independently. Additionally, while the output of the coding of themes as part of the item generation process was

also reviewed by three researchers in total, the initial coding was completed by the primary researcher independently. Finally, the data analysis was verified by a statistician (Lim), but the analysis was primarily conducted independently by the doctoral student researcher.

Lastly, this study was limited somewhat in its scope. As was noted in Chapter 1, the overarching aim of this study was to develop and test an appraisal tool for nursing integrative reviews; however, there is also a need for the development and testing of reporting guidelines for integrative reviews like the reporting guidelines for systematic or scoping reviews. Because reporting guidelines are typically developed with larger-scale Delphi method studies, that was determined to be outside of the scope of this doctoral research. While it doesn't directly reflect on a limitation of this specific study, there is still a need for the development of integrative review reporting guidelines. As noted by select expert reviewers (Appendix G), there may be items that were ultimately included for inclusion in this appraisal tool that are better suited for reporting guidelines ("Were implications discussed for: research, practice, education, or policy?"). This requires further testing.

Future Research

There are several implications for future research following this study. While the pilot tool was deemed to have adequate internal consistency reliability, reliability and construct validity were computed using the same dataset of responses. Further testing of the reliability and validity using additional datasets is needed. Ideally, a large sample of responses could be collected from the same participants using the IRMAT on numerous articles, not just one. It is anticipated that a large pool of participants using the IRMAT on multiple articles which vary in quality would improve the issues seen in this research study with the skewness, kurtosis, and inter-item covariance. Additionally, the majority of the items demonstrated only slight to moderate inter-rater reliability (IRR). The most significant future research needed is for further testing of the IRR of the pilot tool items. Looking holistically at the outcomes of all aspects of this research, future research may benefit from revisiting the extracted methodological elements, as

well as the expert review comments, in further improving the items as a means of improving the IRR. When the primary reviewer was discussing disagreements with the two IRR raters, there was often confusion expressed at how to differentiate between items. Items such as “Were commonalities and differences identified?” and “Was conflicting evidence addressed?” as two specific examples were deemed to be very similar by the two PhD student IRR raters. A recommended future research approach is to revisit the integrative review methodology and improve the differentiation between items. As demonstrated in the expert comments in Appendix G, there was not always agreement amongst the experts as to the interpretation or meaning of the methodological elements and IRMAT items. A cognitive interviewing study to determine the most understandable version of item wording may be warranted. Alternatively, further IRR testing may be needed with different nursing populations. Where this research used PhD nursing research students, further research testing the IRR of the IRMAT items may benefit from using nursing journal editors or authors who have published numerous integrative reviews serving as independent raters.

As mentioned in Chapter 1 and the Chapter 5 *Limitations* section, in addition to this appraisal tool, there is also likely a need for development and testing of integrative review reporting guidelines. The development of reporting guidelines using a Delphi method approach would require significant funding and personnel resources. Fortunately, the response that the primary researcher has received from this study indicate that there is interest from the nursing research community in the area of integrative review methodology. It is likely that enough participants could be secured for a Delphi study to develop integrative review reporting guidelines. The annual conference of the International Academy of Nursing Editors (INANE) would be an ideal location to conduct portions of the Delphi study, as this conference is held biannually in international locations and membership is diverse and representative of the journals most frequently publishing nursing integrative reviews (Table 3.1).

Because this tool was designed to be specifically used for nursing integrative reviews, there are also future research needs for non-nursing disciplines. Cursory literature searches show that literature review papers incorporating “integrative review” or “integrated review” are beginning to appear more commonly in the psychology, sociology, and business literature. More comprehensive research into what specific methodologies these other fields are using is warranted. Appraisal tools for integrative reviews in non-nursing disciplines may be needed.

Finally, there are future research implications for translation of the pilot version of the IRMAT. During aim 2 when the purposive sample of nursing integrative review co-authors were being emailed to participate in the study, the primary researcher received several emails from Brazilian nursing researchers about pursuing a phase of research to develop a Portuguese version of the IRMAT. As is seen by the top nursing journals publishing integrative reviews (Table 3.1), there is a significant output of nursing integrative review research being published in Portuguese language nursing journals (*Revista Brasileira De Enfermagem*, *Revista Gaucha De Enfermagem*, *Revista Da Escola De Enfermagem Da Usp*, and *Revista Latino-Americana De Enfermagem*). The interest expressed from Brazilian nursing researchers, as well as the data demonstrating significant research output in these Portuguese nursing journals, suggest that a Portuguese language version of the IRMAT is a next logical step.

Implications for Education, Practice, and Publishing

Education

The IRMAT will be useful in a broad educational context. The IRMAT may help nursing students evaluate the quality of published integrative reviews. For nursing PhD or DNP students who are authoring integrative reviews as part of their coursework, the IRMAT could be used as a learning tool for students to learn about the components of a published integrative review. As these students are preparing for their own integrative review research endeavors, reading the IRMAT items, descriptions, and rating scales could prepare students for more authoritatively authoring integrative reviews for their coursework. Additionally, the IRMAT could be used

alongside other appraisal tools such as the AMSTAR for systematic reviews as a learning tool for students to examine the differences between different evidence synthesis methodologies.

Practice

Development of the IRMAT will advance practice knowledge in several ways. In the rapid pace of clinical decision-making, the integrative review is intended to provide a high-level, quickly consumable level of knowledge synthesis that "a nurse clinician or scholar may use ... in the development of clinical practice guidelines" (Toronto et al., 2018, p. 30) Through the design and testing of the IRMAT, readers of integrative reviews in the clinical practice setting can be equipped with an appraisal tool to quickly evaluate the methodological and reporting quality of nursing integrative reviews. The ability to appraise integrative reviews more quickly and accurately may lead to an increased use of higher quality evidence, which may in turn lead to higher quality practice and potentially improve clinical outcomes.

Publishing

The IRMAT may have implications for publishing. Editors and peer reviewers of nursing journals could use the IRMAT to appraise the quality of integrative review submissions before publication. While the IRMAT is an appraisal tool and not reporting guidelines and as such is not intended to be authoritative on what methodological items are reporting in publication, the IRMAT can be used to appraise whether or not the methodological considerations of the integrative review format were followed. Editors and peer reviewers as part of the publication submission process may elect to use the IRMAT as an appraisal tool when integrative reviews are submitted by authors for publication.

Conclusion

The Integrative Review Methodology Appraisal Tool (IRMAT) demonstrates evidence of validity and reliability. The IRMAT has implications for educating nursing students (particularly PhD and DNP students) and nurses in clinical practice on appraising integrative reviews in the nursing literature. The IRMAT may also be used by editors and peer reviewers to appraise

submitted integrative reviews prior to publication. While there remains a need for further research surrounding the integrative review methodology – both in nursing and non-nursing disciplines – the IRMAT fills a gap in the health sciences evidence synthesis literature.

References

American Association of Colleges of Nursing. (2019). *AACN's Vision for Academic Nursing*. https://www.aacnnursing.org/Portals/42/News/White-Papers/Vision-Academic-Nursing.pdf

Babyak, M. A., & Green, S. B. (2010). Confirmatory factor analysis: An introduction for psychosomatic medicine researchers. *Psychosomatic Medicine*, *72*(6), 587-597. https://doi.org/10.1097/PSY.0b013e3181de3f8a

Balk, E. M., Chung, M., Chen, M. L., Trikalinos, T. A., & Chang, L. (2013). *Assessing the accuracy of Google Translate to allow data extraction from trials published in non-English languages* (AHRQ Publication No. 12(13)-EHC145-EF). https://www.ncbi.nlm.nih.gov/books/NBK121305/

Beyea, S., & Nicoll, L. H. (1998). Writing an integrative review. *AORN Journal*, *67*(4), 877-880. https://doi.org/10.1016/S0001-2092(06)62653-7

Booth, A., & Carroll, C. (2015). How to build up the actionable knowledge base: The role of 'best fit' framework synthesis for studies of improvement in healthcare. *BMJ Quality & Safety*, *24*(11), 700-708. https://doi.org/10.1136/bmjqs-2014-003642

Bramer, W. M., Giustini, D., de Jonge, G. B., Holland, L., & Bekhuis, T. (2016). De-duplication of database search results for systematic reviews in EndNote. *Journal of the Medical Library Association : JMLA*, *104*(3), 240-243. https://doi.org/10.3163/1536-5050.104.3.014

Broome, M. E. (1993). Integrative literature reviews for the development of concepts. In B. L. Rodgers & K. A. Knafl (Eds.), *Concept development in nursing : Foundations, techniques, and applications* (pp. 193-215). Saunders.

Broome, M. E. (2000). Integrative literature reviews for the development of concepts. In B. L. Rodgers & K. A. Knafl (Eds.), *Concept development in nursing : Foundations, techniques, and applications* (2nd ed., pp. 231-250). Saunders.

Brown, S. J. (2012). *Evidence-based nursing : The research-practice connection* (2nd ed.). Jones & Bartlett Learning.

Bujang, M. A., & Baharum, N. (2017). Guidelines of the minimum sample size requirements for Cohen's Kappa. *Epidemiology Biostatistics and Public Health*, *14*(2), e12267-12261. https://doi.org/10.2427/12267

Cantor, A. B. (1996). Sample-size calculations for Cohen's Kappa. *Psychological Methods*, *1*(2), 150-153. https://doi.org/10.1037/1082-989X.1.2.150

Case, D. O., & Given, L. M. (2016). *Looking for information : A survey of research on information seeking, needs, and behavior* (4th ed.). Emerald Group Publishing Limited.

Christmals, C. D., & Gross, J. J. (2017). An integrative literature review framework for postgraduate nursing research reviews. *European Journal of Research in Medical Sciences*, *5*(1), 7-15.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Conner, B. T. (2014, January 11). *Demystifying literature reviews*. https://www.myamericannurse.com/demystifying-literature-reviews/

Cooper, H. M. (1984). *The integrative research review : A systematic approach*. Sage Publications.

Cooper, H. M. (1989). *Integrating research : A guide for literature reviews* (2nd ed.). Sage Publications.

Cooper, H. M. (1998). *Synthesizing research : A guide for literature reviews* (3rd ed.). Sage Publications.

Coughlan, M., Cronin, P., & Ryan, F. (2013). *Doing a literature review in nursing, health and social care*. Sage Publications.

Critical Appraisal Skills Programme. (2018). *CASP Systematic Review Checklist*. Critical Appraisal Skills Programme (CASP). https://casp-uk.b-cdn.net/wp-content/uploads/2018/03/CASP-Systematic-Review-Checklist-2018_fillable-form.pdf

Crossetti, M. d. G. O. (2012). Integrative review of nursing research: Scientific rigor required. *Revista Gaúcha de Enfermagem*, *33*(2), 12-13. https://doi.org/10.1590/S1983-14472012000200003

Crowe, M., & Sheppard, L. (2011). A general critical appraisal tool: An evaluation of construct validity. *International Journal of Nursing Studies*, *48*(12), 1505-1516. https://doi.org/10.1016/j.ijnurstu.2011.06.004

de Souza, M. T., da Silva, M. D., & de Carvalho, R. (2010). Integrative review: What is it? How to do it? *Einstein (Sao Paulo, Brazil)*, *8*(1), 102-106. https://doi.org/10.1590/s1679-45082010rw1134

DeVellis, R. F. (2017). *Scale development : Theory and applications* (4th ed.). Sage Publications.

Ellis, D. (1984a). The effectiveness of information retrieval systems: The need for improved explanatory frameworks. *Social Science Information Studies*, *4*(4), 261-272. https://doi.org/10.1016/0143-6236(84)90002-4

Ellis, D. (1984b). Theory and explanation in information retrieval research. *Journal of Information Science*, *8*(1), 25-38. https://doi.org/10.1177/016555158400800105

Ellis, D. (1993). A comparison of the information seeking patterns of researchers in the physical and social sciences. *Journal of Documentation*, *49*(4), 356-359. https://doi.org/10.1108/eb026919

Ellis, D. (2005). Ellis's Model of Information-Seeking Behavior. In K. E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of information behavior* (pp. 138-142). Information Today, Inc.

Evans, D. (2007). Overview of methods. In C. Webb & B. H. Roe (Eds.), *Reviewing research evidence for nursing practice : Systematic reviews* (pp. 137-148). Blackwell Publishing.

Flanagan, J. (2018). The integrative review. *International Journal of Nursing Knowledge*, *29*(2), 81. https://doi.org/10.1111/2047-3095.12208

Gamer, M., Lemon, J., Fellow, I., & Singh, P. (2019). *Package 'irr'*. https://cran.r-project.org/package=irr

Ganong, L. H. (1987). Integrative reviews of nursing research. *Research in Nursing & Health*, *10*(1), 1-11. https://doi.org/10.1002/nur.4770100103

Garrard, J. (2017). *Health sciences literature review made easy : The matrix method* (5th ed.). Jones & Bartlett Learning.

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

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. https://doi.org/https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377-381. https://doi.org/https://doi.org/10.1016/j.jbi.2008.08.010

Higgins, J. P. T., & Cochrane Collaboration. (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). Wiley-Blackwell.

Holly, C. (2014). *Scholarly inquiry and the DNP capstone*. Springer Publishing Company, LLC.

Holly, C. (2017). Other types of reviews: Rapid, scoping, integrated, and reviews of text and opinion. In C. Holly, S. W. Salmond, & M. Saimbert (Eds.), *Comprehensive systematic*

review for advanced practice nursing (2nd ed., pp. 321-335). Springer Publishing Company.

Holly, C. (2019). Integrative review. In *Practice-based scholarly inquiry and the DNP project* (2nd ed., pp. 171-183). Springer Publishing Company, LLC,.

Hong, Q. N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice*, *24*(3), 459-467. https://doi.org/10.1111/jep.12884

Hong, Q. N., Pluye, P., Fabregues, S., bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *Mixed Methods Appraisal Tool (MMAT), version 2018.* (Registration of Copyright (#1148552)). Industry Canada.

Hooper, D., Coughlan, J., & Mullen, M. (2007). Structural equation modeling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, *6*(1), 53-60.

Hopia, H., Latvala, E., & Liimatainen, L. (2016). Reviewing the methodology of an integrative review. *Scandinavian Journal Of Caring Sciences*, *30*(4), 662-669. https://doi.org/10.1111/scs.12327

International Academy of Nursing Editors. (2019). *Directory of Nursing Journals*. https://airtable.com/shrjqveaKHtS9xku8/tblNXTxmTr18CC1If

Jackson, G. B. (1980). Methods for Integrative Reviews. *Review of Educational Research*, *50*(3), 438-460. https://doi.org/10.3102/00346543050003438

Joanna Briggs Institute. (2020). *Checklist for Systematic Reviews*. https://jbi.global/sites/default/files/2021-03/Checklist_for_Systematic_Reviews_and_Research_Syntheses.docx

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401-415. https://doi.org/10.1007/bf02291817

Kean, E. B. (2021, March 24-27, 2021). Integrative review of integrative reviews: Quality of nursing literature reviews, 2005-2020. Midwest Nursing Research Society 45th Annual Research Virtual Conference.

Kirkevold, M. (1997). Integrative nursing research - An important strategy to further the development of nursing science and nursing practice. *Journal of Advanced Nursing*, *25*(5), 977-984. https://doi.org/10.1046/j.1365-2648.1997.1997025977.x

Kline, R. B. (2016). *Principles and practice of structural equation modeling, 4th ed*. Guilford Press.

Knafl, K., & Whittemore, R. (2017). Top 10 tips for undertaking synthesis research. *Research in Nursing & Health*, *40*(3), 189-193. https://doi.org/10.1002/nur.21790

Knafl, K. A., & Deatrick, J. D. (2000). Knowledge synthesis and concept development in nursing. In B. L. Rodgers & K. A. Knafl (Eds.), *Concept development in nursing : Foundations, techniques, and applications* (2nd ed., pp. 39-54). Saunders.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.

LoBiondo-Wood, G. (2014). Systematic reviews and clinical practice guidelines. In G. LoBiondo-Wood & J. Haber (Eds.), *Nursing research : Methods and critical appraisal for evidence-based practice* (8th ed., pp. 218-230). Elsevier.

Lobiondo-Wood, G. (2018). Systematic reviews and clinical practice guidelines. In G. LoBiondo-Wood & J. Haber (Eds.), *Nursing research : Methods and critical appraisal for evidence-based practice* (9th ed., pp. 199-211). Elsevier.

LoBiondo-Wood, G., & Haber, J. (2010a). Integrating the processes of research and evidence-based practice. In G. LoBiondo-Wood & J. Haber (Eds.), *Nursing research : Methods and critical appraisal for evidence-based practice* (7th ed., pp. 5-26). Mosby/Elsevier.

LoBiondo-Wood, G., & Haber, J. (2010b). Nonexperimental designs. In G. LoBiondo-Wood & J. Haber (Eds.), *Nursing research : Methods and critical appraisal for evidence-based practice* (7th ed., pp. 195-219). Mosby/Elsevier.

LoBiondo-Wood, G., & Haber, J. (2010c). *Nursing research : Methods and critical appraisal for evidence-based practice* (7th ed.). Mosby/Elsevier.

LoBiondo-Wood, G., & Haber, J. (2014a). Integrating research, evidence-based practice, and quality improvement processes. In G. LoBiondo-Wood & J. Haber (Eds.), *Nursing research : Methods and critical appraisal for evidence-based practice* (8th ed., pp. 5-24). Elsevier.

LoBiondo-Wood, G., & Haber, J. (2014b). *Nursing research : Methods and critical appraisal for evidence-based practice* (8th ed.). Elsevier.

LoBiondo-Wood, G., & Haber, J. (2018a). Integrating research, evidence-based practice, and quality improvement processes. In G. LoBiondo-Wood & J. Haber (Eds.), *Nursing research : Methods and critical appraisal for evidence-based practice* (9th ed., pp. 5-22). Elsevier.

LoBiondo-Wood, G., & Haber, J. (2018b). *Nursing research : Methods and critical appraisal for evidence-based practice* (9th ed.). Elsevier.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing research (New York)*, *35*(6), 382-385.

Maydeu-Olivares, A. (2017). Maximum likelihood estimation of structural equation models for continuous data: standard errors and goodness of fit. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(3), 383-394. https://doi.org/10.1080/10705511.2016.1269606

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica (Zagreb)*, *22*(3), 276-282.

Meho, L. I., & Tibbo, H. R. (2003). Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology*, *54*(6), 570-587. https://doi.org/10.1002/asi.10244

Mendes, K., Silveira, R. C., & Galvão, C. (2008). Revisão integrativa: Método de pesquisa para a incorporação de evidências na saúde e na enfermagem. *Texto & Contexto Enfermagem*, *17*(4), 758-764. https://doi.org/10.1590/S0104-07072008000400018

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis : A methods sourcebook* (3rd ed.).

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Journal of Clinical Epidemiology*, *62*(10), 1006-1012. https://doi.org/10.1016/j.jclinepi.2009.06.005

Moher, D., Schulz, K. F., Simera, I., & Altman, D. G. (2010). Guidance for developers of health research reporting guidelines. *PLoS Medicine*, *7*(2), e1000217. https://doi.org/10.1371/journal.pmed.1000217

National Library of Medicine. (2020). *PubMed nursing journals*. https://www.ncbi.nlm.nih.gov/nlmcatalog/?term=nursing%5Bst%5D+AND+currentlyindexed%5BAll%5D

Nursing and Allied Health Resources and Services. (2016). *NAHRS Selected List of Nursing Journals 2016* Medical Library Association. https://www.mlanet.org/d/do/5982

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Pace, R., Pluye, P., Bartlett, G., Macaulay, A., Salsberg, J., Jagosh, J., & Seller, R. (2010, 2010). *Reliability of a tool for appraising methodological quality of qualitative, quantitative and mixed methods studies.* North American Primary Care Research Group Annual Meetings, Seattle, WA.

Pace, R., Pluye, P., Bartlett, G., Macaulay, A. C., Salsberg, J., Jagosh, J., & Seller, R. (2012). Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. *International Journal of Nursing Studies*, *49*(1), 47-53. https://doi.org/10.1016/j.ijnurstu.2011.07.002

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology*, *134*, 178-189. https://doi.org/10.1016/j.jclinepi.2021.03.001

Peters, M. D. J., Godfrey, C., McInerney, P., Munn, Z., Tricco, A. C., & Khalil, H. (2020). Scoping reviews (2020 version). In E. Aromataris & Z. Munn (Eds.), *JBI Manual for Evidence Synthesis*. https://doi.org/https://doi.org/10.46658/JBIMES-20-12

Pituch, K. A., & Stevens, J. (2016). Exploratory factor analysis. In *Applied multivariate statistics for the social sciences : Analyses with SAS and IBM's SPSS* (6th ed., pp. 339-390). Routledge/Taylor & Francis Group.

Pluye, P., Gagnon, M. P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative,

quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies*, *46*(4), 529-546. https://doi.org/10.1016/j.ijnurstu.2009.01.009

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489-497. https://doi.org/10.1002/nur.20147

Polit, D. F., & Beck, C. T. (2017). *Nursing research : Generating and assessing evidence for nursing practice* (10th ed.). Lippincott Williams & Wilkins.

Raiche, G. M., D. (2022). *nFactors: An R package for parallel analysis and non graphical solutions to the Cattell scree test.* (Version 2.4.1.1) [Computer software]. Universite du Quebec a Montreal. https://cran.r-project.org/web/packages/nFactors/index.html

Revelle, W. (2023). *psych: Procedures for Psychological, Psychometric, and Personality Research.* In (Version 2.3.3) [Computer software]. Northwestern University, Evanston, Illinois. https://CRAN.R-project.org/package=psych

Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response analysis. *Journal of Statistical Software*, *17*(5), 1 - 25. https://doi.org/10.18637/jss.v017.i05

Rodgers, B. L., & Knafl, K. A. (1993). *Concept development in nursing : Foundations, techniques, and applications*. Saunders.

Rodgers, B. L., & Knafl, K. A. (2000). *Concept development in nursing : Foundations, techniques, and applications* (2nd ed.). Saunders.

Roman, A. R., & Friedlander, M. (1998). Revisão integrativa de pesquisa aplicada à enfermagem (Integrative research review applied to nursing). *Cogitare Enfermagem*, *3*(2), 109-112.

Rosseel, Y. (2011). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Russell, C. L. (2005). An overview of the integrative research review. *Progress in Transplantation*, *15*(1), 8-13.

Shea, B. J., Bouter, L. M., Peterson, J., Boers, M., Andersson, N., Ortiz, Z., Ramsay, T., Bai, A., Shukla, V. K., & Grimshaw, J. M. (2007). External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS ONE*, *2*(12), e1350. https://doi.org/10.1371/journal.pone.0001350

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7*, 10. https://doi.org/10.1186/1471-2288-7-10

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, j4008. https://doi.org/10.1136/bmj.j4008

Shuler, J. B. (2014). Integrative review. In C. Holly (Ed.), *Scholarly inquiry and the DNP capstone* (pp. 125-136). Springer Publishing Company, LLC.

Sim, J., & Wright, C. C. (2005). The Kappa Statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy*, *85*(3), 257-268. https://doi.org/10.1093/ptj/85.3.257

Soares, C. B., Hoga, L. A. K., Peduzzi, M., Sangaleti, C., Yonekura, T., & Silva, D. R. A. D. (2014). Integrative review: Concepts and methods used in nursing. *Revista da Escola de Enfermagem*, *48*(2), 335-345. https://doi.org/10.1590/S0080-6234201400002000020

Stevens, K. R. (2001). Systematic reviews: the heart of evidence-based practice. *AACN Clinical Issues*, *12*(4), 529-538.

Sutton, A., Clowes, M., Preston, L., & Booth, A. (2019). Meeting the review family: Exploring review types and associated information retrieval requirements. *Health Information and Libraries Journal*, *36*(3), 202-222. https://doi.org/10.1111/hir.12276

Thivant, E. (2005). Information seeking and use behaviour of economists and business analysts. *Information Research*, *10*(4).

Toronto, C. E., Quinn, B. L., & Remington, R. (2018). Characteristics of reviews published in nursing literature: A methodological review. *ANS Advances in Nursing Science*, *41*(1), 30-40. https://doi.org/10.1097/ANS.0000000000000180

Toronto, C. E., & Remington, R. (2020). *A Step-by-step guide to conducting an integrative review*. Springer International Publishing AG. https://doi.org/10.1007/978-3-030-37504-1

Torpey, E. (2018 April). *Employment outlook for bachelor's-level occupations*. U.S. Bureau of Labor Statistics. Retrieved 2020, January 26 from https://www.bls.gov/careeroutlook/2018/article/bachelors-degree-outlook.htm#Healthcare%20and%20science

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467. https://doi.org/10.7326/m18-0850

VanGeest, J., & Johnson, T. P. (2011). Surveying nurses: Identifying strategies to improve participation. *Evaluation & the Health Professions*, *34*(4), 487-511. https://doi.org/10.1177/0163278711399572

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

Whittemore, R. (2007). Rigour in integrative reviews. In C. Webb & B. H. Roe (Eds.), *Reviewing research evidence for nursing practice : Systematic reviews* (pp. 149-156). Blackwell Publishing.

Whittemore, R., Chao, A., Jang, M., Minges, K. E., & Park, C. (2014). Methods for knowledge synthesis: An overview. *Heart & Lung*, *43*(5), 453-461. https://doi.org/10.1016/j.hrtlng.2014.05.014

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, *52*(5), 546-553.

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, *55*(3), 249-270. https://doi.org/10.1108/EUM0000000007145

Appendix A: ***Integrative Review (IR) Definitions Extracted from Nursing Sources***

Source	Definition	Selected IR References Cited by Source Material
Beyea and Nicoll (1998)	"An integrative review summarizes past research and draws overall conclusions from the body of literature on a particular topic. The body of literature comprises all studies that address related or identical hypothesis." (p. 877)	Cooper (1984); Ganong (1987)
Broome in Rodgers and Knafl (1993)	"An integrative review is defined as one in which past research is summarized by drawing overall conclusions from many studies." (p. 195)	Cooper (1987); Cooper (1989); Ganong (1987); Glass (1976); Glass (1981); Glass (1978)
Broome in Rodgers and Knafl (2000)	"An integrative review is defined as one in which past research is summarized by drawing overall conclusions from many studies." (p. 234)	Cooper (1987); Cooper (1989); Cooper (1994); Ganong (1987); Glass (1976); Kirkevold (1997); Smith & Stullenbarger (1995)
Brown (2012)	"Integrative research reviews (IRRs) are also called state-of-the-science summaries, narrative reviews, and qualitative systematic reviews (to differentiate them from quantitative research reviews, that is, meta-analyses)." (p. 210)	None
Christmals and Gross (2017)	"An integrative literature review is a non-experimental design in which the researchers objectively critique, summarise and make conclusions about a subject matter through a systematic search, categorization and thematic analysis of past qualitative and quantitative research studies on the subject (LoBiondo-Wood & Haber, 2010; Sparbel & Anderson, 2000; Torraco, 2005). An integrative review, a comprehensive and inclusive methodology, permits the inclusion of both quantitative and qualitative research articles (Whittemore & Knafl, 2005)." (p. 7)	LoBiondo-Wood & Haber (2010); Sparbel & Anderson (2000); Torraco (2005); Whittemore & Knafl (2005)

Source	Definition	Selected IR References Cited by Source Material
Conner (2014)	"The term *integrative review* sometimes is used interchangeably with *systematic review.* Yet the two aren't identical. An integrative review includes data from experimental and nonexperimental research studies as well as theoretical literature. It may serve a wide range of purposes, such as defining concepts, reviewing theories, or analyzing methodological issues. Typically, it uses an orderly, specified methodology and process but may lack true quantification of results. Integrative reviews usually aren't found in the hierarchies of evidence but may be considered to provide a midlevel of evidence." (para. 7)	Whittemore & Knafl (2005)
Coughlan, Cronin, and Ryan (2013)	"According to Broome (2000), an integrative review is one which summarizes past research and draws conclusions on a given topic. In this definition the term research is interpreted in its broadest sense in literature that is sourced is not limited to empirical (primary) research studies. Theoretical or conceptual literature is also considered important (Whittemore and Knafl, 2005)." (p. 17)	Broome (2000); Whittemore & Knafl (2005); Torraco (2005)
Coughlan, Cronin, and Ryan (2013)	"Summarises and draws conclusions on past research on a given topic. Research is interpreted in its broadest sense and literature that is sourced can include primary research, and theoretical and conceptual literature." (p. 138)	Broome (2000); Whittemore & Knafl (2005); Torraco (2005)
Crossetti (2012)	"An integrative review summarizes results from previous studies, that is, those that have already been conducted and, most notably, shows conclusions of the body of literature on a particular phenomenon, and compassing, thus, all studies related to the question guiding the literature search." (p. 12)	Cooper (1984); Russel [sic] (2005); Whittemore & Knafl (2005); Ganong (1987); Broome (2000)
de Souza, da Sukvam, and de Carvalho (2010)	"integrative review is the methodology that provides synthesis of knowledge and applicability of results of significant studies into practice (1)." (p. 102)	Beyea & Nicoll (1998); Roman & Friedlander (1998); Whittemore (2005); Russell (2005); Whittemore & Knafl (2005)

Source	Definition	Selected IR References Cited by Source Material
de Souza, da Sukvam, and de Carvalho (2010)	"Although important, the review methods most often used – systematic review and meta-analysis – do not contemplate relevant nursing issues related to care and/or impact of disease or of treatment. Due to its methodological approach, the integrative review allows including diverse methods, which can potentially play the significant role in EBP in Nursing (3)." (p. 102)	Beyea & Nicoll (1998); Roman & Friedlander (1998); Whittemore (2005); Russell (2005); Whittemore & Knafl (2005)
Evans in Webb and Roe (2007)	"Integrative reviews provided much broader summary of the literature than is produced by systematic reviews, and have been described as the broadest category of research reviews, encompassing both empirical and theoretical literature (Whittemore, 2005). They commonly incorporate the findings from a range of different research designs. However, because their focus encompasses multiple methodological perspectives, the complexity of the review process is increased." (p. 137-8)	Whittemore (2005); Cooper (1984)
Flanagan (2018)	"Integrative reviews provide important syntheses of what is known about areas of concern. The integrative review is used widely in nursing because it is the broadest of reviews and typically includes case reports, the theoretical literature, and all types of qualitative and quantitative research." (p. 81)	None
Ganong (1987)	"An integrative review may be defined as one in which the reviewer is “primarily interested in inferring generalizations about substantial issues from a set of studies directly bearing on those issues” (Jackson, 1980, p. 438). Such reviews include examination of research support for competing hypothesis, suggestions for new theoretical issues, and identification of needed research." (p. 1)	Jackson (1980); Cooper (1982); Glass, McGaw, and Smith (1981); Jackson (1980)
Holly (2019)	"An integrative review is not a research method... An integrative literature review summarizes the main points of past research to draw general conclusions from a literary source on a certain topic using literature that includes studies that address related or identical hypotheses (https://guides.temple.edu/c.php?g=78618&p=4260131). As such, an integrative review appraises and combines available evidence, including theory." (p. 172)	Broome (2000); Cooper (1998); Russell (2005); Torraco (2005); Whittemore & Knafl (2005)

Source	Definition	Selected IR References Cited by Source Material
Holly in Holly, Salmond, and Saimbert (2017)	"An integrative literature review is a critical summary of the main points of a selected set of literature on a particular topic. An integrative review appraises and combines available evidence, including theory and other information found in books (Holly, 2013). The integrative review method can incorporate diverse methodologies to capture the circumstances, processes, and individual elements of the topic under study (Whittemore & Knafl, 2005)." (p. 327)	Broome (2000); Whittemore & Knafl (2005)
Hopia, Latvala, and Liimatainen (2016)	"The integrative method is defined as that which includes both empirical and theoretical publications (4). Therefore, the diversity of the sampling frame is the hallmark of this type of review (4, 5)." (p. 662)	Cooper (2010); Evans (2007); Whittemore (2007); Whittemore & Knafl (2005); Cooper (1998); Smith & Stullenbarger (1991); Whittemore (2005)
Kirkevold (1997)	"By integrative nursing research I refer to the collection, analysis and integration of separate research findings into meaningful wholes." (p. 977)	Cooper (1989)
Kirkevold (1997)	"The focus of this kind of research review, then, is the integration of findings from several empirical research studies pertaining to one particular topic. I shall call this empirical integrative research in order to clearly differentiate it from theoretical reviews." (p. 980)	Cooper (1989)
Kirkevold (1997)	"The purpose of theoretical reviews is to critique, refined or develop consistent theoretical accounts about a particular phenomenon or relationship. Within this approach, the analysis of existing research is on a more general level than the empirical integrative review." (p. 980)	Cooper (1989)
Knafl and Whittemore (2017)	"Synthesis of research or literature with heterogeneous designs. Can include theoretical articles." (p. 190)	Whittemore et al. (2014)
LoBiondo-Wood and Haber (2010)	"Integrative Research Review Synthesis review of the literature on a specific concept or topic." (p. 579)	None
LoBiondo-Wood and Haber (2014)	"Synthesis review of the literature on a specific concept or topic." (p. 578)	None

Source	Definition	Selected IR References Cited by Source Material
LoBiondo-Wood and Haber (2018)	"Synthesis review of the literature on a specific concept or topic." (p. 511)	None
LoBiondo-Wood and Haber in LoBiondo-Wood and Haber (2010)	"An integrative review is a focused review and synthesis of the literature on a specific area that follows specific steps of literature integration and synthesis without statistical analysis." (p. 9)	None
LoBiondo-Wood and Haber in LoBiondo-Wood and Haber (2010)	"An integrative review is the broadest category of review (Whittemore, 2005). It can include either theoretical or research literature or both. In integrative review does not include a statistical analysis. An integrative review may include methodology studies, a theory review, or the results of differing research studies with wide ranging clinical implications (Whittemore, 2005). An integrative review can include both quantitative or qualitative research period statistics are not used summarized and make conclusions about the studies." (p. 212)	Whittemore (2005)
LoBiondo-Wood and Haber in LoBiondo-Wood and Haber (2014)	"An integrative review is a focused review and synthesis of research or theoretical literature in a particular focus area, and includes specific steps of literature integration and synthesis without statistical analysis; it can include both quantitative and qualitative articles (Whittemore, 2005)." (p. 21)	Whittemore (2005)
LoBiondo-Wood and Haber in LoBiondo-Wood and Haber (2018)	"An integrative review is a focused review and synthesis of research or theoretical literature in a particular focus area, and includes specific steps of literature integration and synthesis without statistical analysis; it can include both quantitative and qualitative articles (Cochrane Consumer Network, 2016; Uman, 2011; Whittemore, 2005)." (p. 19)	Whittemore (2005)
LoBiondo-Wood in LoBiondo-Wood and Haber (2014)	“Integrative reviews critically appraised the literature in an area without a statistical analysis and are the broadest category of review (Whittemore, 2005; Whittemore & Knafl, 2005).” (p. 219)	Whittemore (2005); Whittemore & Knafl (2005)
LoBiondo-Wood in LoBiondo-Wood and Haber (2014)	"An integrative review is the broadest category of review (Whittemore, 2005; Whittemore & Knafl, 2005). It can include theoretical literature, research literature, or both. An integrative review may include methodology studies, a theory review, or the results of differing research studies with wide ranging clinical implications (Whittemore, 2005). An integrative review can include quantitative or qualitative research, or both. Statistics are not used summarized and make conclusions about the studies." (p. 225)	Whittemore (2005); Whittemore & Knafl (2005)

Source	Definition	Selected IR References Cited by Source Material
LoBiondo-Wood in LoBiondo-Wood and Haber (2018)	"Integrative reviews critically appraise the literature in an area but without a statistical analysis and are the broadest category of review (Whittemore, 2005; Whittemore & Knafl, 2005)." (p. 200)	Whittemore (2005); Whittemore & Knafl (2005)
LoBiondo-Wood in LoBiondo-Wood and Haber (2018)	"You will also find critical reviews of an area of research without a statistical analysis or a theory synthesis termed integrative reviews. An integrative review is the broadest category of review (Whittemore, 2005; Whittemore & Knafl, 2005). It can include theoretical literature, research literature, or both. An integrative review may include methodology studies, a theory review, or the results of differing research studies with wide-ranging clinical implications (Whittemore, 2005). An integrative review can include quantitative or qualitative research, or both. Statistics are not used to summarize and generate conclusions about the studies. " (p. 205)	Whittemore (2005); Whittemore & Knafl (2005)
Mendes, Silveira, and Galvao (2008)	"An analysis of studies important to a topic, which supports decision-making and the improvement of clinical practice. It permits general conclusions on a specific topic; allows the synthesis of various studies, including experimental and quasi-experimental research; is an ample analysis of the literature; permits the combination of data from theoretical and empirical literature; and can be used to define concepts, review theories, or for the methodological analysis of studies. The inclusion of differently designed studies increases the potential of IR to increase the depth and scope of the conclusions. IR is an incipient research method in Brazilian nursing." (p. 333)	Cooper (1984); Ganong (1987); Broome (1993); Stetler et al (1998); Beyea & Nicoll (1998); Roman & Friedlander (1998); Whittemore (2005); Whittemore & Knafl (2005)

Source	Definition	Selected IR References Cited by Source Material
Roman and Friedlander (1998)	"Integrative Research Review is a method that has the purpose of synthesizing results obtained in research on a delimited theme or issue, in a systematic and orderly, with the aim of contributing to knowledge of that theme or issue. According to COOPER (1982, 1989) it is a method that groups the results of primary research on the same subject in order to synthesize and analyze these data to develop a more comprehensive explanation of a specific phenomenon. KIRKEVOLD (1995) adds that the purpose of Integrative Research Review is about interconnecting elements isolated from existing studies. Claims that the Reviews of Research focuses on both empirical and theoretical frameworks developed in primary research. IT'S integrative because it provides more comprehensive information about a particular event, based on data taken from surveys without mandatory historical connotation." (p. 109)	Cooper (1982); Cooper (1989); Kirkvold [sic] (1995); Smith & Stullenbarger (1995); Ganong (1987)
Russell (2005)	“An integrative review of the literature is defined as one in which ‘past research is summarized by drawing overall conclusions from many studies.’ (p47) [Broome in Rodgers & Knafl, 2000]” (p. 8)	Cooper (1998); Broome (2000); Beyea & Nicoll (1998); Stevens (2001); Ganong (1987); Cooper (1989)
Shuler in Holly (2014)	"An integrative review is not a research method... An integrative literature review summarizes the main points of past research to draw general conclusions from a literary source on a certain topic using literature that includes studies that address related or identical hypotheses (unipapers.org/guide/integrative-literature-review). As such, an integrative review appraises and combines available evidence, including theory." (p. 125)	Broome (1993); Russell (2005); Torraco (2005); Whittemore & Knafl (2005)
Soares, Hoga, Peduzzi, Sangaleti, Yonekura, and Silva (2014)	"it is a type of review that can go beyond the analysis and synthesis of findings from primary studies allowing exploiting other research dimensions, and that presents potentialities for the development of new theories and new problems for research." (p. 329)	de Souza et al. (2010); Whittemore & Knafl (2005); Jackson (1980); Cooper (1982); Cooper (1984); Ganong (1987); Ellis (1991); Broome (1993); Kirkevold (1997); Stetler (1998);

Source	Definition	Selected IR References Cited by Source Material
		Beyea & Nicoll (1998); Roman & Friedlander (1998); Whittemore (2005); Torraco (2005); Russell (2005); Mendes & Silveira (2008); Pompeo et al. (2009); Carliner (2011); Cooper (1998)
Stevens (2001)	"Integrative review refers to a broader, sometimes less rigorous, method used to systematically combine results from a body of studies." (p. 530)	None
Toronto and Remington (2020)	"An integrative review looks more broadly at a phenomenon of interest than a systematic review and allows for diverse research, which may contain theoretical and methodological literature to address the aim of the review." (p. 2)	Cooper (1982); Cooper (1984); Ganong (1987); de Souza (2010); Jackson (1980); Russell (2005); Soares et al (2014); Torraco (2005); Torraco (2016); Whittemore & Knafl (2005)
Toronto, Quinn, and Remington (2018)	"Similar to the systematic review, an integrative review uses explicit and systematic methods to avoid incomplete searches and selection bias. Search strategies, inclusion and exclusion criteria, and results of searches are clearly outlined in integrative reviews." (p. 31)	Holly (2012); Whittemore et al (2014); Cooper (1998); Ganong (1987); Coughlin & Cronin (2017); Whittemore & Knafl (2005); Jackson (1980); Booth, Papaioannou, & Sutton (2012)

Source	Definition	Selected IR References Cited by Source Material
Whittemore (2005)	"Integrative reviews are the broadest category of research reviews and can encompass empirical or theoretical literature, or both, depending on the purpose of the research. Integrative reviews can be focused on methodology, theory, or the results of differing empirical studies with a wide range of implications, Table 2 (Broome, 1993). One of the distinct advantages to the integrative review approach is the ability to combine data from different types of research designs and include theoretical as well as empirical literature. Although the inclusion of multiple research designs can complicate the analysis, greater variety in the sampling frame has the potential to increase the depth and breadth of conclusions. The richness of the sampling frame also can contribute to a comprehensive portrayal of the topic of interest." (p. 57)	Jackson (1980); Cooper (1998); Broome (1993); Glass (1976)
Whittemore and Knafl (2005)	"An integrative review is a specific review method that summarizes past empirical or theoretical literature to provide a more comprehensive understanding of a particular phenomenon or healthcare problem (Broome 1993)." (p. 546)	Ganong (1987); Cooper (1998); Kirkevold (1997); Jackson (1980); Broome (1993)
Whittemore, Chao, Jang, Minges, and Park (2014)	"Integrative reviews are a category of knowledge synthesis methods that include a broad approach and sampling frame that can include empirical or theoretical literature, or both, depending on the purpose of the review.15 Integrative reviews can focus on methodology (e.g., evaluating the conceptual and operational definitions of a concept), theory (e.g., evaluating different behavior change theories), and/or research.70 Due to the methodological development of knowledge synthesis methods, the integrative review method is best suited for synthesizing knowledge on primary research combined with methodological and/or theoretical manuscripts. If all of the articles in a review are primary research, then a systematic review, meta-analysis, or mixed studies review method should be used." (p. 456, 458)	Whittemore & Knafl (2005); Whittemore (2005); Kirkevold (1997)
Whittemore in Webb and Roe (2007)	"The integrative review is a method that summarizes the literature on a clinical problem or phenomenon of concern that incorporates multiple perspectives and types of literature. The diversity of the sampling frame is the hallmark of this type of review period integrative reviews are the broadest type of review, and include experimental and non-experimental research as well as the theoretical literature if appropriate." (p. 149)	Whittemore (2005)

Appendix B: ***Extracted Methodological Elements Mapped to Ellis's Model***

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
de Souza, da Sukvam, and de Carvalho (2010)		104	including search in electronic databases,	X								
Toronto and Remington (2020)	2	12	Previous reviews may further the development of the topic and justify the need for proceeding with a new review, while at the same time helping to identify a gap in current knowledge.	X								
Toronto and Remington (2020)	1	4	include empirical or theoretical literature, or both (Cooper 1984)	X								
Toronto and Remington (2020)	1	4	provide synthesis on: (1) empirical research (review of quantitative and/or qualitative empirical studies on a particular topic), (2) methodological (review and analyses of designs and methodologies of different studies), and (3) theoretical (review of theories on a particular topic) (Whittemore et al. 2014; Soares et al. 2014)	X								
Toronto and Remington (2020)	1	5	broad purpose and review question(s) are clearly stated	X								

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	1	6	clearly identifying a problem from a gap in the literature	X								
Toronto and Remington (2020)	1	6	concepts of interest related to the research problem need to be clearly defined	X								
Toronto and Remington (2020)	1	6	background and significance for the research problem will provide justification for why the review is necessary	X								
Toronto and Remington (2020)	2	12	The findings from this preliminary background literature review could then support proceeding with the development of a review purpose or review questions to address the phenomenon of interest.	X								
Toronto and Remington (2020)	2	12	The introduction of the IR provides the background and rationale for conducting the review.	X								
Toronto and Remington (2020)	2	12	theoretical literature may help to define concepts	X								
Toronto and Remington (2020)	2	12	a review of the phenomenon of interest in the broadest context to provide the reader with a background to the topic	X								

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	2	12	Once the conceptual overview is complete, the introduction moves into the more specific description of how the phenomenon is to be addressed in the review, the setting or context, and the population being studied.	X								
Toronto and Remington (2020)	2	12	The purpose describes the goal of the review, or why the review is being conducted.	X								
Toronto and Remington (2020)	2	12	The review question(s) succinctly identifies what the review proposes to answer and suggests how it might contribute to a better understanding of the phenomenon of interest (Aveyard 2014).	X								
Toronto and Remington (2020)	2	13	identifies and describes the phenomenon to be investigated;	X								
Toronto and Remington (2020)	2	13	operationalizes the phenomenon, population, and context;	X								
Toronto and Remington (2020)	2	13	summarizes relevant research to present the gap in the literature that the review is to fill (Hudson-Barr 2004)	X								

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	2	13	The importance of previous work that was reported on the topic area should be discussed, and any unanswered questions and conflicting or challenging findings should be described (Hudson-Barr 2004).	X								
Toronto and Remington (2020)	2	14	clearly describing what is meant by the variables and how they will be used in the review	X								
Toronto and Remington (2020)	2	14	The conceptual and operational definitions of variables to be examined need to be developed.	X								
Toronto and Remington (2020)	2	16	The IR review question(s) allows one to explore issues relevant to nursing (Russell 2005).	X								
Brown (2012)	9	214	The process that was used to search for study reports is described in detail, including databases searched, key terms used, and any exclusion or exclusion criteria used.		X							
Christmals and Gross (2017)		12	Conduct literature search		X							
Crossetti (2012)		12	the literature search must be exhaustive		X							

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Crossetti (2012)		12	The list of articles will be complete once data have been saturated		X							
de Souza, da Sukvam, and de Carvalho (2010)		104	the search in databases should be broad and diverse …		X							
Evans in Webb and Roe (2007)	10	141	careful selection of databases and search terms, searching bibliography lists, hand searching relevant journals and searching conference abstracts		X							
Holly (2017)	14	328	The search strategy should encompass a minimum of two databases, although the search can be more extensive.		X							
Holly (2017)	14	329	For each database searched, a log should be kept that details the database searched, the range of dates searched, and the results of the search in terms of both the number of “hits” obtained and the number of relevant articles or other material retrieved.		X							
Holly (2019)	10	175	The search should be broad and encompass a minimum of two databases, although		X							

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			a search of three to six databases is preferred.									
Hopia, Latvala, and Liimatainen (2016)		667	Therefore, to find both theoretical and empirical literature for a given topic, search strategies should involve a combination of electronic database searches and grey literature hand searches.		X							
Shuler in Holly (2014)	8	128	The search should be broad and encompass a minimum of two databases, although a search of three to six databases is preferred.		X							
Toronto and Remington (2020)	1	5	comprehensive and replicable search strategy		X							
Toronto and Remington (2020)	3	22	include a clearly documented and comprehensive literature search, defining in detail all databases, search terms, limiters, eligibility (inclusion/exclusion), and criteria used, and describing any additional search methods (Cooper 1982; Whittemore and Knafl 2005)		X							

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	3	23	A narrative description of all information sources, including databases, that were used; limiters used to narrow search results, such as year of publication, language, and publication status; and search terms used.		X							
Toronto and Remington (2020)	3	33	All search history (searches, search terms, results from those searches, and article citations) should be saved—even search results that may be excluded later. This information will be combined into a final reporting format, such as the PRISMA Flow Diagram (2015) or other type of search flow diagram.		X							
Toronto and Remington (2020)	3	34	Search limiters, filters, and Boolean operators have been used correctly.		X							
Toronto and Remington (2020)	3	24	includes a comprehensive search… [that] includes multiple strategies (Whittemore and Knafl 2005)		X							
Whittemore and Knafl (2005)		548	Ideally, all of the relevant literature on the problem or topic		X							

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			of interest is included in the review;									
Whittemore and Knafl (2005)		548	other recommended approaches to searching the literature include ancestry searching, journal hand searching, networking, and searching research registries (Conn et al. 2003b)		X							
Whittemore and Knafl (2005)		548	addressing the issue of publication bias may also be relevant to the literature search stage (Conn et al. 2003b, Soeken & Sripusanapan 2003)		X							
Whittemore and Knafl (2005)		548	a comprehensive search for an integrative review identifies the maximum number of eligible primary sources, using at least two to three strategies (Jadad et al. 1998, Conn et al. 2003b)		X							
Whittemore (2005)		61	The search protocol and the eligibility criteria for primary studies are clearly stated and the results of the search are comprehensive (using 2–3 search strategies).		X							

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Whittemore (2005)		58	includes all of the relevant literature on the specified problem		X							
Whittemore (2005)		58	Ancestry searching, journal hand-searching, networking, and searching research registries are approaches also recommended for searching the literature (Conn, Isaramalai, et al., 2003).		X							
Whittemore (2005)		59	A comprehensive search for a research review identifies the maximum number of eligible studies, utilizing as many search strategies as resources allow (at least 2–3 strategies; Jadad et al., 1998).		X							
Whittemore in Webb and Roe (2007)	11	151	Often multiple strategies such as computerized database searching, ancestry searching, journal hand searching, networking and or searching research registries are required to obtain a representative sample (Conn et al. 2003a).		X							
Brown (2012)	9	213	Reviewers typically searched several health care databases using a			X						

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			variety of search terms, combinations of search terms, and search options.									
de Souza, da Sukvam, and de Carvalho (2010)		104	manual search in journals			X						
Holly (2019)	10	175	The strategy developed should include an electronic database search, a hand search in journals relevant to the guiding questions …			X						
Holly (2019)	10	175	A search for grey or unpublished literature can be considered, but it is not necessary.			X						
Mendes, Silveira, and Galvao (2008)		761	the search begins in the databases to identify the studies that will be included in the review			X						
Shuler in Holly (2014)	8	128	The strategy developed should include an electronic database search, a hand search in journals relevant to the guiding questions …			X						
Shuler in Holly (2014)	8	128	A search for grey or unpublished literature can be considered.			X						
Toronto and Remington (2020)	1	6	two or more methods for literature search			X						

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	3	24	includes a comprehensive search, one that captures as much literature pertaining to the topic as possible (Evans 2007; Whittemore 2007)			X						
Toronto and Remington (2020)	3	24	should utilize more than one or two search terms or phrases			X						
Toronto and Remington (2020)	3	29	Both natural language and controlled language searches are important to include in the search strategy			X						
Toronto and Remington (2020)	3	34	The search concepts are clear, not too narrow or too broad.			X						
Toronto and Remington (2020)	3	34	All spelling variants and synonyms have been searched, including abbreviations.			X						
Toronto and Remington (2020)	3	34	Appropriate subject headings, or controlled language terms, have been identified and searched.			X						
Brown (2012)	9	213	Thus, retrieval of eligible studies from database is only a starting point. The panel should go on to peruse reference list, go to research registries, …				X					

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
de Souza, da Sukvam, and de Carvalho (2010)		104	the references described in the selected studies				X					
Holly (2017)	14	329	An examination of the reference list of those articles selected for the review is also helpful in identifying additional resources for the review.				X					
Holly (2019)	10	175	The strategy developed should include … a search of the reference list of studies selected to be included in the review.				X					
Shuler in Holly (2014)	8	128	The strategy developed should include … a search of the reference list of studies selected to be in the review.				X					
Toronto and Remington (2020)	3	34	The search strategy has been modified by adding terms based on citations highly relevant to the topic.				X					
Brown (2012)	9	213	Thus, retrieval of eligible studies from database is only a starting point. The panel should go on to ... contact colleagues, and even run searches using Web search engines.					X				
de Souza, da		104	contact with researchers and the					X				

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Sukvam, and de Carvalho (2010)			use of unpublished material(2).									
Holly (2019)	10	175	Authors who have previously published on the topic for review can be contacted to determine whether any other sources are available or if they have published similar papers.					X				
Shuler in Holly (2014)	8	128	Also, authors who have previously published on the topic for review can be contacted to determine whether any other sources are available.					X				
Toronto and Remington (2020)	3	34	New searches return no new, unique, and relevant results.					X				
Toronto and Remington (2020)	3	34	Author searches on the most prolific authors of the topic show no new citations.					X				
Toronto and Remington (2020)	3	34	All databases likely to contain the highest number of citations have been searched.					X				
Brown (2012)	9	212	review panels decide how they will handle studies that are of dubious methodological quality.						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Brown (2012)	9	213	Using a pre specified set of inclusion criteria, two persons (ideally) decide which studies qualify for full review.						X			
Brown (2012)	9	213	The panel will then sort the articles into stacks by subtopics						X			
Brown (2012)	9	214	the methods used to evaluate the quality of the studies [is] also described.						X			
Christmals and Gross (2017)		12	Adopt a data collection tool						X			
Christmals and Gross (2017)		12	Set rules of inference for data analysis and interpretation						X			
de Souza, da Sukvam, and de Carvalho (2010)		104	The ideal procedure is to include all the studies found or a randomized selection of them;						X			
de Souza, da Sukvam, and de Carvalho (2010)		104	however, if both choices are not feasible due to the amount of works, the inclusion and exclusion criteria adopted for the articles must be clearly explained and discussed(8).						X			
de Souza, da Sukvam, and de Carvalho (2010)		105	the studies should be divided into subgroups according to a previously established classification,						X			
Evans in Webb and Roe (2007)	10	140	the process of selection … should be guided by clearly						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			stated criteria that delineate the area of interest									
Evans in Webb and Roe (2007)	10	143	The inclusion of multiple research designs means that criteria are needed for each type of research, which serves to increase the complexity of the appraisal process.						X			
Holly (2017)	14	329	A critical assessment of papers to be included in an integrated review may be done by one person…						X			
Holly (2019)	10	176	The search should involve seeking primary, secondary, and tertiary literature.						X			
Holly (2019)	10	177	each of the selected studies needs to be critically appraised						X			
Holly (2019)	10	177	Studies not meeting the appraisal criteria are excluded.						X			
Mendes, Silveira, and Galvao (2008)		762	This procedure for inclusion and exclusion of articles must be conducted in a judicious and transparent manner…						X			
Mendes, Silveira, and Galvao (2008)		762	The level of evidence of the studies must be evaluated						X			
Russell (2005)		11	establishing inclusion criteria for the studies						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Shuler in Holly (2014)	8	129	The search should involve seeking primary, secondary, and tertiary literature.						X			
Shuler in Holly (2014)	8	129	each of the selected studies needs to be critically appraised						X			
Shuler in Holly (2014)	8	129	Studies not meeting the appraisal criteria are excluded.						X			
Toronto and Remington (2020)	1	5	uses diverse data sources						X			
Toronto and Remington (2020)	1	5	methodological quality and relevance of selected literature are appraised						X			
Toronto and Remington (2020)	1	7	effective inclusion and exclusion criteria						X			
Toronto and Remington (2020)	1	7	screening, which involves reviewing the citations resulting from a search and selecting those deemed relevant for full-text retrieval						X			
Toronto and Remington (2020)	1	7	critical appraisal of the retrieved studies						X			
Toronto and Remington (2020)	2	17	All decisions about literature to include or exclude should be justified and documented in the methods to demonstrate that an unbiased process was followed.						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	2	17	Types of studies or literature.						X			
Toronto and Remington (2020)	2	17	The phenomenon under investigation.						X			
Toronto and Remington (2020)	2	17	The characteristics of the population being studied.						X			
Toronto and Remington (2020)	2	17	Publication language.						X			
Toronto and Remington (2020)	2	17	Time period covered by the review and its justification.						X			
Toronto and Remington (2020)	2	17	Setting (Garrard 2014; Stern et al. 2014).						X			
Toronto and Remington (2020)	2	17-8	If the studies are to be excluded for methodological quality, this should be stated with a clear description of the measures used to determine acceptable methodological quality.						X			
Toronto and Remington (2020)	2	18	the inclusion and exclusion criteria are aligned with the review purpose and/or question(s)						X			
Toronto and Remington (2020)	3	34	Database search results … are next reviewed to determine which						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			studies will be included in the review sample.									
Toronto and Remington (2020)	3	36	The most important reason to consider searching beyond databases is minimize publication bias. This type of bias is defined as a situation that leads to a report not being published due to the nature of its results (Russell 2005).						X			
Toronto and Remington (2020)	3	41	screening by relevance						X			
Toronto and Remington (2020)	3	41	then selecting by full text						X			
Toronto and Remington (2020)	3	41, 42	finally sorting into studies ("Instead of counting those as two studies, it would be counted as one study, using both of the reports to collect data about the project. Sometimes one report is multiple studies, which should be treated separately.")						X			
Toronto and Remington (2020)	4	45	the inclusion and exclusion criteria should identify whether inferior studies will be included or excluded after the appraisal process						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	4	46	it is essential that all evidence be assessed for quality before inclusion in the IR						X			
Toronto and Remington (2020)	4	52	it is important that [gray literature] be critically appraised before inclusion in the IR, as the quality of gray literature is variable						X			
Toronto and Remington (2020)	4	52	Unpublished research studies, such as dissertations, should be appraised using the appropriate critical appraisal tool as would be used for a published study.						X			
Toronto and Remington (2020)	4	52	Other textual papers should be assessed for aspects such as accuracy, objectivity, authority, evidence, and significance.						X			
Toronto and Remington (2020)	4	46	Reviewers should make judgments about the methodological strengths and weaknesses of all included studies before making inferences about the phenomenon of interest (Jackson 1980)						X			
Whittemore and Knafl (2005)		550	The primary sources included in the integrative review need to be divided into subgroups						X			

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			according to some logical system									
Whittemore and Knafl (2005)		548	purposive sampling can be combined with a comprehensive search						X			
Whittemore and Knafl (2005)		549	any sampling decision must be justified and made explicit						X			
Whittemore and Knafl (2005)		550	how quality is evaluated in an integrative review will vary depending on the sampling frame						X			
Whittemore and Knafl (2005)		550	consideration of the quality of primary sources in an integrative review is addressed in a meaningful way						X			
Whittemore in Webb and Roe (2007)	11	151	any sampling decision must be methodologically justified and made explicit						X			
Whittemore in Webb and Roe (2007)	11	151	Data collection in integrative reviews requires a standard and thorough examination of each primary source (Cooper, 1998).						X			
Brown (2012)	9	213	Panel members read each report and extract basic information about design, sample, variables, and results. They then enter the information into tables and							X		

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			create list to help them identify differences, commonality's, and patterns across the studies.									
Brown (2012)	9	214	The process used to extract information from the reports [is] also described.							X		
Christmals and Gross (2017)		12	Revise data collection tool to fit review purpose							X		
Christmals and Gross (2017)		12	Extract relevant information from included articles							X		
de Souza, da Sukvam, and de Carvalho (2010)		104	To extract data from the articles selected, it is necessary to use a previously prepared instrument							X		
de Souza, da Sukvam, and de Carvalho (2010)		104	The data should include definition of the subjects, methodology, size of the sample, measuring variables, method of analysis and concepts used as bases							X		
de Souza, da Sukvam, and de Carvalho (2010)		105	data are extracted from the primary sources using the prepared instrument to simplify, summarize and organize the findings,							X		
Evans in Webb and Roe (2007)	10	143	The aim of the analysis phase is to ensure complete documentation of all							X		

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			relevant data from each included study.									
Holly (2017)	14	329	A Table of Inclusion should be developed that includes information regarding the paper (author, year, title, journal), purpose of the study, study design, description of the sample, data analysis, major findings, and conclusions.							X		
Holly (2019)	10	178	When the critical appraisal has been completed, a table can be created that contains information about each of the studies in the review.							X		
Mendes, Silveira, and Galvao (2008)		762	defining the information to be extracted from the selected studies, using an instrument to gather and synthesize key information. 12 [Beyea & Nicoll, 1998]							X		
Mendes, Silveira, and Galvao (2008)		762	organize and summarize the information in a concise manner, forming a database that is easy to access and manage							X		

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Mendes, Silveira, and Galvao (2008)		762	the information should cover the study sample (subjects), the objectives, the methodology used, results and the main conclusions of each study.14 [Broome, 2000]							X		
Russell (2005)		11	a data collection tool should be developed							X		
Shuler in Holly (2014)	8	130	When the critical appraisal has been completed, a table can be created that contains information about each of the studies in the review.							X		
Toronto and Remington (2020)	5	64	selecting, focusing, simplifying, and abstracting data from the sample of primary sources (Whittemore and Knafl 2005; Miles and Huberman 1994a)							X		
Toronto and Remington (2020)	1	5	data abstraction							X		
Toronto and Remington (2020)	1	7	extract data into matrices (tables)							X		
Toronto and Remington (2020)	4	52	Notation of the results of the appraisal of each study should be entered into the matrix to support the credibility of the data							X		

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			analysis and findings of the review.									
Toronto and Remington (2020)	5	58	creation of a review matrix							X		
Toronto and Remington (2020)	5	60	more than one reviewer abstract data and to do so independently							X		
Toronto and Remington (2020)	5	60	order, code, and categorize data from multiple sources that may have used diverse methodological perspectives (Whittemore and Knafl 2005; Cooper 1998)							X		
Toronto and Remington (2020)	5	64	The primary sources in the sample will be initially reduced into subcategories (Whittemore and Knafl 2005).							X		
Whittemore and Knafl (2005)		550	the data from primary sources are ordered, coded, categorized, and summarized into a unified and integrated conclusion about the research problem (Cooper 1998)							X		
Whittemore and Knafl (2005)		550	Predetermined and relevant data of each subgroup classification are extracted from all primary data sources							X		

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			and compiled into a matrix or spreadsheet (Miles & Huberman 1994, Garrard 2004).									
Whittemore and Knafl (2005)		551	converting the extracted data from individual sources into a display that assembles the data from multiple primary sources around particular variables or subgroups							X		
Whittemore and Knafl (2005)		551	data displays can be in the form of matrices, graphs, charts, or networks							X		
Whittemore and Knafl (2005)		551	different data displays are likely to be required for each subgroup classification							X		
Whittemore (2005)		59	common data are extracted from primary studies for subsequent analysis							X		
Whittemore (2005)		59	methodological features of primary studies are also extracted to evaluate the overall quality of individual studies							X		
Whittemore in Webb and Roe (2007)	11	151	Pre-determined relevant data from each primary source need to be extracted, preferably using tables, charts or matrices (Miles & Huberman, 1994; Garrarand [sic], 1999).							X		

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Whittemore in Webb and Roe (2007)	11	151-2	Data extraction forms are ideally piloted with several primary sources and revised to ensure an accurate portrayal of all primary sources.							X		
Brown (2012)	9	214	The panel or authors indicate whether and how their conclusions square with any prior summarization work that has been done on the topic, summarized the limitations of the body of research, and offer opinions regarding the ramifications of the conclusions.								X	
Brown (2012)	9	213	Similarities and differences in results and findings would be identified and reasons for the differences explored.								X	
Christmals and Gross (2017)		12	Systematically analyse data								X	
de Souza, da Sukvam, and de Carvalho (2010)		104-5	In this stage, based on interpretation and synthesis of results, the data demonstrated in the analysis of the articles are compared to the theoretical reference.								X	
de Souza, da Sukvam, and de		104	this [critical analysis] phase demands an organized approach to weigh rigor and								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Carvalho (2010)			characteristics of each study									
de Souza, da Sukvam, and de Carvalho (2010)		105	identification of patterns, differences and redistribution of these topics may be included as part of the general discussion.								X	
Holly (2019)	10	178-9	As data are conceptualized at higher and higher levels of abstraction, each primary source is re-read and reviewed to verify that the new conceptualization was congruent with primary sources.								X	
Holly (2019)	10	179	The overall process involves data reduction, data display, data comparison, drawing conclusion, and verification (Whittemore & Knafl, 2005).								X	
Mendes, Silveira, and Galvao (2008)		762	the selected studies must be analyzed in detail								X	
Mendes, Silveira, and Galvao (2008)		762	The analysis must be performed critically, looking for explanations for the different or conflicting results in different studies.12,16 [Beyea & Nicoll, 1998; Ganong, 1987]								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Russell (2005)		11	The reviewer must identify which studies were not able to be included, describe why they were not included, and then discuss how they might have shown different results from those selected.								X	
Russell (2005)		12	critically judges whether the data element or result is worthy of remaining in the study data set								X	
Russell (2005)		12	The reviewer evaluates the data for unreliable values, which might include whether the findings from one study or two different from the other studies to be considered, and whether data recording errors exist.								X	
Russell (2005)		12	The reviewer also must evaluate the reliability of each study's findings, both in and of itself and in comparison with all of the other studies included in the review.								X	
Shuler in Holly (2014)	8	131	As data are conceptualized at higher and higher levels of abstraction, each primary source is re-read and reviewed to verify that the new								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			conceptualization was congruent with primary sources.									
Toronto and Remington (2020)	5	64	The constant comparison method consists of four phases: data reduction, data display, data comparison, and conclusion drawing and verification (Whittemore and Knafl 2005; Miles and Huberman 1994a)								X	
Toronto and Remington (2020)	1	5	comparison								X	
Toronto and Remington (2020)	1	5	synthesis								X	
Toronto and Remington (2020)	1	7	analyze for similarities and differences (patterns) in relation to the stated review purpose or questions								X	
Toronto and Remington (2020)	1	7	patterns are then synthesized								X	
Toronto and Remington (2020)	1	8	Comparisons and contrasts are made of the findings of the review with background literature, and work of others								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Toronto and Remington (2020)	5	65	examining the data displays for patterns, themes, commonalities, and differences across the review sample (Whittemore and Knafl 2005)								X	
Toronto and Remington (2020)	4	46	The method of assessing the risk of bias should be transparent and reproducible.								X	
Whittemore and Knafl (2005)		551	Conclusions or conceptual models that are developed are continually revised in order to be inclusive of as much data as possible (Miles & Huberman 1994).								X	
Whittemore and Knafl (2005)		551	patterns, themes, relationships, or conclusions requires verification with primary source data for accuracy and confirmability (Miles & Huberman 1994)								X	
Whittemore and Knafl (2005)		551	addressing conflicting evidence								X	
Whittemore and Knafl (2005)		551	exploration of confounding influences								X	
Whittemore and Knafl (2005)		551	data comparison which involves an iterative process of examining data displays of primary source data in order to identify								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			patterns, themes, or relationships									
Whittemore and Knafl (2005)		551	patterns and processes are isolated								X	
Whittemore and Knafl (2005)		551	commonalities and differences are identified								X	
Whittemore (2005)		60	Conclusions must have clear supportive evidence from the findings of all individual studies (Moher, Jadad, & Klassen, 1998).								X	
Whittemore (2005)		59	Quality is determined by evaluating the internal validity of primary studies or “the extent to which study design, conduct, and analysis systematically avoid or minimize potential sources of bias” (Conn & Rantz, 2003, p. 323).								X	
Whittemore (2005)		60	Coded data of each individual study are compared with all other studies for similarities and differences around variables of interest.								X	
Whittemore (2005)		60	Outliers are identified and examined for reasons for variability explored								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Whittemore (2005)		61	The quality of primary studies should be evaluated and incorporated in the analysis and interpretation of findings (Conn & Rantz, 2003).								X	
Whittemore in Webb and Roe (2007)	11	152	consideration of the quality of primary sources in an integrative review is addressed in a meaningful way								X	
Whittemore in Webb and Roe (2007)	11	152	Whittemore and Knafl (2005) have specifically addressed the complexity of data analysis and integrative reviews and have proposed that the qualitative method developed by Miles & Huberman (1994) assists in obtaining thorough, integrated inaccurate conclusions. The method consists of data reduction, data display, dated comparison, conclusion drawing and verification (Table 11.2) (Miles & Huberman, 1994).								X	
Whittemore in Webb and Roe (2007)	11	152	Data analysis procedures also need to include the opportunity for the analyst to identify conflicting evidence								X	

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			in tenuous results, as well as to consider several different explanations for interpreting the data.									
Brown (2012)	9	214	The goal of analysis is to reach conclusions that represent the findings of the individual studies as a body of findings, which is different from looking at each one in isolation from the others.									X
Christmals and Gross (2017)		12	Discuss and interpret data									X
Holly (2017)	14	329	Data for an integrative review are generally analyzed qualitatively by using a thematic synthesis approach, which arranges studies into common areas or themes that can be summarized.									X
Holly (2019)	10	179	Drawing conclusions and summarizing results is the final step in an integrative review, and is generally done as a narrative.									X
Holly (2019)	10	179	Similarities and differences in the findings are identified and a description of generalizations representative of the									X

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
			defined categories and the integrative review in its entirety, as possible (Whittemore & Knafl , 2005).									
Knafl & Whittemore (2017)		192	it often is desirable to update your database search as you near completion of the project to enhance the currency of the review.									X
Russell (2005)		12	“reducing the separate data points collected by The Enquirer into a unified statement about the research problem.” [Cooper]									X
Shuler in Holly (2014)	8	131	Drawing conclusions and summarizing results is the final step in an integrative review, and is generally done as a narrative.									X
Shuler in Holly (2014)	8	131	Similarities and differences in the findings are identified with a description of generalizations representative of the defined categories and the integrative review in its entirety, as possible (Whittemore & Knafl , 2005).									X

Source	Ch.	Pg.	Element	Starting	Searching	Browsing	Chaining	Monitoring	Differentiating	Extracting	Verifying	Ending
Whittemore (2005)		60	Data analysis in research reviews requires that the investigator order, categorize, and summarize data from primary individual studies into a unified conclusion about the research problem (Cooper, 1998).									X
Whittemore (2005)		60	Visual aids such as tables or schematic representations of the relationships among codes, concepts, and ideas often are employed (Miles & Huberman, 1994).									X
Whittemore (2005)		61	Studies ideally are assessed for their ability to be combined logically or statistically (heterogeneity tests), or both ways.									X

Appendix C: *R Syntax for Data Analyses*

```
#####################################################################
#                    DATA FILES                     #
#####################################################################

###28 items; no missing values, n=160###
data160 <- read.csv("Data_160.csv",header=TRUE)

###25 items - excludes 18,10,6; no missing values, n=160###
data160pruned <- read.csv("Data_160pruned.csv",header=TRUE)

###23 items - excludes 18,10,6,13,19; no missing values, n=183###
data183 <-read.csv("Data_183.csv",header=TRUE)

###22 items - excludes 18,10,6,13,19,2; no missing values, n=183###
data183pruned <-read.csv("Data_183pruned.csv",header=TRUE)

###19 items - excludes 18,10,6,13,19,2,14,20,4; no missing values, n=198###
data198 <-read.csv("Data_198.csv",header=TRUE)

###18 items - excludes 18,10,6,13,19,2,14,20,4,7; no missing values, n=198###
dataPruned <-read.csv("Data_198pruned.csv",header=TRUE)

###16 items - excludes 18,10,6,13,19,2,14,20,4,7,3,26; no missing values, n=198###
dataPrunedB <-read.csv("Data_198prunedB.csv",header=TRUE)
```

```
F1orig <-read.csv("F1.csv",header=TRUE)

F2orig <-read.csv("F2.csv",header=TRUE)


F1pruned <-read.csv("F1pruned.csv",header=TRUE)

F2pruned <-read.csv("F2pruned.csv",header=TRUE)



### Required packages ----------------------------------------------------


library(psych)

library(nFactors)

library(lavaan)

library(performance)

library(ltm)



# Data Subscales ----------------------------------------------------------


# Subset of 28 items (excludes Survey ID)

subset=data160[,2:29]


# EFAmodel0: 28 items

EFAmodel0data=data160[,2:29]


# EFAmodelA: 25 items (excludes 18,10,6)

EFAmodelAdata=data160pruned[,2:26]
```

```
# EFAmodelB: 23 items (excludes 18,10,6,13,19)
EFAmodelBdata=data183[,2:24]

# EFAmodelC: 22 items (excludes 18,10,6,13,19,2)
EFAmodelCdata=data183pruned[,2:23]

# EFAmodelD: 19 items (excludes 18,10,6,13,19,2,14,20,4)
EFAmodelDdata=data198[,2:20]

# Subset of 18 items (pruned model - excludes 18,10,6,13,19,2,14,20,4,7)
subset18=dataPruned[,2:19]

# Subset of 16 items (pruned model - excludes 18,10,6,13,19,2,14,20,4,7,3,26)
subset16=dataPrunedB[,2:17]


#########################################################################
#              Item Scores                      #
#########################################################################

# Frequencies for Each Item
solution<-as.data.frame(table(unlist(subset)))

# Each items' total score
colSums(subset)
```

```
# Each items' mean score
colMeans(subset)

# histogram plots
multi.hist(subset,freq=TRUE)

# Covariance Matrix
covMatrix = cov(subset)

# Item Variance
itemVar = diag(covMatrix)

# Item Co-Variance
itemcor = t(covMatrix)[lower.tri(t(covMatrix))]


##########################################################################
#              Descriptive Statistics                    #
##########################################################################

# basic descriptive statistics
describe(subset)

# Sample Variance
round(diag(cov(subset)),3)
```

```
#Sample Co-Variance
round(t(cov(subset))[lower.tri(t(cov(subset)))],3)

#Sample Correlation Matrix
round(cor(subset),3)
cor.plot(subset,3)


#########################################################################
#           Explanatory Factor Analysis (EFA)                 #
#########################################################################

# Scree Test
ev <- eigen(cor(subset)) # get eigenvalues
ap <- parallel(subject=nrow(subset),var=ncol(subset),rep=100,cent=.05)
nS <- nScree(x=ev$values, aparallel=ap$eigen$qevpea)
plotnScree(nS)


### Exploratory Factor Analyses--------------------------------------

KMO(EFAmodelAdata)
bartlett.test(EFAmodelAdata)

KMO(EFAmodelBdata)
```

```
bartlett.test(EFAmodelBdata)

KMO(EFAmodelCdata)
bartlett.test(EFAmodelCdata)

KMO(EFAmodelDdata)
bartlett.test(EFAmodelDdata)

KMO(EFAmodelEdata)
bartlett.test(EFAmodelEdata)

KMO(subset18)
bartlett.test(subset18)

KMO(subset16)
bartlett.test(subset16)

##EFAmodel0 - 28 items, n=160##
EFAmodel0 <- fa(EFAmodel0data, nfactors=2, n.obs=1312, rotate="promax", fm="pa")
print(EFAmodel0, digits=3, cutoff=0.30, sort=TRUE)
factor.stats(EFAmodel0data,EFAmodel0,n.obs=2000)  #get model fit indices
fa.diagram(EFAmodel0)
fa.diagram(EFAmodel0$Structure)
EFAmodel0$loadings[,1:2] #extract factor loading

##EFAmodelA - 25 items: items 18,10,6 removed, n=160##
```

```
EFAmodelA <- fa(EFAmodelAdata, nfactors=2, n.obs=1312, rotate="promax", fm="pa")
print(EFAmodelA, digits=3, cutoff=0.30, sort=TRUE)
factor.stats(EFAmodelAdata,EFAmodelA,n.obs=2000)  #get model fit indices
fa.diagram(EFAmodelA)
fa.diagram(EFAmodelA$Structure)
EFAmodelA$loadings[,1:2] #extract factor loading

##EFAmodelB - 23 items: 13,19 removed, n=183##
EFAmodelB <- fa(EFAmodelBdata, nfactors=2, n.obs=1312, rotate="promax", fm="pa")
print(EFAmodelB, digits=3, cutoff=0.30, sort=TRUE)
factor.stats(EFAmodelB,EFAmodelB,n.obs=2000)  #get model fit indices
fa.diagram(EFAmodelB)
fa.diagram(EFAmodelB$Structure)
EFAmodelB$loadings[,1:2] #extract factor loading

##EFAmodelC - 22 items: 2 removed, n=183##
EFAmodelC <- fa(EFAmodelCdata, nfactors=2, n.obs=1312, rotate="promax", fm="pa")
print(EFAmodelC, digits=3, cutoff=0.30, sort=TRUE)
factor.stats(EFAmodelCdata,EFAmodelC,n.obs=2000)  #get model fit indices
fa.diagram(EFAmodelC)
fa.diagram(EFAmodelC$Structure)
EFAmodelC$loadings[,1:2] #extract factor loading

##EFAmodelD - 19 items: 14,20,4 removed, n=198##
EFAmodelD <- fa(EFAmodelDdata, nfactors=2, n.obs=1312, rotate="promax", fm="pa")
print(EFAmodelD, digits=3, cutoff=0.30, sort=TRUE)
```

```
factor.stats(EFAmodelDdata,EFAmodelD,n.obs=2000)  #get model fit indices
fa.diagram(EFAmodelD, digits=3)
fa.diagram(EFAmodelD$Structure, digits=3)
EFAmodelD$loadings[,1:2] #extract factor loading

##EFAmodelE - 18 items: 7 removed, n=198##
EFAmodelE <- fa(subset18, nfactors=2, n.obs=1312, rotate="promax", fm="pa")
print(EFAmodelE, digits=3, cutoff=0.30, sort=TRUE)
factor.stats(subset18,EFAmodelE,n.obs=2000)  #get model fit indices
fa.diagram(EFAmodelE, digits=3)
fa.diagram(EFAmodelE$Structure, digits=3)
EFAmodelE$loadings[,1:2] #extract factor loading


#####################################################################
#     MODEL A:  Single Factor Model All 28 Items                    #
#####################################################################


SingleFactor = '
# Fully Z-Scored Factor Identification Approach
F =~ item2 + item3 + item4 + item5 + item6 + item8 + item10 + item11 + item12 + item13
+ item15 + item17 + item18 + item20 + item21 + item22 + item23 + item24 + item25 + item26 +
item27 + item28 + item31 + item33 + item34

# Item intercepts --> ~ 1 indicates means or intercepts
```

```
item2 ~ 1; item3 ~ 1; item4 ~ 1; item5 ~ 1; item6 ~ 1; item8 ~ 1; item10 ~ 1; item11 ~ 1;
item12 ~ 1; item13 ~ 1; item15 ~ 1; item17 ~ 1; item18 ~ 1; item21 ~ 1; item22 ~ 1; item23 ~ 1;
item24 ~ 1; item25 ~ 1; item26 ~ 1; item27 ~ 1; item28 ~ 1; item31 ~ 1; item33 ~ 1; item34 ~ 1;

# Item error (unique) variances and covariances --> use the ~~ command
item2 ~~ item2; item3 ~~ item3; item4 ~~ item4; item5 ~~ item5; item6 ~~ item6; item8
~~ item8; item10 ~~ item10; item11 ~~ item11; item12 ~~ item12; item13 ~~ item13; item15 ~~
item15; item17 ~~ item17; item18 ~~ item18; item20 ~~ item20; item21 ~~ item21; item22 ~~
item22; item23 ~~ item23; item24 ~~ item24; item25 ~~ item25; item26 ~~ item26; item27 ~~
item27; item28 ~~ item28; item31 ~~ item31; item33 ~~ item33; item34 ~~ item34;

# Factor variance
F ~~ F

# Factor mean (intercept)
F ~ 0
"
modelA = lavaan(model = SingleFactor, data = subset, estimator = "MLR", mimic =
"mplus", std.lv = FALSE)
summary(modelA, fit.measures = TRUE, rsquare = TRUE, standardized = TRUE,
header = TRUE)
model_performance(modelA)

#########################################################################
#   MODEL B:   Original 28Q, 6F Multiple Factor (Ellis) Model            #
#########################################################################
```

```
OrigMultiFactor = ''
# F1 loadings (Starting)
F1 =~ item2 + item3 + item4 + item5 + item6 + item7 + item8
# F2 loadings (Searching)
F2 =~ item10 + item11 + item12 + item13 + item14
# F3 loadings (Differentiating)
F3 =~ item15 + item17 + item18 + item19
#F4 loadings (Extracting)
F4 =~ item20 + item21 + item22 + item23
#F5 loadings (Verifying)
F5 =~ item24 + item25 + item26 + item27 + item28
#F6 loadings (Ending)
F6 =~ item31 + item33 + item34
''
modelB = sem(model = OrigMultiFactor, data = subset, estimator = "MLR", mimic =
"mplus", std.lv = TRUE)
summary(modelB, fit.measures = TRUE, rsquare = TRUE, standardized = TRUE,
header = TRUE)
model_performance(modelB)



#########################################################################
#   MODEL C:   Pruned 18Q, 2F Multiple Factor (EFA) Model           #
#########################################################################
```

```
PrunedMultiFactor = "
# F1 loadings (F1)
F1 =~ item 3 + item5 + item8 + item11 + item12 + item15 + item17 + item24 + item34
# F2 loadings (F2)
F2 =~ item21 + item22 + item23 + item25 + item26 + item27 + item28 + item31 + item33
"
modelC = sem(model = PrunedMultiFactor, data = subset18, estimator = "MLR", mimic =
"mplus", std.lv = TRUE)
summary(modelC, fit.measures = TRUE, rsquare = TRUE, standardized = TRUE,
header = TRUE)
model_performance(modelC)



#########################################################################
#   MODEL D:   Pruned 16Q (removing Poor IRR items), 2F Multiple Factor (EFA) Model
#
#########################################################################

PrunedMultiFactorB = "
# F1 loadings (F1)
F1 =~ item5 + item8 + item11 + item12 + item15 + item17 + item24 + item34
# F2 loadings (F2)
F2 =~ item21 + item22 + item23 + item25 + item27 + item28 + item31 + item33
"
```

```
modelD = sem(model = PrunedMultiFactorB, data = subset16, estimator = "MLR", mimic = "mplus", std.lv = TRUE)
summary(modelD, fit.measures = TRUE, rsquare = TRUE, standardized = TRUE, header = TRUE)
model_performance(modelD)

#########################################################################
#       Model ANOVA & Cronbach Alpha                    #
#########################################################################

### model comparison
anova(modelA, modelB, modelC)



###Cronbach Alpha
cronbach.alpha(F1orig)
cronbach.alpha(F2orig)



cronbach.alpha(F1pruned)
cronbach.alpha(F2pruned)



#### Residual Matrix-----------------------------------------------------------
resid(object = modelA, type = "raw")
resid(object = modelB, type = "raw")
```

```
resid(object = modelC, type = "raw")



#########################################################################
#               Factor Score                            #
#                                                       #
#########################################################################

# Function to estimate factor scores----------------------------------------------

factorScores = function(lavObject){
  output = inspect(object = lavObject, what = "est")
  sigma = output$lambda %*% output$psi %*% t(output$lambda) + output$theta
  modelData = lavObject@Data@X[[1]]
  scores = t(output$alpha%*%matrix(1, nrow=1, ncol = dim(modelData)[1]) +
output$psi %*% t(output$lambda) %*% solve(sigma)%*%(t(modelData) -
output$nu%*%matrix(1, nrow=1, ncol=dim(modelData)[1])))
  varscores = output$psi - output$psi %*% t(output$lambda) %*% solve(sigma) %*%
output$lambda %*% output$psi
  factorSE = sqrt(diag(varscores))
  names(factorSE) = paste0(names(factorSE), ".SE")
  factorSEmat = matrix(1, nrow=nrow(scores), ncol = 1) %*% matrix(factorSE, nrow = 1,
ncol = ncol(scores))
  colnames(factorSEmat) = names(factorSE)
  result = data.frame(cbind(scores, factorSEmat))
```

```
  names(result)
  odds = seq(1, ncol(result)-1, 2)
  evens = seq(2, ncol(result), 2)
  result = result[c(odds,evens)]
  factorCov = varscores
  if (dim(varscores)[1] == 1 & dim(varscores)[2] == 1){
   factorCorr = solve(sqrt(varscores)) %*% varscores %*% solve(sqrt(varscores))
  } else {
   factorCorr = solve(sqrt(diag(diag(varscores)))) %*% varscores %*%
solve(sqrt(diag(diag(varscores))))
  }
  return(list(scores = result, factorCov = factorCov, factorCorr = factorCorr))
 }

 factorscore = factorScores(modelC)
 head(factorscore)
 var(factorscore$F1)
 var(factorscore$F2)


 ### Function to estimate factor score reliability----------------------------------------------------
------

 factorScoreReliability = function(lavObject){
  output = inspect(object = lavObject, what = "est")
  sigma = output$lambda %*% output$psi %*% t(output$lambda) + output$theta
```

```
  varscores = output$psi - output$psi %*% t(output$lambda) %*% solve(sigma) %*%
output$lambda %*% output$psi
  return(diag(output$psi)/(diag(output$psi) + diag(varscores)))
}

factorScoreReliability(modelC)


### Histogram overlaid with kernel density curve----------------------------------------------------
---------

# For F1

a =ggplot(factorscore$scores, aes(x=F1)) +
  geom_histogram(aes(y=after_stat(density)),      # Histogram with density instead of
count on y-axis
                 binwidth=.5,
                 colour="black", fill="white") + xlim(c(-4,4)) + labs(title = "F1 Factor Score") +
  geom_density(alpha=.2, fill="#FF6666")    # Overlay with transparent density plot

# For F2
b= ggplot(factorscore$scores, aes(x=F2)) +
  geom_histogram(aes(y=after_stat(density)),      # Histogram with density instead of
count on y-axis
                 binwidth=.5,
                 colour="black", fill="white") + xlim(c(-4,4)) + labs(title = "F2 Factor Score") +
```

```
  geom_density(alpha=.2, fill="#FF6666")    # Overlay with transparent density plot
ggarrange(a,b,ncol=2, nrow=1)

### Plot for observed score and factor score----------------------------------------------------------
---------

F1Mean = apply(X = subset18[c("item3", "item5", "item8", "item11", "item12", "item15",
"item17", "item24", "item34")], MARGIN = 1, FUN = mean)
F2Mean = apply(X = subset18[c("item21", "item22", "item23", "item25", "item26",
"item27", "item28", "item31", "item33")], MARGIN = 1, FUN = mean)
par(mfrow = c(1,2))
plot(y = F1Mean, x = factorscore$scores$F1, xlim = c(-3.5,2), ylim = c(1,7), xlab = "F1
Factor Score", ylab = "F1 Items Mean")
text(x = -3, y = 6, labels = paste("r = ", as.character(round(cor(F1Mean,
factorscore$scores$F1), digits = 3))))
plot(y = F2Mean, x = factorscore$scores$F2, xlim = c(-3.5,2), ylim = c(1,7), xlab = "F2
Factor Score", ylab = "F2 Items Mean")
text(x = -3, y = 6, labels = paste("r = ", as.character(round(cor(F2Mean,
factorscore$scores$F2), digits = 3))))


### Plot for Predicted Item Response vs Factor Score


cfaPlots = function(lavObject){
  output = inspect(object = lavObject, what = "est")
```

```
#get factor scores
fscores = factorScores(lavObject = lavObject)
nfactors = ncol(fscores$scores)/2

#get max observed data
itemMax = max(apply(X = lavObject@Data@X[[1]], MARGIN = 2, FUN = max))
itemMin = min(apply(X = lavObject@Data@X[[1]], MARGIN = 2, FUN = min))

#get range for all scores
factorMax = max(apply(X = fscores$scores[seq(1, ncol(fscores$scores), 2)], MARGIN =
2, FUN = max))
factorMin = min(apply(X = fscores$scores[seq(1, ncol(fscores$scores), 2)], MARGIN =
2, FUN = min))

#set up x values
x = seq(factorMin, factorMax, .01)

par(mfrow = c(1, nfactors))
#make plots by factor
factor=1
for (factor in 1:nfactors){
  xmat = NULL
  ymat = NULL
  inames = NULL
  for (item in 1:nrow(output$lambda)){
```

```
        if (output$lambda[item, factor] != 0){
          inames = c(inames, rownames(output$lambda)[item])
          y = output$nu[item] + output$lambda[item, factor]*x
          xmat = cbind(xmat, x)
          ymat = cbind(ymat, y)
        }
      }
      matplot(x = xmat, y = ymat, type = "l", lwd = 5, lty=2:(ncol(xmat)+1), ylim = c(itemMin-
1, itemMax+1), xlim = c(factorMin, factorMax),
              ylab = "Predicted Item Response", xlab = colnames(output$lambda)[factor], col =
2:(ncol(xmat)+1))
      lines(x = c(factorMin,factorMax), y = c(itemMin, itemMin), lty = 3, lwd = 5)
      lines(x = c(factorMin,factorMax), y = c(itemMax, itemMax), lty = 3, lwd = 5)
      legend(x = -3, y = 7, legend = inames, lty = 2:(ncol(xmat)+1), lwd = 5, col =
2:(ncol(xmat)+1))
    }
    par(mfrow = c(1,1))
  }
  par(mfrow = c(1,2))
  cfaPlots(lavObject = modelC)



  ####################################################################
  #           Inter-Rater Reliability                 #
  #                                    #
  ####################################################################
```

```
# Function to estimate inter-rater reliability----------------------------------------------

rater1Q3 <-
c(1,1,1,1,1,1,1,1,0,0,1,0,0,1,1,1,0,1,1,0,1,1,1,1,1,1,0,1,0,0,1,0,0,1,0,0,1,1,1,1,1,1,1,1,1,1,0,0,1
)
rater2Q3 <-
c(1,1,0,1,1,0,1,0,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,0,0,1,1,1,0,1,1,1,1,1,1,1,1,0,1,0,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q3,rater2Q3))

rater1Q5 <-
c(1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,0,0,1,1,1,1,1,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,0,1
)
rater2Q5 <-
c(1,1,1,1,0,0,1,0,0,1,0,1,1,0,1,1,0,1,1,0,1,1,1,0,0,0,1,0,1,1,0,1,0,0,1,0,1,1,1,1,1,1,0,1,0,0,1,0,1,0,1
)
cohen.kappa(x=cbind(rater1Q5,rater2Q5))

rater1Q8 <-
c(0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0
)
rater2Q8 <-
c(1,1,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
)
cohen.kappa(x=cbind(rater1Q8,rater2Q8))
```

```
rater1Q11 <-
c(1,1,0,1,0,1,1,1,1,0,1,1,0,0,1,0,0,0,1,1,1,1,0,1,0,0,1,1,0,1,0,1,1,1,1,0,0,1,0,0,0,1,0,0,1,1,0,0,0,1
)
rater2Q11 <-
c(1,1,0,1,0,0,1,1,1,1,0,0,1,1,0,0,0,1,1,1,1,1,1,1,0,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q11,rater2Q11))

rater1Q12 <-
c(1,0,0,0,0,0,0,1,0,1,1,0,0,0,0,0,0,1,1,0,0,1,1,1,0,0,0,1,1,1,0,1,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,1
)
rater2Q12 <-
c(1,0,0,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,1,0,0,0,0,0
)
cohen.kappa(x=cbind(rater1Q12,rater2Q12))

rater1Q17 <-
c(1,0,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0,0,1,1,0,1,0,1,1,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1
)
rater2Q17 <-
c(1,1,1,1,1,1,1,1,0,1,1,0,1,0,0,1,0,1,1,1,1,1,1,1,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q17,rater2Q17))
```

```
rater1Q21 <-
c(1,0,0,1,0,0,1,1,1,1,1,0,0,0,0,0,0,0,1,0,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,1,1,0,0,0,1
)
rater2Q21 <-
c(1,1,1,1,1,1,1,1,1,1,1,0,1,0,0,0,0,1,1,0,1,1,1,1,0,0,0,0,1,1,0,1,1,1,1,0,1,1,0,1,0,0,0,0,1,0,1,1,0,1
)
cohen.kappa(x=cbind(rater1Q21,rater2Q21))

rater1Q22 <-
c(1,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
)
rater2Q22 <-
c(1,0,0,1,1,1,1,1,1,1,1,0,1,0,0,0,0,1,1,1,1,1,0,1,0,0,1,1,1,1,0,1,1,1,0,1,1,1,0,1,0,1,0,0,1,1,1,1,0,1
)
cohen.kappa(x=cbind(rater1Q22,rater2Q22))

rater1Q23 <-
c(1,0,0,0,0,1,1,1,1,0,1,0,0,1,1,1,0,0,1,1,1,1,1,1,0,1,1,1,1,1,1,0,0,1,1,1,1,1,0,1,0,0,0,1,1,0,1,0,0,1
)
rater2Q23 <-
c(1,0,0,1,1,1,0,1,1,0,1,0,1,1,1,1,0,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,0,1,1,1,1,0,1,0,1,0,1,1,0,1,0,1,1
)
cohen.kappa(x=cbind(rater1Q23,rater2Q23))
```

```
rater1Q24 <-
c(1,0,0,1,1,0,1,1,1,0,1,0,1,0,0,0,0,1,0,1,0,1,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,1,0,0,1,0,0,0,0,1
)
rater2Q24 <-
c(0,0,0,1,0,0,1,1,1,0,1,0,1,0,0,0,1,1,1,1,1,0,1,0,1,1,1,1,1,0,1,1,1,0,1,1,1,0,1,0,1,0,0,1,1,1,1,0,1
)
cohen.kappa(x=cbind(rater1Q24,rater2Q24))

rater1Q25 <-
c(0,1,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
)
rater2Q25 <-
c(0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q25,rater2Q25))

rater1Q26 <-
c(0,1,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0
)
rater2Q26 <-
c(1,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q26,rater2Q26))
```

```
rater1Q27 <-
c(1,0,1,1,0,0,1,1,1,0,1,0,1,0,0,0,0,1,0,1,1,1,1,1,0,0,0,0,0,0,0,0,1,0,0,0,1,1,0,1,0,1,0,0,1,1,1,0,0,1
)
rater2Q27 <-
c(1,1,1,1,0,1,1,1,1,0,1,0,1,0,0,0,0,1,1,0,1,1,1,0,0,0,0,0,0,1,0,1,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,0,1,1
)
cohen.kappa(x=cbind(rater1Q27,rater2Q27))

rater1Q28 <-
c(1,1,1,1,0,0,1,0,1,0,1,1,1,0,0,0,0,1,0,0,1,1,1,1,0,0,1,1,0,0,0,1,1,1,0,0,1,0,0,1,0,0,0,0,1,1,1,0,0,0
)
rater2Q28 <-
c(1,0,1,1,0,0,1,1,1,0,0,1,1,0,0,0,0,1,1,0,1,1,1,0,0,0,0,0,0,0,0,1,1,1,1,1,1,0,0,1,1,0,0,0,0,0,1,0,1,1
)
cohen.kappa(x=cbind(rater1Q28,rater2Q28))

rater1Q31 <-
c(1,1,0,1,1,1,1,1,1,0,1,0,1,0,0,1,1,1,0,0,1,1,1,1,0,0,1,0,0,1,0,1,1,0,1,1,1,1,0,1,0,1,0,0,1,0,0,0,0,0
)
rater2Q31 <-
c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q31,rater2Q31))
```

```
rater1Q33 <-
c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,0,1,1,1,1,0,1,1,0,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1
)
rater2Q33 <-
c(1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1
)
cohen.kappa(x=cbind(rater1Q33,rater2Q33))

rater1Q34 <-
c(1,1,1,0,1,1,0,0,0,1,1,1,1,0,0,0,1,1,1,1,1,1,1,1,0,1,1,1,1,0,1,1,0,1,0,1,1,0,1,1,1,0,0,1,1,0,1,0,1
)
rater2Q34 <-
c(1,1,1,0,1,0,0,0,0,1,1,1,1,0,0,0,1,1,1,0,1,1,1,0,0,0,1,1,1,0,1,0,1,1,1,1,0,1,1,1,0,0,1,1,0,1,0,1
)
cohen.kappa(x=cbind(rater1Q34,rater2Q34))
```

Appendix D: ***Items with Subtheme References Coded with Percentage Coverage***

Theme	Subtheme References Coded with % Coverage	IRMAT Item
Starting	Protocol : 6 references coded, 1.72% coverage	Was the protocol described?
Starting	gap_literature : 9 references coded, 2.59% coverage; Rationale : 5 references coded, 1.44% coverage	Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?
Starting	PICO_elements : 8 references coded, 2.30% coverage	Were the variables of interest clearly stated?
Starting	conceptual-operational_definitions : 8 references coded, 2.30% Coverage	Were operational and conceptual definitions of variables provided?
Starting	problem_identification : 9 references coded, 2.59% coverage	Was the problem stated unambiguously and was it easy to identify?
Starting	Purpose : 6 references coded, 1.72% coverage	Was there an explicit statement of the purpose or aim of the review?
Starting	review_questions : 10 references coded, 2.87% coverage	Was there an explicit statement of the research question(s) the review addresses?
Starting	sampling_frame : 4 references coded, 1.15% coverage	Was the sampling frame clearly stated?
Starting	theoretical-conceptual_framework : 9 references coded, 2.59% coverage	Is a theoretical or conceptual framework used to guide the review?
Searching	comprehensiveness : 12 references coded, 3.45% coverage; number_databases : 3 references coded, 0.86% coverage; number_strategies : 6 references coded, 1.72% coverage	Is a description of a comprehensive search provided?
Searching	database_search : 11 references coded, 3.16% coverage; keywords : 11 references coded, 3.16% coverage; term_variations : 1 reference coded, 0.29% coverage	Were detailed descriptions of the database search processes provided?

Theme	Subtheme References Coded with % Coverage	IRMAT Item
	limiter_rationale : 5 references coded, 1.44% coverage search_dates : 3 references coded, 0.86% coverage; strategy_modification : 1 reference coded, 0.29% coverage; subject_headings : 2 references coded, 0.57% coverage	
Searching	search_example : 6 references coded, 1.72% coverage; reproducible_search : 2 references coded, 0.57% coverage	Was a reproducible line-by-line search strategy (or a sequence of terms for simpler interfaces) provided for at least one database?
Searching	non-database_search : 20 references coded, 5.75% coverage; contacting_authors : 8 references coded, 2.30% coverage; forward_backward : 10 references coded, 2.87% coverage; grey_lit : 3 references coded, 0.86% coverage; hand_searching : 8 references coded, 2.30% coverage; internet_searches : 2 references coded, 0.57% coverage; research_registries : 4 references coded, 1.15% coverage	Were other recommended approaches to searching the literature used?
Searching	publication_bias : 2 references coded, 0.57% coverage	If applicable to the purpose and type of literature included, was publication bias addressed?
Differentiating	inclusion-exclusion : 7 references coded, 2.01% coverage; criteria_specifics : 14 references coded, 4.02% coverage	Were the inclusion/exclusion criteria for the eligible sources clearly stated?
Differentiating	Sampling : 5 references coded, 1.44% coverage	Was purposive sampling of the search results used?

Theme	Subtheme References Coded with % Coverage	IRMAT Item
Differentiating	Screening : 8 references coded, 2.30% coverage	Were the search results screened for relevance using a pre-specified set of eligibility criteria?
Differentiating	flowchart : 7 references coded, 2.01% coverage	Is a flowchart of search results provided?
Differentiating	Subgroups : 4 references coded, 1.15% coverage	Were the retrieved sources divided into subgroups?
Extracting	data_extraction : 9 references coded, 2.59% coverage; extraction_criteria : 7 references coded, 2.01% coverage; extraction_tool : 8 references coded, 2.30% coverage	Were the relevant data extracted from all sources using a pre-determined extraction instrument (for all subgroups, if applicable)?
Extracting	process : 9 references coded, 2.59% coverage	Is the data extraction process explicit, unbiased, and reproducible?
Extracting	matrix : 9 references coded, 2.59% coverage	Were data analyzed using a review matrix?
Extracting	data_display : 8 references coded, 2.30% coverage	Is a data display assembling the data from retrieved sources provided?
Verifying	analysis : 26 references coded, 7.47% coverage	Was a systematic analytic method explicitly identified?
Verifying	comparison : 10 references coded, 2.87% coverage	Were commonalities and differences identified?
Verifying	conflicting_evidence : 6 references coded, 1.72% coverage	Was conflicting evidence addressed?
Verifying	critical_appraisal : 18 references coded, 5.17% coverage; quality_criteria : 11 references coded, 3.16% coverage; risk_of_bias : 4 references coded, 1.15% coverage; study_reliability : 2 references coded, 0.57% coverage	Was the quality of retrieved sources addressed in a meaningful way?
Verifying	number_reviewers : 7 references coded, 2.01% coverage	Was quality verified by two independent reviewers?
Verifying	exclusion_quality : 10 references coded, 2.87% coverage	Were any articles excluded based on quality appraisal?
Ending	conclusions : 20 references coded, 5.75% coverage	Were patterns, themes, relationships, or conclusions verified with sources?

Theme	Subtheme References Coded with % Coverage	IRMAT Item
Ending	synthesis : 18 references coded, 5.17% coverage	Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?
Ending	visualization : 8 references coded, 2.30% coverage; conceptual_model : 3 references coded, 0.86% coverage; tables : 2 references coded, 0.57% coverage	Were the conclusions reported in table or diagrammatic form?
Ending	implications : 6 references coded, 1.72% coverage	Were implications discussed for: research, practice, education, or policy?
Ending	limitations : 2 references coded, 0.57% coverage	Were methodological or other limitations of the review explicitly stated?

Appendix E: *IRMAT Items (V 1) Generated with Methodological Elements*

STARTING THEME – 9 Items

Was the protocol described?

DESCRIPTION: A protocol is a predetermined documented plan for the review, including the rationale, research question, and proposed methods for the review. Detailed protocols should be developed *a priori* and made publicly available prior to carrying out the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no description of a protocol.	A protocol is not explicitly described; however, there are indicators (description of methods established *a priori,* etc.) that a protocol may have been completed. Or the protocol is provided only via online supplement not accessible to reader.	A protocol is explicitly referenced, but there is no description nor link provided.	A protocol is referenced, and there is a clear description provided. Or there is a link to an accessible online supplement or protocol registry.

Protocol : 6 references coded, 1.72% coverage

- *[the protocol] should provide a detailed description of the proposed activity*
- *A written protocol is developed prior to the beginning of the review and specifies the research review plan in detail.*
- *After the purpose and/or review question(s) is clearly articulated, the reviewer will then determine whether an integrative review is the method of choice.*
- *the methods for the review of the literature should be established prior to commencement*
- *the proposed method of data analysis and synthesis should be documented in a review protocol before commencing the review*

- *The search protocol … for primary studies are clearly stated and the results of the search are comprehensive (using 2–3 search strategies).*

Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?

DESCRIPTION: To identify the gap in the literature and necessity for the review, a preliminary background literature review should be conducted.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clear statement of a preliminary background literature review being conducted nor a reference to the gaps in the literature for this review topic.	There may be references to gaps in the literature for this review topic, but there is no statement of preliminary literature reviews being conducted.	There are references to preliminary literature search(es) being conducted, but there is no clear statement that the preliminary searches were used to identify the gap.	There is a clear a statement that a preliminary background literature review was conducted to identify the gap in the literature for this review topic.

gap_literature : 9 references coded, 2.59% coverage

- *A rationale for conducting the review will be articulated based on a preliminary review of what is known and not known.*
- *a review of the phenomenon of interest in the broadest context to provide the reader with a background to the topic*
- *Any gap in the literature is then presented and should be explicitly described.*
- *Before the IR takes place, a preliminary review of the literature is conducted to support the need for the review.*
- *clearly identifying a problem from a gap in the literature*
- *Previous reviews may further the development of the topic and justify the need for proceeding with a new review, while at the same time helping to identify a gap in current knowledge.*

- *summarizes relevant research to present the gap in the literature that the review is to fill (Hudson-Barr 2004)*
- *The findings from this preliminary background literature review could then support proceeding with the development of a review purpose or review questions to address the phenomenon of interest.*
- *The importance of previous work that was reported on the topic area should be discussed, and any unanswered questions and conflicting or challenging findings should be described (Hudson-Barr 2004).*

Rationale : 5 references coded, 1.44% coverage

- *A rationale for conducting the review will be articulated based on a preliminary review of what is known and not known.*
- *background and significance for the research problem will provide justification for why the review is necessary*
- *IRR reports open by stating the issue they examine and why the reviewers think it is important.*
- *Rationale for conducting the integrative review*
- *The introduction of the IR provides the background and rationale for conducting the review.*

Were the variables of interest clearly stated?

DESCRIPTION: Variables of interest may be defined as concepts, target population, health care problem, etc. Variables may or may not be defined in terms of PICO(S) elements (population, intervention, comparison, outcome, setting).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES

There is no explicit statement of variables of interest, and variables cannot be inferred.	Variables of interest are not explicitly stated but may be inferred.	There is an explicit statement of the variables of interest, but variables are not clearly defined within specific contexts (population, outcomes, etc.).	There is an explicit statement of the variables of interest, and variables are clearly defined within specific contexts (population, outcomes, etc.).

PICO_elements : 8 references coded, 2.30% coverage

- *defining the question … includes the definition of participants, the interventions to be evaluated and the results to be measured(2).*
- *Identifying the target and accessible population are 2 key steps in this [data collection or literature search] stage.*
- *Key concepts and variables, such as population, intervention in health problem, must then be determined (Whittemore & Knafl, 2005).*
- *Key concepts such as the population for study, any interventions of interest, comparison groups, and outcomes should be evident in the guiding questions. [Shuler in Holly, 2014]*
- *Key concepts such as the population for study, any interventions of interest, comparison groups, and outcomes, should be evident in the guiding questions. [Holly, 2019]*
- *the method and the variables of interest (i.e., target population, health care problem or intervention, clinical outcomes) are determined*
- *the variables of interest (that is, concepts, target population, health care problem) … are determined*
- *You should note if the review focused on a certain population or setting, and whether it is focused on one or several outcomes.*

Were operational and conceptual definitions of variables provided?

DESCRIPTION: Operational definitions define variables in terms of how they are measured. Conceptual definitions are the abstract or theoretical meanings of variables.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Neither operational nor conceptual definitions are provided.	It is unclear whether operational and conceptual definitions are provided, or definitions are provided but not for all relevant variables.	Only operational -or- only conceptual definitions are provided.	Both operational and conceptual definitions are provided for all variables.

conceptual-operational_definitions : 8 references coded, 2.30% Coverage

- *clearly describing what is meant by the variables and how they will be used in the review*
- *concepts of interest related to the research problem need to be clearly defined*
- *definitions of conceptual variables must be clearly and concisely delineated by the reviewer*
- *operationalizes the phenomenon, population, and context;*
- *The conceptual and operational definitions of variables to be examined need to be developed.*
- *The introduction should include the list of guiding questions developed for the review and definitions of the conceptual and operational variables that were a part of the review. [Shuler in Holly, 2014]*
- *The introduction should include the list of guiding questions developed for the review and definitions of the conceptual and operational variables that were a part of the review. [Holly, 2019]*

- *The reviewer should also delineate the relationships between the variables under study. Using a theoretical framework to guide this early phase of the review process will enhance the ability of the reviewer to "fit" the results of the process into the body of developing nursing knowledge.*

Was the problem stated unambiguously and was it easy to identify?

DESCRIPTION: A problem statement outlines the areas under examination by the review. An easily identifiable problem statement may begin with a statement such as "This review aims to address the problem of…", or "Previous literature searches highlighted problems surrounding…", etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clearly stated problem.	A problem statement is alluded to but is not clearly stated.	The problem statement is clear but is not easily identifiable.	The problem statement is clear and is easily identifiable.

problem_identification : 9 references coded, 2.59% coverage

- *a clear identification of the problem that the review is addressing and the review purpose*
- *clear identification of the clinical or conceptual problem the review is addressing and the purpose of the review*
- *clear identification of the problem and the associated review purpose*
- *clearly identifying a problem from a gap in the literature*
- *definition of a problem and formulation of a hypothesis or research question that is relevant to health and nursing*
- *Having clearly stated review question(s), based on a carefully developed introduction and background, then sets the stage for the undertaking of the IR.*

- *identifies and describes the phenomenon to be investigated;*
- *Once the conceptual overview is complete, the introduction moves into the more specific description of how the phenomenon is to be addressed in the review, the setting or context, and the population being studied.*
- *the method and the variables of interest (i.e., target population, health care problem or intervention, clinical outcomes) are determined*

Was there an explicit statement of the purpose or aim of the review?

DESCRIPTION: The purpose and aim of the review differs from the problem statement in that the purpose/aim are actionable items and/or address the problem in some way. An explicit statement of the purpose/aim may commonly be represented by a phrase such as "The purpose/aim of this review is…".

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of the purpose or aim of the review, and it cannot be inferred from other context.	There is no explicit statement of the purpose or aim of the review, but it may be inferred from other context.	The purpose or aim is clear but is not explicitly stated.	The purpose or aim is clear and is explicitly stated.

Purpose : 6 references coded, 1.72% coverage

- *broad purpose and review question(s) are clearly stated*
- *clear identification of the clinical or conceptual problem the review is addressing and the purpose of the review*
- *clear identification of the problem and the associated review purpose*
- *Formulate review purpose and questions*
- *Purpose and aim of review*

- *The purpose describes the goal of the review, or why the review is being conducted.*

Was there an explicit statement of the research question(s) the review addresses?

DESCRIPTION: The research question(s) of the review differ from the problem statement and the purpose/aim in that the research question(s) are written in question form. An explicit statement of the research questions may commonly be represented by a phrase such as "The research questions guiding this review are…".

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of the research question(s) guiding the review, and it cannot be inferred from other context.	There is no explicit statement of the research question(s) guiding the review, but it may be inferred from other context.	The research question(s) guiding the review are clear but are not explicitly stated in question format.	The research question(s) guiding the review are clear and are explicitly stated in question format.

review_questions : 10 references coded, 2.87% coverage

- *broad purpose and review question(s) are clearly stated*
- *defining the question ... should be clearly and specifically prepared, and related to a theoretical principle that comprises theories and rationale learned by the researcher (1,9).*
- *defining the question … includes the definition of participants, the interventions to be evaluated and the results to be measured (2).*
- *definition of a problem and formulation of a hypothesis or research question that is relevant to health and nursing*
- *Formulate review purpose and questions*
- *Having clearly stated review question(s), based on a carefully developed introduction and background, then sets the stage for the undertaking of the IR.*

- *Review question*
- *The introduction should include the list of guiding questions developed for the review and definitions of the conceptual and operational variables that were a part of the review. [Shuler in Holly, 2014]*
- *The introduction should include the list of guiding questions developed for the review and definitions of the conceptual and operational variables that were a part of the review. [Holly, 2019]*
- *The review question(s) succinctly identifies what the review proposes to answer and suggests how it might contribute to a better understanding of the phenomenon of interest (Aveyard 2014).*

Was the sampling frame clearly stated?

DESCRIPTION: A sampling frame for an integrative review may include (but is not limited to): the type of empirical studies, specific research design(s), inclusion of methodological or theoretical literature/framework, etc. Providing a rationale for choice of sampling frame strengthens the choice of using a sampling frame. Attention should also be paid to the alignment of the choice of sampling frame with the integrative review methodology.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clear statement of sampling frame nor is rationale provided. Or, the sampling frame does not align with the integrative review methodology, i.e., limiting to only randomized controlled trials, where a systematic review would be more appropriate.	There is no clear statement of whether a sampling frame was used, nor is rationale provided; however, a sampling frame may be alluded to in the screening (inclusion/exclusion) or data analysis descriptions.	A sampling frame is clearly stated but no rationale is provided.	A sampling frame with rationale is clearly stated. If no sampling frame is used, this is also clearly stated.

sampling_frame : 4 references coded, 1.15% coverage

- *include empirical or theoretical literature, or both (Cooper 1984)*
- *reviewers consider the kinds of studies they will include in the review and how far back they will go in the search for studies*
- *the appropriate sampling frame are determined (that is, type of empirical studies, inclusion of theoretical literature)*
- *The reviewers also decide whether studies using the full range of designs will be included or just those with certain design characteristics.*

Is a theoretical or conceptual framework used to guide the review?

DESCRIPTION: A theoretical or conceptual framework (also called theoretical or conceptual models) are established descriptions of relationships between concepts and are often represented by visualizations or diagrams. The use of a theoretical or conceptual framework to guide the review is one of the elements that may distinguish integrative reviews from other review types (systematic, scoping, etc.).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no statement of a theoretical or conceptual model being used to guide the review.	A theoretical or conceptual framework may be inferred but is not clearly stated. It is also unclear how it is guiding or shaping the review.	A theoretical or conceptual framework is clearly stated, but it is unclear how it is guiding or shaping the review.	A theoretical or conceptual framework is clearly stated and is explicitly used throughout the review (screening, data extraction, synthesis, etc.).

theoretical-conceptual_framework : 9 references coded, 2.59% coverage

- *background and significance for the research problem will provide justification for why the review is necessary*

- *concepts of interest related to the research problem need to be clearly defined*
- *defining the question ... should be clearly and specifically prepared, and related to a theoretical principle that comprises theories and rationale learned by the researcher(1,9).*
- *If a theoretical framework is used to guide the review, an explanation of how this framework will be organizing the integrative review should be included.*
- *The findings from this preliminary background literature review could then support proceeding with the development of a review purpose or review questions to address the phenomenon of interest.*
- *The introduction of the IR provides the background and rationale for conducting the review.*
- *The reviewer should also delineate the relationships between the variables under study. Using a theoretical framework to guide this early phase of the review process will enhance the ability of the reviewer to "fit" the results of the process into the body of developing nursing knowledge.*
- *theoretical literature may help to define concepts*
- *when appropriate, experts suggest using a theoretical framework to guide the IR process (Soares et al. 2014; Russell 2005; Denney and Tewksbury 2013; Torraco 2005)*

SEARCHING THEME – 5 items

Is a description of a comprehensive search provided?

DESCRIPTION: A comprehensive search consists of a minimum of two to three search strategies (databases, grey literature, forward/backward searching, etc.) and a minimum of two databases, with three to six databases being preferred.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES

There is no explicit description of a search process, nor can it be inferred from the text.	There is a limited description of a search process which only includes: • 1 search strategy **AND** • 1+ database(s)	There is an explicit description of a search process which includes: • 2-3 search strategies **AND** • 1-2 databases	There is an explicit description of a search process which includes: • 2-3 search strategies **AND** • 3-6+ databases

comprehensiveness : 12 references coded, 3.45% coverage

- *comprehensive and replicable search strategy*
- *Ideally, all of the relevant literature on the problem or topic of interest is included in the review;*
- *include a clearly documented and comprehensive literature search, defining in detail all databases, search terms, limiters, eligibility (inclusion/exclusion), and criteria used, and describing any additional search methods (Cooper 1982; Whittemore and Knafl 2005)*
- *includes a comprehensive search, one that captures as much literature pertaining to the topic as possible (Evans 2007; Whittemore 2007)*
- *includes all of the relevant literature on the specified problem*
- *the search in databases should be broad and diverse …*
- *The search should involve seeking primary, secondary, and tertiary literature. [Shuler in Holly, 2014]*
- *The search should involve seeking primary, secondary, and tertiary literature. [Holly, 2019]*
- *uses diverse data sources*
- *All databases likely to contain the highest number of citations have been searched.*
- *Author searches on the most prolific authors of the topic show no new citations.*

- *New searches return no new, unique, and relevant results.*

number_databases : 3 references coded, 0.86% coverage

- *The search strategy should encompass a minimum of two databases, although the search can be more extensive.*
- *The search should be broad and encompass a minimum of two databases, although a search of three to six databases is preferred. [Shuler in Holly, 2014]*
- *The search should be broad and encompass a minimum of two databases, although a search of three to six databases is preferred. [Holly, 2019]*

number_strategies : 6 references coded, 1.72% coverage

- *A comprehensive search for a research review identifies the maximum number of eligible studies, utilizing as many search strategies as resources allow (at least 2–3 strategies; Jadad et al., 1998).*
- *a comprehensive search for an integrative review identifies the maximum number of eligible primary sources, using at least two to three strategies (Jadad et al. 1998, Conn et al. 2003b)*
- *includes a comprehensive search… [that] includes multiple strategies (Whittemore and Knafl 2005)*
- *Often multiple strategies such as computerized database searching, ancestry searching, journal hand searching, networking and or searching research registries are required to obtain a representative sample (Conn et al. 2003a).*
- *The search protocol … for primary studies are clearly stated and the results of the search are comprehensive (using 2–3 search strategies).*
- *two or more methods for literature search*

Were detailed descriptions of the database search processes provided?

DESCRIPTION: A detailed description of a database search will include (at a minimum): database name(s), keywords, whether subject headings were used, any applied limits (year, language), etc. A more comprehensive description will include (at a minimum): database name AND vendor, keywords AND variations, which controlled vocabularies were used, list of all limiters WITH rationale, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no description of the database search process. There may be a statement such as “Two databases were searched…”, but the database names, keywords, etc. are not provided.	There is a minimal description of the database search process provided, which may include: • database name(s) OR • keywords OR • whether or not subject headings were used OR • any applied limits, etc.	The following are explicitly stated: • database name(s) AND • keywords AND • whether or not subject headings were used AND • any applied limits, etc.	The following are explicitly stated: • database name(s) AND vendor AND • keywords AND variations AND • which controlled vocabularies were used AND • list of all limiters WITH rationale, etc.

database_search : 11 references coded, 3.16% coverage

- *A narrative description of all information sources, including databases, that were used; limiters used to narrow search results, such as year of publication, language, and publication status; and search terms used.*
- *careful selection of databases and search terms, searching bibliography lists, hand searching relevant journals and searching conference abstracts*
- *Conduct literature search*
- *Electronic databases searched*

- *include a clearly documented and comprehensive literature search, defining in detail all databases, search terms, limiters, eligibility (inclusion/exclusion), and criteria used, and describing any additional search methods (Cooper 1982; Whittemore and Knafl 2005)*
- *including search in electronic databases,*
- *Reviewers typically searched several health care databases using a variety of search terms, combinations of search terms, and search options.*
- *The literature search process should be clearly documented in the methods section, including the search terms, databases used, additional search strategies, and inclusion and exclusion criteria for the determination of relevant primary sources.*
- *the search begins in the databases to identify the studies that will be included in the review*
- *The strategy developed should include an electronic database search, a hand search in journals relevant to the guiding questions … [Shuler in Holly, 2014]*
- *The strategy developed should include an electronic database search, a hand search in journals relevant to the guiding questions … [Holly, 2019]*

keywords : 11 references coded, 3.16% coverage

- *A narrative description of all information sources, including databases, that were used; limiters used to narrow search results, such as year of publication, language, and publication status; and search terms used.*
- *Both natural language and controlled language searches are important to include in the search strategy*

- *careful selection of databases and search terms, searching bibliography lists, hand searching relevant journals and searching conference abstracts*
- *include a clearly documented and comprehensive literature search, defining in detail all databases, search terms, limiters, eligibility (inclusion/exclusion), and criteria used, and describing any additional search methods (Cooper 1982; Whittemore and Knafl 2005)*
- *Key search terms*
- *Reviewers typically searched several health care databases using a variety of search terms, combinations of search terms, and search options.*
- *should utilize more than one or two search terms or phrases*
- *The literature search process should be clearly documented in the methods section, including the search terms, databases used, additional search strategies, and inclusion and exclusion criteria for the determination of relevant primary sources.*
- *the literature search process should clearly document the search terms, the databases used, the search strategies, and the inclusion criteria for determining relevant studies*
- *All spelling variants and synonyms have been searched, including abbreviations.*
- *The search concepts are clear, not too narrow or too broad.*

term_variations : 1 reference coded, 0.29% coverage

- *All spelling variants and synonyms have been searched, including abbreviations.*

limiter_rationale : 5 references coded, 1.44% coverage

- *Searches should also be described narratively to provide additional details such as limiters that were applied or reasons behind decisions*
- *Any limits applied ... are important to document so that the scope of the search is clearly described*
- *include a clearly documented and comprehensive literature search, defining in detail all databases, search terms, limiters, eligibility (inclusion/exclusion), and criteria used, and describing any additional search methods (Cooper 1982; Whittemore and Knafl 2005)*
- *A narrative description of all information sources, including databases, that were used; limiters used to narrow search results, such as year of publication, language, and publication status; and search terms used.*
- *Search limiters, filters, and Boolean operators have been used correctly.*

search_dates : 3 references coded, 0.86% coverage

- *dates covered ... are important to document so that the scope of the search is clearly described*
- *dates of the search are important to document so that the scope of the search is clearly described*
- *it is essential to provide a clear and comprehensive discussion of how the gray literature was identified to provide transparency in search methods.*

strategy_modification : 1 reference coded, 0.29% coverage

- *The search strategy has been modified by adding terms based on citations highly relevant to the topic.*

subject_headings : 2 references coded, 0.57% coverage

- *Both natural language and controlled language searches are important to include in the search strategy*
- *Appropriate subject headings, or controlled language terms, have been identified and searched.*

Was a reproducible line-by-line search strategy (or a sequence of terms for simpler interfaces) provided for at least one database?

DESCRIPTION: A reproducible line-by-line search strategy includes sufficient detail that a reader may recreate the search process and results in a manner that the search results are replicated.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no search strategy provided. There may be a statement such as "A comprehensive search was conducted …", but keywords or terms, etc. are not provided sufficiently that a search could be replicated.	A reproducible search strategy is not explicitly described; however, there may be enough details provided that the search could possibly be replicated. Or the reproducible search strategy is provided only via online supplement not accessible to reader.	All reproducible search strategies are explicitly described; however, none are provided in the manuscript text, but reproducible line-by-line search strategies may be referenced and/or supplied only as online supplements.	All reproducible search strategies are explicitly described, and at least one is provided in the manuscript text (not as an online supplement). Additional reproducible line-by-line search strategies may be referenced and/or supplied as online supplements.

search_example : 6 references coded, 1.72% coverage

- *All search history (searches, search terms, results from those searches, and article citations) should be saved—even search results that may be excluded later. This information will be combined into a final reporting format, such as the PRISMA Flow Diagram (2015) or other type of search flow diagram.*

- *An example of the search strategy should be included in a table or in an appendix. [Shuler in Holly, 2014]*
- *An example of the search strategy should be included in a table or in an appendix. [Holly, 2019]*
- *comprehensive and replicable search strategy*
- *the literature search process should clearly document the search terms, the databases used, the search strategies, and the inclusion criteria for determining relevant studies*
- *the search report ... needs to be presented in a way that each search could be reproducible or can at least be properly evaluated*

reproducible_search : 2 references coded, 0.57% coverage

- *comprehensive and replicable search strategy*
- *All search history (searches, search terms, results from those searches, and article citations) should be saved—even search results that may be excluded later. This information will be combined into a final reporting format, such as the PRISMA Flow Diagram (2015) or other type of search flow diagram.*

Were other recommended approaches to searching the literature used?

DESCRIPTION: Non-database search approaches may include searching grey literature, ancestry/descendancy (forward/backward) searching, hand searching, networking, or searching research registries.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
No non-database search approaches are described.	Non-database approaches may be inferred but are unclear or not	Only one non-database search approach is used and/or minimal or	One or more non-database search approaches are used and described with

	described with sufficient detail.	insufficient detail is provided.	sufficient detail for replication.

non-database_search : 20 references coded, 5.75% coverage

conference_abstracts : 1 reference coded, 0.29% coverage

- *careful selection of databases and search terms, searching bibliography lists, hand searching relevant journals and searching conference abstracts*

contacting_authors : 8 references coded, 2.30% coverage

- *Also, authors who have previously published on the topic for review can be contacted to determine whether any other sources are available.*
- *Ancestry searching, journal hand-searching, networking, and searching research registries are approaches also recommended for searching the literature (Conn, Isaramalai, et al., 2003).*
- *Authors who have previously published on the topic for review can be contacted to determine whether any other sources are available or if they have published similar papers.*
- *Before undertaking the IR, the reviewer may choose to reach out to notable scholars of prior reviews or authors of primary research related to the phenomenon of interest.*
- *contact with researchers and the use of unpublished material(2).*
- *Often multiple strategies such as computerized database searching, ancestry searching, journal hand searching, networking and or searching research registries are required to obtain a representative sample (Conn et al. 2003a).*

- *other recommended approaches to searching the literature include ancestry searching, journal hand searching, networking, and searching research registries (Conn et al. 2003b)*
- *Thus, retrieval of eligible studies from database is only a starting point. The panel should go on to peruse reference list, go to research registries, contact colleagues, and even run searches using Web search engines.*

forward_backward : 10 references coded, 2.87% coverage

- *An examination of the reference list of those articles selected for the review is also helpful in identifying additional resources for the review.*
- *Ancestry searching, journal hand-searching, networking, and searching research registries are approaches also recommended for searching the literature (Conn, Isaramalai, et al., 2003).*
- *careful selection of databases and search terms, searching bibliography lists, hand searching relevant journals and searching conference abstracts*
- *Often multiple strategies such as computerized database searching, ancestry searching, journal hand searching, networking and or searching research registries are required to obtain a representative sample (Conn et al. 2003a).*
- *other recommended approaches to searching the literature include ancestry searching, journal hand searching, networking, and searching research registries (Conn et al. 2003b)*
- *Other search activities, such as: bibliography and reference lists searched; internet search strategy; journals hand-searched*
- *the references described in the selected studies*

- *The strategy developed should include … a search of the reference list of studies selected to be in the review.*
- *The strategy developed should include … a search of the reference list of studies selected to be included in the review.*
- *Thus, retrieval of eligible studies from database is only a starting point. The panel should go on to peruse reference list, go to research registries, contact colleagues, and even run searches using Web search engines.*

grey_lit : 3 references coded, 0.86% coverage

- *it is essential to provide a clear and comprehensive discussion of how the gray literature was identified to provide transparency in search methods.*
- *A search for grey or unpublished literature can be considered, but it is not necessary.*
- *A search for grey or unpublished literature can be considered.*

hand_searching : 8 references coded, 2.30% coverage

- *Ancestry searching, journal hand-searching, networking, and searching research registries are approaches also recommended for searching the literature (Conn, Isaramalai, et al., 2003).*
- *careful selection of databases and search terms, searching bibliography lists, hand searching relevant journals and searching conference abstracts*
- *manual search in journals*
- *Often multiple strategies such as computerized database searching, ancestry searching, journal hand searching, networking and or searching research registries are required to obtain a representative sample (Conn et al. 2003a).*

- *other recommended approaches to searching the literature include ancestry searching, journal hand searching, networking, and searching research registries (Conn et al. 2003b)*
- *Other search activities, such as: bibliography and reference lists searched; internet search strategy; journals hand-searched*
- *The strategy developed should include an electronic database search, a hand search in journals relevant to the guiding questions … [Shuler in Holly, 2014]*
- *The strategy developed should include an electronic database search, a hand search in journals relevant to the guiding questions … [Holly, 2019]*

internet_searches : 2 references coded, 0.57% coverage

- *Other search activities, such as: bibliography and reference lists searched; internet search strategy; journals hand-searched*
- *Thus, retrieval of eligible studies from database is only a starting point. The panel should go on to peruse reference list, go to research registries, contact colleagues, and even run searches using Web search engines.*

research_registries : 4 references coded, 1.15% coverage

- *Thus, retrieval of eligible studies from database is only a starting point. The panel should go on to peruse reference list, go to research registries, contact colleagues, and even run searches using Web search engines.*
- *Ancestry searching, journal hand-searching, networking, and searching research registries are approaches also recommended for searching the literature (Conn, Isaramalai, et al., 2003).*

- *other recommended approaches to searching the literature include ancestry searching, journal hand searching, networking, and searching research registries (Conn et al. 2003b)*
- *Often multiple strategies such as computerized database searching, ancestry searching, journal hand searching, networking and or searching research registries are required to obtain a representative sample (Conn et al. 2003a).*

If applicable to the purpose and type of literature included, was publication bias addressed?

DESCRIPTION: The use of gray literature as a non-database search approach may be one way of addressing publication bias. An explicit statement in the review acknowledging how publication bias was (or was not) addressed may strengthen the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Publication bias is applicable to the purpose and type of literature included but is not addressed by the review.	Publication bias is not applicable to the purpose and type of literature included, which is explicitly acknowledged by the review.	There is a not an explicit statement of how publication bias was acknowledged by the review, but there is evidence that grey literature or sources outside of traditional publication avenues were retrieved.	There is an explicit statement of how publication bias was acknowledged by the review.

publication_bias : 2 references coded, 0.57% coverage

- *addressing the issue of publication bias may also be relevant to the literature search stage (Conn et al. 2003b, Soeken & Sripusanapan 2003)*
- *The most important reason to consider searching beyond databases is minimize publication bias. This type of bias is defined as a situation that leads to a report not being published due to the nature of its results (Russell 2005).*

DIFFERENTIATION THEME – 5 items

Were the inclusion/exclusion criteria for the eligible sources clearly stated?

DESCRIPTION: Specific eligibility criteria for inclusion and exclusion are clearly stated. Inclusion and exclusion criteria should be aligned with the purpose and aim of the review. Ideally, rationale will be provided for all limits, particularly those such as excluding non-English sources or limiting to a recent span of years.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Inclusion/exclusion criteria for the eligible sources are not clearly stated.	Inclusion/exclusion criteria are provided but may not be clearly aligned with the review purpose. Or, rationale for eligibility criteria may not be clearly stated.	Inclusion/exclusion criteria which are clearly aligned with the review purpose are explicitly stated. Rationale for eligibility criteria is not clearly stated.	Inclusion/exclusion criteria which are clearly aligned with the review purpose are explicitly stated. Rationale for eligibility criteria is stated.

inclusion-exclusion : 7 references coded, 2.01% coverage

- *All decisions about literature to include or exclude should be justified and documented in the methods to demonstrate that an unbiased process was followed.*
- *Delineate inclusion and exclusion criteria*
- *effective inclusion and exclusion criteria*
- *however, if both choices are not feasible due to the amount of works, the inclusion and exclusion criteria adopted for the articles must be clearly explained and discussed(8).*
- *the inclusion and exclusion criteria are aligned with the review purpose and/or question(s)*

- *the literature search process of an integrative review should be clearly documented in the method section including … the inclusion and exclusion criteria for determining relevant primary sources*
- *This procedure for inclusion and exclusion of articles must be conducted in a judicious and transparent manner…*

criteria_specifics : 14 references coded, 4.02% coverage

- *all decisions made regarding the inclusion and exclusion criteria for studies are documented and justified in the description of the review methodology*
- *Any exclusion criteria such as language or date of publication*
- *Criteria for selecting studies and articles, including: population or setting; intervention, condition, concept or phenomenon; outcome or data of interest; research designs or type of articles; other criteria such as sampling time frame*
- *establishing inclusion criteria for the studies*
- *How the criteria were used*
- *How the criteria were used to select studies and articles*
- *the eligibility criteria for primary studies are clearly stated*
- *the process of selection … should be guided by clearly stated criteria that delineate the area of interest*
- *Publication language.*
- *Setting (Garrard 2014; Stern et al. 2014).*
- *The characteristics of the population being studied.*
- *The phenomenon under investigation.*
- *Time period covered by the review and its justification.*
- *Types of studies or literature.*

Was purposive sampling of the search results used?

DESCRIPTION: Purposive sampling is the deliberate selection of a sample of the retrieved sources, as opposed to evaluating the entirety of retrieved sources. Purposive sampling differs from the use of a sampling frame for the entire review, but purposive sampling is also strengthened by explicitly stating rationale.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Purposive sampling was not used. All retrieved sources were included.	There is no explicit statement that purposive sampling was used, but the number of results and sources cannot be reconciled.	Purposive sampling was used but no rationale was provided.	Purposive sampling was used, and an explicit statement of rationale was provided.

Sampling : 5 references coded, 1.44% coverage

- *any sampling decision must be justified and made explicit*
- *any sampling decision must be methodologically justified and made explicit*
- *Any sampling decision should be made explicit and justified (Whittemore & Knafl, 2005).*
- *purposive sampling can be combined with a comprehensive search*
- *The ideal procedure is to include all the studies found or a randomized selection of them;*

Were the search results screened for relevance using a pre-specified set of eligibility criteria?

DESCRIPTION: The pre-specified eligibility criteria for inclusion/exclusion are used to screen retrieved sources. A strong review will include a statement of both title/abstract and full text relevance screening.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Retrieved sources were not screened for inclusion/exclusion using eligibility requirements.	Insufficient detail is provided to determine whether retrieved sources are screened. Or, there is a statement of screening but no statement of eligibility criteria.	Retrieved sources are screened for inclusion/exclusion, but there is not an explicit statement that screening occurred in in two phases (title/abstract and full text). Or, the set of eligibility requirements was not pre-specified.	Retrieved sources are screened for inclusion/exclusion in two phases – title/abstract and full text – using a pre-specified set of eligibility requirements.

Screening : 8 references coded, 2.30% coverage

- *A review panel has greater potential to conduct an IRR that is free of error and bias than does an individual – the panel members act as checks and balances to each other's' work and uncover unconscious biases.*
- *Database search results … are next reviewed to determine which studies will be included in the review sample.*
- *screening by relevance*
- *screening, which involves reviewing the citations resulting from a search and selecting those deemed relevant for full-text retrieval*
- *The search and selection of articles included in the review should preferably be performed by two reviewers independently.11 [Polit & Beck, 2006]*
- *then selecting by full text*
- *This procedure for inclusion and exclusion of articles must be conducted in a judicious and transparent manner…*
- *Using a pre specified set of inclusion criteria, two persons (ideally) decide which studies qualify for full review.*

Is a flowchart of search results provided?

DESCRIPTION: A flowchart of search results presents the number of retrieved results, deduplicated results, results remaining after title/abstract and full text screening, and the final number of eligible sources. The most comprehensive flowcharts include detailed breakdowns of retrieved results by database/source and number of results excluded with reasons.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no flowchart of search results provided.	Flowchart of search results is not readily available or is provided only via online supplement not accessible to reader.	A flowchart with the following elements is provided: • number of retrieved results • deduplicated results • number of results remaining after title/abstract screening • number of results remaining after full text screening • final number of eligible sources	A flowchart with the following elements is provided: • number of retrieved results by database/source • deduplicated results • number of results remaining after title/abstract screening • number of results remaining after full text screening by reason for exclusion • final number of eligible sources

flowchart : 7 references coded, 2.01% coverage

- *A search diagram format that depicts the flow of information through different phases of the review. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) is an example of a reporting model that provides both a checklist and a search flow diagram that can be adapted for IR use (Moher et al. 2009).*
- *All search history (searches, search terms, results from those searches, and article citations) should be saved—even search results that may be excluded*

later. This information will be combined into a final reporting format, such as the PRISMA Flow Diagram (2015) or other type of search flow diagram.

- *finally sorting into studies ("Instead of counting those as two studies, it would be counted as one study, using both of the reports to collect data about the project. Sometimes one report is multiple studies, which should be treated separately.")*
- *For each database searched, a log should be kept that details the database searched, the range of dates searched, and the results of the search in terms of both the number of "hits" obtained and the number of relevant articles or other material retrieved.*
- *Results of the search process*
- *Results of the selection process*
- *The flow diagram will include details of databases searched, records identified through database searching, number of records after duplicates removed, number of records screened, number of full-text records excluded (with reasons), and number of studies included in the final sample, as illustrated in the PRISMA Flow Diagram (2015).*

Were the retrieved sources divided into subgroups?

DESCRIPTION: Examples of subgroups may include (but are not limited to): types of evidence, chronology, setting, sample characteristics, etc. The use of subgroups is strengthened by explicitly stating rationale for which group type(s) were chosen.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Retrieved sources were not divided into subgroups.	Retrieved sources were not divided into subgroups, where subgroups would have been appropriate. Or,	Retrieved sources were divided into subgroups but rationale for the group type(s) is not explicitly stated.	Retrieved sources were divided into subgroups, and the rationale for the group type(s) is explicitly stated.

	subgroups were not appropriate for the types of sources retrieved.		

Subgroups : 4 references coded, 1.15% coverage

- *The panel will then sort the articles into stacks by subtopics*
- *The primary sources in the sample will be initially reduced into subcategories (Whittemore and Knafl 2005).*
- *The primary sources included in the integrative review need to be divided into subgroups according to some logical system*
- *the studies should be divided into subgroups according to a previously established classification,*

EXTRACTING THEME– 4 items

Were the relevant data extracted from all sources using a pre-determined extraction instrument (for all subgroups, if applicable)?

DESCRIPTION: Relevant data should be defined by the authors in relation to the problem, purpose, and research questions guiding the review. An explicit statement of designing the extraction instrument before extraction begins will strengthen the review. If applicable, a mention of data extracted for all subgroups should be present.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
The relevant data were not extracted from all sources and/or were not extracted for all subgroups (if applicable).	It is unclear if all relevant data were extracted from all sources for all subgroups (if applicable), and there may also be no evidence that a pre-determined extraction instrument was used.	The relevant data were extracted from all sources for all subgroups (if applicable), but there is no evidence that a pre-determined extraction instrument was used.	The relevant data were extracted from all sources for all subgroups (if applicable), and there is evidence that a pre-determined extraction instrument was used.

data_extraction : 9 references coded, 2.59% coverage

- *common data are extracted from primary studies for subsequent analysis*
- *data abstraction*
- *data are extracted from the primary sources using the prepared instrument to simplify, summarize and organize the findings,*
- *Extract relevant information from included articles*
- *order, code, and categorize data from multiple sources that may have used diverse methodological perspectives (Whittemore and Knafl 2005; Cooper 1998)*
- *Predetermined and relevant data of each subgroup classification are extracted from all primary data sources and compiled into a matrix or spreadsheet (Miles & Huberman 1994, Garrard 2004).*
- *selecting, focusing, simplifying, and abstracting data from the sample of primary sources (Whittemore and Knafl 2005; Miles and Huberman 1994a)*
- *The aim of the analysis phase is to ensure complete documentation of all relevant data from each included study.*
- *the data from primary sources are ordered, coded, categorized, and summarized into a unified and integrated conclusion about the research problem (Cooper 1998)*

extraction_criteria : 7 references coded, 2.01% coverage

- *A Table of Inclusion should be developed that includes information regarding the paper (author, year, title, journal), purpose of the study, study design, description of the sample, data analysis, major findings, and conclusions.*
- *Details of studies and articles included in review*

- *Extracted data should include definition of the subjects, methodology, size of the sample, variables, method of analysis, and results. [Shuler in Holly, 2014]*
- *Extracted data should include definition of the subjects, methodology, size of the sample, variables, method of analysis, and results. [Holly, 2019]*
- *methodological features of primary studies are also extracted to evaluate the overall quality of individual studies*
- *The data should include definition of the subjects, methodology, size of the sample, measuring variables, method of analysis and concepts used as bases*
- *the information should cover the study sample (subjects), the objectives, the methodology used, results and the main conclusions of each study.14 [Broome, 2000]*

extraction_tool : 8 references coded, 2.30% coverage

- *a data collection tool should be developed*
- *Adopt a data collection tool*
- *data are extracted from the primary sources using the prepared instrument to simplify, summarize and organize the findings,*
- *Data extraction forms are ideally piloted with several primary sources and revised to ensure an accurate portrayal of all primary sources.*
- *defining the information to be extracted from the selected studies, using an instrument to gather and synthesize key information. 12 [Beyea & Nicoll, 1998]*
- *Extract relevant information from included articles*
- *Revise data collection tool to fit review purpose*
- *To extract data from the articles selected, it is necessary to use a previously prepared instrument*

Is the data extraction process explicit, unbiased, and reproducible?

DESCRIPTION: Like the information retrieval/search process, explicit detail should be presented about the data extraction process so that it could be reproduced, with an emphasis on minimizing transcription error. An unbiased data extraction process would involve more than one reviewer extracting and/or verifying data.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
The data extraction process is not described explicitly so that it is reproducible.	The data extraction process was completed by more than one person but is not sufficiently described so that it is reproducible.	The data extraction process was completed by only one person and is described explicitly so that it is reproducible.	The data extraction process was completed by more than one person and is described explicitly so that it is reproducible.

process : 9 references coded, 2.59% coverage

- *Approach to collection of data from articles and study reports*
- *How the process was managed*
- *more than one reviewer abstract data and to do so independently*
- *Panel members read each report and extract basic information about design, sample, variables, and results. They then enter the information into tables and create list to help them identify differences, commonality's, and patterns across the studies.*
- *patterns and processes are isolated*
- *Strategies used to minimize transcription error*
- *The data extraction process should be explicit, unbiased, and reproducible (Friedland et al., 1998).*
- *The process used to extract information from the reports [is] also described.*

- *To minimize the risk of transcription error, it is common practice to have two reviewers extract the data from studies in dependently.*

Were data analyzed using a review matrix?

DESCRIPTION: A review matrix arranges the extracted data from all sources into columns and rows for a visual representation of data and as a means of arranging data for further analysis. If using subgroups, several matrices may be presented as opposed to one single matrix.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Data were not analyzed using a review matrix.	Data matrix is not readily available or is provided only via online supplement not accessible to reader.	Data were analyzed using a review matrix, but not all extracted data (or subgroups, if applicable) are present.	Data were analyzed using a review matrix, and all extracted data (including subgroups, if applicable) are present.

matrix : 9 references coded, 2.59% coverage

- *A table of key elements of each individual study ideally is included and enough detail is provided to evaluate that conclusions are based on the evidence.*
- *Categories and data-display matrices are then developed to display all of the coded data from each study by its category.*
- *Categories and data-display matrices are then developed to display all of the coded data from each study by its category.*
- *creation of a review matrix*
- *extract data into matrices (tables)*
- *Panel members read each report and extract basic information about design, sample, variables, and results. They then enter the information into tables and*

create list to help them identify differences, commonality's, and patterns across the studies.

- *using tables to organize data in a clear and concise format*
- *When the critical appraisal has been completed, a table can be created that contains information about each of the studies in the review.*
- *When the critical appraisal has been completed, a table can be created that contains information about each of the studies in the review.*

Is a data display assembling the data from retrieved sources provided?

DESCRIPTION: Examples of data displays may include (but are not limited to): matrices, graphs, charts, networks, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no visualization of data extracted from retrieved sources.	Data display is not readily available or is provided only via online supplement not accessible to reader.	Only partial data extracted from retrieved sources are displayed visually.	All extracted data from all subgroups (if applicable) are provided via visual data display.

data_display : 8 references coded, 2.30% coverage

- *converting the extracted data from individual sources into a display that assembles the data from multiple primary sources around particular variables or subgroups*
- *data displays can be in the form of matrices, graphs, charts, or networks*
- *different data displays are likely to be required for each subgroup classification*
- *grafts, charts, networks, or, as most frequently used, matrices (Miles and Huberman 1994a)*

- *graphic displays of data are also included (i.e., plotting quality of studies with a particular outcome)*
- *organize and summarize the information in a concise manner, forming a database that is easy to access and manage*
- *Pre-determined relevant data from each primary source need to be extracted, preferably using tables, charts or matrices (Miles & Huberman, 1994; Garrarand [sic], 1999).*
- *Predetermined and relevant data of each subgroup classification are extracted from all primary data sources and compiled into a matrix or spreadsheet (Miles & Huberman 1994, Garrard 2004).*

VERIFYING THEME – 6 items

Was a systematic analytic method explicitly identified?

DESCRIPTION: There is an explicit statement of the systematic analytic method (such as constant comparison, content analysis, or thematic analysis, etc.) used for data analysis.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
A systematic analytic method is not named and is not vaguely described.	Analysis is only vaguely described or may be inferred through context but is not explicitly identified by name.	A systematic analytic method process is thoroughly described but is not explicitly named.	A systematic analytic method is explicitly identified by name (constant comparison, content analysis, thematic analysis, etc.), and the process is thoroughly described.

analysis : 26 references coded, 7.47% coverage

- *a systematic analytic [data analysis] method should be explicitly identified before undertaking the review*

- *A systematic method for data analysis should be identified explicitly before undertaking integrative review.*
- *coding, coding sheets, grouping, categorization, and abstraction (Elo and Kynga 2008)*
- *constant comparison, content analysis, and thematic analysis are commonly used approaches in IRs (Hopia et al. 2016)*
- *Cooper (1998) in an old, but still relevant publication, reminds us that this step reduces individual data points for aggregation of findings.*
- *Data analysis in research reviews requires that the investigator order, categorize, and summarize data from primary individual studies into a unified conclusion about the research problem (Cooper, 1998).*
- *Data collection in integrative reviews requires a standard and thorough examination of each primary source (Cooper, 1998).*
- *Discuss and interpret data*
- *examining the data displays for patterns, themes, commonalities, and differences across the review sample (Whittemore and Knafl 2005)*
- *Explicit records of all coding structures and data analysis decisions must be maintained to increase the rigor and the transparency of the process.*
- *familiarizing yourself with your data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report (Braun and Clarke 2006)*
- *In presenting the results, the goal is to make clear how the data were collected, how the analysis was carried out, and how the conclusions were derived from the data analysis.*

- *provide explicit details during dissemination about approaches used during the data analysis stage*
- *reviewers should look to use inductive analysis approaches generally associated with qualitative or mixed-method research (Whittemore and Knafl 2005)*
- *Set rules of inference for data analysis and interpretation*
- *Systematically analyse data*
- *The analysis stage should be explicit, with conclusions demonstrating thoughtful consideration of all studies.*
- *The approach to analyzing the data and generating themes or categories should be described. [Shuler in Holly, 2014]*
- *The approach to analyzing the data and generating themes or categories should be described. [Holly, 2019]*
- *The constant comparison method consists of four phases: data reduction, data display, data comparison, and conclusion drawing and verification (Whittemore and Knafl 2005; Miles and Huberman 1994a)*
- *The goal of analysis is to reach conclusions that represent the findings of the individual studies as a body of findings, which is different from looking at each one in isolation from the others.*
- *The overall process involves data reduction, data display, data comparison, drawing conclusion, and verification (Whittemore & Knafl, 2005).*
- *The procedures for data analysis should also be reported in the methods section of an integrative review report.*
- *To further ensure accuracy, two independent reviewers may code individual primary sources.*

- *Whittemore & Knafl (2005) have specifically addressed the complexity of data analysis and integrative reviews and have proposed that the qualitative method developed by Miles & Huberman (1994) assists in obtaining thorough, integrated inaccurate conclusions. The method consists of data reduction, data display, dated comparison, conclusion drawing and verification (Table 11.2) (Miles & Huberman, 1994).*
- *"reducing the separate data points collected by The Enquirer into a unified statement about the research problem." [Cooper]*

Were commonalities and differences identified?

DESCRIPTION: Identification of commonalties and differences most commonly refers to an examination and comparison between retrieved sources but may also refer to identifying commonalties and differences between the retrieved sources and other extant materials, such as background literature or statistics.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Commonalities and differences are not identified and cannot be inferred from the text.	Commonalities and differences are identified may be inferred from the text but are not explicitly described.	Commonalities and differences are identified and present in the text but are not explicitly described.	Commonalities and differences are identified and explicitly described.

comparison : 10 references coded, 2.87% coverage

- *analyze for similarities and differences (patterns) in relation to the stated review purpose or questions*
- *clustering, counting, and making contrasts and comparisons*
- *Coded data of each individual study are compared with all other studies for similarities and differences around variables of interest.*

- *commonalities and differences are identified*
- *comparison*
- *Comparisons and contrasts are made of the findings of the review with background literature, and work of others*
- *data comparison which involves an iterative process of examining data displays of primary source data in order to identify patterns, themes, or relationships*
- *Similarities and differences in results and findings would be identified and reasons for the differences explored.*
- *Similarities and differences in the findings are identified and a description of generalizations representative of the defined categories and the integrative review in its entirety, as possible (Whittemore & Knafl , 2005 quoted by Shuler in Holly, 2014).*
- *Similarities and differences in the findings are identified with a description of generalizations representative of the defined categories and the integrative review in its entirety, as possible (Whittemore & Knafl , 2005 quoted by Holly, 2019).*

Was conflicting evidence addressed?

DESCRIPTION: Addressing conflicting evidence differs from looking at identified differences in that an examination of conflicts seeks to determine a reason or explanation for the differences.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Conflicting evidence is clearly present but is not addressed.	It cannot be determined if conflicting evidence is present. Or, conflicting evidence is present and	Conflicting evidence is examined with possible explanations for differences able to	Conflicting evidence is examined with possible explanations for differences explicitly stated.

	discussed but not adequately.	be inferred but not explicitly stated.	

conflicting_evidence : 6 references coded, 1.72% coverage

- *addressing conflicting evidence*
- *Data analysis procedures also need to include the opportunity for the analyst to identify conflicting evidence in tenuous results, as well as to consider several different explanations for interpreting the data.*
- *exploration of confounding influences*
- *In addition, reasons for variability are explored.*
- *Outliers are identified and examined for reasons for variability explored*
- *The analysis must be performed critically, looking for explanations for the different or conflicting results in differerent studies.12,16 [Beyea & Nicoll, 1998; Ganong, 1987]*

Was the quality of retrieved sources addressed in a meaningful way?

DESCRIPTION: Quality of retrieved sources must be addressed in a meaningful way, with attention paid to aligning assessment of quality with the methodology or research design of identified sources. Ideally, the methods and tools used to assess quality are explicitly stated. Thorough assessment of quality will also include an examination of risk of bias and study reliability, if applicable.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Quality of retrieved sources was not addressed.	Quality of retrieved sources was addressed but not in a meaningful way. Alignment of tools with research designs of sources may be inadequate	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to

	or may not be addressed.	assess quality. Risk of bias and study reliability were not assessed.	assess quality. Risk of bias and study reliability are also assessed.

critical_appraisal : 18 references coded, 5.17% coverage

- *consideration of the quality of primary sources in an integrative review is addressed in a meaningful way*
- *consideration of the quality of primary sources in an integrative review is addressed in a meaningful way*
- *critical appraisal of the retrieved studies*
- *each of the selected studies needs to be critically appraised*
- *each of the selected studies needs to be critically appraised*
- *it is essential that all evidence be assessed for quality before inclusion in the IR*
- *it is important that [gray literature] be critically appraised before inclusion in the IR, as the quality of gray literature is variable*
- *Notation of the results of the appraisal of each study should be entered into the matrix to support the credibility of the data analysis and findings of the review.*
- *Other textual papers should be assessed for aspects such as accuracy, objectivity, authority, evidence, and significance.*
- *Results of the critical appraisal process*
- *Reviewers should make judgments about the methodological strengths and weaknesses of all included studies before making inferences about the phenomenon of interest (Jackson 1980)*
- *the methods used to evaluate the quality of the studies [is] also described.*
- *The quality of primary studies should be evaluated and incorporated in the analysis and interpretation of findings (Conn & Rantz, 2003).*

- *The report of the review should clearly indicate how each article was appraised…*
- *the selected studies must be analysed in detail*
- *this [critical analysis] phase demands an organized approach to weigh rigor and characteristics of each study*
- *Unpublished research studies, such as dissertations, should be appraised using the appropriate critical appraisal tool as would be used for a published study.*
- *… , and what the results of the appraisal were.*

quality_criteria : 11 references coded, 3.16% coverage

- *Approach to appraisal of study or article quality*
- *Criteria used to determine quality*
- *how quality is evaluated in an integrative review will vary depending on the sampling frame*
- *it is important that [gray literature] be critically appraised before inclusion in the IR, as the quality of gray literature is variable*
- *methodological quality and relevance of selected literature are appraised*
- *Quality is determined by evaluating the internal validity of primary studies or "the extent to which study design, conduct, and analysis systematically avoid or minimize potential sources of bias" (Conn & Rantz, 2003, p. 323).*
- *Results of the selection process*
- *The inclusion of multiple research designs means that criteria are needed for each type of research, which serves to increase the complexity of the appraisal process.*
- *The level of evidence of the studies must be evaluated*

- *Unpublished research studies, such as dissertations, should be appraised using the appropriate critical appraisal tool as would be used for a published study.*
- *… what criteria were used to determine quality …*

risk_of_bias : 4 references coded, 1.15% coverage

- *Quality is determined by evaluating the internal validity of primary studies or "the extent to which study design, conduct, and analysis systematically avoid or minimize potential sources of bias" (Conn & Rantz, 2003, p. 323).*
- *The method of assessing the risk of bias should be transparent and reproducible.*
- *To further ensure accuracy and minimize bias in the critical appraisal process, it is preferable to have two reviewers independently review all literature for quality and relevance (Whittemore 2007).*
- *The constant comparison method consists of four phases: data reduction, data display, data comparison, and conclusion drawing and verification (Whittemore and Knafl 2005; Miles and Huberman 1994a)*

study_reliability : 2 references coded, 0.57% coverage

- *The reviewer evaluates the data for unreliable values, which might include whether the findings from one study or two different from the other studies to be considered, and whether data recording errors exist.*
- *The reviewer also must evaluate the reliability of each study's findings, both in and of itself and in comparison with all of the other studies included in the review.*

Was quality verified by two independent reviewers?

DESCRIPTION: The use of two independent raters is preferable as this ensures accuracy and reduces bias.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no evidence that quality was verified as part of the review.	It cannot be determined if quality was verified. Or, there is evidence that quality was verified but the number of reviewers is not stated.	There is evidence that quality was verified by one reviewer.	There is evidence that quality was verified by two reviewers. Inter-rater agreement calculations may have also been provided.

number_reviewers : 7 references coded, 2.01% coverage

- *A critical assessment of papers to be included in an integrated review may be done by one person…*
- *Critical appraisal for an integrative review can be completed by one investigator. [Shuler in Holly, 2014]*
- *Critical appraisal for an integrative review can be completed by one investigator. [Holly, 2019]*
- *To ensure accuracy, often two independent reviewers code individual studies for content and quality, and interrater agreement is calculated.*
- *To further ensure accuracy and minimize bias in the critical appraisal process, it is preferable to have two reviewers independently review all literature for quality and relevance (Whittemore 2007).*
- *To further ensure accuracy, two independent reviewers may code individual primary sources.*
- *Using a pre specified set of inclusion criteria, two persons (ideally) decide which studies qualify for full review.*

Were any articles excluded based on quality appraisal?

DESCRIPTION: Based on the outcome of quality appraisal, low quality articles may be excluded. Rationale for inclusion or exclusion of low quality should be explicltly stated. If low quality articles are excluded, it is ideal that a detailed log of excluded sources is provided with rationale explicitly stated.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Quality appraisal was not completed.	Quality appraisal was completed AND there were articles of low quality, but no articles were excluded based on quality appraisal nor was rationale for inclusion of low-quality articles provided.	Quality appraisal was completed AND there were articles of low quality which were excluded, but there is no detailed log of excluded articles nor rationale provided.	Quality appraisal was completed AND there were articles of low quality which were excluded, and there is a detailed log of excluded articles or rationale provided.

exclusion_quality : 10 references coded, 2.87% coverage

- *A detailed log of the papers determined not to be good enough should be kept and the reason for exclusion noted. [Shuler in Holly, 2014]*
- *A detailed log of the papers determined not to be good enough should be kept and the reason for exclusion noted.[Holly, 2019]*
- *Additionally, the report should identify whether any works were excluded based on the quality appraisal.*
- *critically judges whether the data element or result is worthy of remaining in the study data set*
- *If the studies are to be excluded for methodological quality, this should be stated with a clear description of the measures used to determine acceptable methodological quality.*

- *review panels decide how they will handle studies that are of dubious methodological quality.*
- *Studies not meeting the appraisal criteria are excluded. [Shuler in Holly, 2014]*
- *Studies not meeting the appraisal criteria are excluded. [Holly, 2019]*
- *the inclusion and exclusion criteria should identify whether inferior studies will be included or excluded after the appraisal process*
- *The reviewer must identify which studies were not able to be included, describe why they were not included, and then discuss how they might have shown different results from those selected.*

ENDING THEME – 5 items

Were patterns, themes, relationships, or conclusions verified with sources?

DESCRIPTION: Patterns, themes, relationships, or conclusions are clearly stated. Explicit statements of verification with retrieved sources will strengthen the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Patterns, themes, relationships, or conclusions are not identified.	Patterns, themes, relationships, or conclusions may be identified, but it is unclear whether conclusions are verified with sources.	Patterns, themes, relationships, or conclusions are identified, but verification with sources is not explicit but may be easily inferred.	Patterns, themes, relationships, or conclusions are identified and are explicitly verified with sources.

conclusions : 20 references coded, 5.75% coverage

- *A new conceptualization of the primary sources integrates all subgroups into a comprehensive portrayal of the topic of concern, thus completing the review process.*

- *As data are conceptualized at higher and higher levels of abstraction, each primary source is re-read and reviewed to verify that the new conceptualization was congruent with primary sources. [Shuler in Holly, 2014]*
- *As data are conceptualized at higher and higher levels of abstraction, each primary source is re-read and reviewed to verify that the new conceptualization was congruent with primary sources. [Holly, 2019]*
- *Conclusions must have clear supportive evidence from the findings of all individual studies (Moher, Jadad, & Klassen, 1998).*
- *Cooper (1998) in an old, but still relevant publication, reminds us that this step reduces individual data points for aggregation of findings.*
- *Drawing conclusions and summarizing results is the final step in an integrative review, and is generally done as a narrative. [Shuler in Holly, 2014]*
- *Drawing conclusions and summarizing results is the final step in an integrative review, and is generally done as a narrative. [Holly, 2019]*
- *explicit details from primary sources and evidence to support conclusions need to be provided*
- *Finally, in the Discussion section, inferences should be drawn and discussed in relation to the guiding questions. [Shuler in Holly, 2014]*
- *Finally, in the discussion section, inferences should be drawn and discussed in relation to the guiding questions. [Holly, 2019]*
- *Findings can be presented as a research agenda that poses new questions for investigation or a taxonomy that can be used to classify previous research.*
- *Findings or key issues from individual papers*
- *identification of patterns, differences and redistribution of these topics may be included as part of the general discussion.*

- *In this stage, based on interpretation and synthesis of results, the data demonstrated in the analysis of the articles are compared to the theoretical reference.*
- *include a concise summary of their major findings and key contributions to the state of science*
- *patterns, themes, relationships, or conclusions requires verification with primary source data for accuracy and confirmability (Miles & Huberman 1994)*
- *The panel or authors indicate whether and how their conclusions square with any prior summarization work that has been done on the topic, summarized the limitations of the body of research, and offer opinions regarding the ramifications of the conclusions.*
- *the results capture the depth and breadth of the topic and contribute to a new understanding of the phenomenon of concern*
- *The Results section should summarize and synthesize the themes that were uncovered in the review.*
- *This should be a fully synthesized report showing the themes that emerged across the included studies.*

Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?

DESCRIPTION: The review is concluded by explicitly synthesizing elements or conclusions into a summation of the topic or phenomenon. Merely listing identified elements is not an explicit means of synthesis. If applicable, synthesized elements or conclusions are present for all subgroups.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES

Important elements or conclusions are not synthesized into a summation of the topic or phenomenon.	Important elements or conclusions are explicitly listed or discussed separately but are not clearly synthesized into a summation of the topic or phenomenon.	Important elements or conclusions are presented in a manner that is more consistent with listing than synthesis or summation of the topic or phenomenon. Or, if applicable, synthesized summations are not present for all subgroups.	Important elements or conclusions are explicitly synthesized into a summation of the topic or phenomenon. If applicable, synthesized summations are present for all subgroups.

synthesis : 18 references coded, 5.17% coverage

- *a narrative or thematic synthesis*
- *a thematic synthesis for each of the stated review questions*
- *Approach to data synthesis*
- *Cooper (1998) in an old, but still relevant publication, reminds us that this step reduces individual data points for aggregation of findings.*
- *Data for an integrative review are generally analyzed qualitatively by using a thematic synthesis approach, which arranges studies into common areas or themes that can be summarized.*
- *familiarizing yourself with your data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report (Braun and Clarke 2006)*
- *Findings or key issues from individual papers*
- *patterns are then synthesized*
- *Results of the synthesis process*
- *Specific methods used for the synthesis of data*

- *Studies ideally are assessed for their ability to be combined logically or statistically (heterogeneity tests), or both ways.*
- *synthesis*
- *Synthesis within themes continues to be the most common approach to present the results of IRs (Toronto et al. 2018). These themes are developed during the data analysis stage and used as an organizing structure in the results section.*
- *The Results section should summarize and synthesize the themes that were uncovered in the review.*
- *The results section should summarize and synthesize the themes that were uncovered in the review.*
- *the synthesis of important elements or conclusions of each subgroup into an integrated summation of the topic or phenomenon*
- *This should be a fully synthesized report showing the themes derived across studies.*
- *This should be a fully synthesized report showing the themes that emerged across the included studies.*

Were the conclusions reported in table or diagrammatic form?

DESCRIPTION: Tables or diagrammatic forms of conclusions differ from visualizations of extracted data in that they represent the synthesis or outcomes of the review. Examples of visualization of conclusions may include (but are not limited to): matrices, graphs, charts, diagrams, networks, joint displays, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Conclusions are not reported in table or diagrammatic form.	Table(s) or diagrammatic form(s) of conclusions are not	Conclusions are reported in table or diagrammatic form but not explicitly in a	Conclusions are reported in table or diagrammatic form in a meaningful way

	readily available or are provided only via online supplement not accessible to reader.	meaningful way that contributes to new knowledge or understanding of the topic or phenomenon.	that contributes to new knowledge or understanding of the topic or phenomenon.

visualization : 8 references coded, 2.30% coverage

- *Conclusions of integrative reviews can be reported in table or diagrammatic form.*
- *Findings can be presented as a research agenda, a conceptual model, or a taxonomy (see Table 8.2).*
- *Visual aids such as tables or schematic representations of the relationships among codes, concepts, and ideas often are employed (Miles & Huberman, 1994).*

- *conceptual_model : 3 references coded, 0.86% coverage*
- *Conclusions or conceptual models that are developed are continually revised in order to be inclusive of as much data as possible (Miles & Huberman 1994).*
- *generates a new model, conceptual framework, or other unique conception informed by the author's intimate knowledge of the topic*
- *models, conceptual systems, conceptual mapping, or categories (Elo and Kynga 2008)*

tables : 2 references coded, 0.57% coverage

- *tables to demonstrate the linkages between the data, categories created, and final results*

- *tables document the evidence of primary sources … multiple tables are often used to show the logical sequence of analysis, explicitly linking conclusions to data displays.*

Were implications discussed for: research, practice, education, or policy?

DESCRIPTION: Implications for research, practice, education, or policy are explicitly stated. Reviews may be strengthened by directly aligning implications with both the purpose and synthesized findings of the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Implications for research, practice, education, or policy are not explicitly stated nor may they be inferred from the text.	Implications for research, practice, education, or policy may be inferred but do not clearly align to the purpose of the review or its findings.	Implications for research, practice, education, or policy and alignment with review purpose may be inferred but are not explicitly stated. Or, implications for research, practice, education, or policy are stated but do not clearly align to the purpose of the review or its findings.	Implications for research, practice, education, or policy are explicitly stated and discussed, with clear alignment with purpose and findings of the review.

implications : 6 references coded, 1.72% coverage

- *Cooper (1998) in an old, but still relevant publication, reminds us that this step reduces individual data points for aggregation of findings.*
- *Findings can be presented as a research agenda that poses new questions for investigation or a taxonomy that can be used to classify previous research.*
- *importance of disseminating the findings is also addressed*
- *interpretation of findings and implications for research; practice; and policy are presented*

- *Recommendations and implications for research, practice, education, theory, and policy when applicable are made*
- *Results of the synthesis process*

Were methodological or other limitations of the review explicitly stated?

DESCRIPTION: Methodological limitations of the review may include (but are not limited to): only one individual for screening, quality appraisal, thematic analysis, etc.; limiting to English-language or recent years with no rationale; limited access to databases or full text resources for retrieval of sources, etc. Methodological limitations of the review do not include limitations of the retrieved sources, which is more appropriately discussed during quality appraisal.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There are no methodological limitations listed.	Limitations are only mentioned regarding retrieved sources, not methodological limitations of the review itself.	Methodological limitations of the review may be inferred but are not explicitly stated. Or, methodological limitations of the review are present but are not discussed.	All methodological or other limitations of the review are explicitly stated.

limitations : 2 references coded, 0.57% coverage

- *limitations of the review are presented*
- *comment on methodological limitations of their review*

Appendix F: *Non-Human Subjects Designation for Face and Content Validity*

University of CINCINNATI
Institutional Review Board
FWA #: 000003152

NOT HUMAN RESEARCH DETERMINATION

January 19, 2022

Emily Kean
CON Adjunct Faculty

Dear Emily Kean,

Type of Submission:	Initial Study
Title of Study:	Development and Expert Testing of Face and Content Validity of the IRMAT: Integrative Review Methodology Appraisal Tool
Investigator:	Emily Kean
IRB ID:	2022-0025
Funding:	None
Documents Reviewed:	• 2022-01-11 HRP-503N TEMPLATE Not Human Subjects rev 5MAY2021.docx

On **1/19/2022**, the IRB reviewed the above submission and determined that the proposed activity is not research involving human subjects as defined by DHHS, DOJ and FDA regulations.

IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities are research involving human in which the organization is engaged, please submit a new request to the IRB for a determination. You can create a modification by clicking **Create Modification/CR** within the study.

Statement regarding International Conference on Harmonization and Good Clinical Practices
The Institutional Review Board is duly constituted (fulfilling FDA requirements for diversity), has written procedures for initial and continuing review of clinical trials: prepares written minutes of convened meetings and retains records pertaining to the review and approval process all in compliance with requirements defined in 21 CFR Parts 50, 56 and 312 Code of Federal Regulations. This institution is in compliance with the ICH GCP as adopted by FDA/DHHS.

Please note: This review is through the IRB only. You may be responsible for reporting to other regulatory officials. Please check with your institution and department to ensure you have met all reporting requirements.

Thank you for your cooperation during the review process.

v 2.12.20

Appendix G: *IRMAT Items (V 1.2) with Expert Ratings and Feedback*

STARTING THEME – 7 items (Previously 9 items)

DELETED: Was the protocol described?

DESCRIPTION: A protocol is a predetermined documented plan for the review, including the rationale, research question, and proposed methods for the review. Detailed protocols should be developed *a priori* and made publicly available prior to carrying out the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no description of a protocol.	A protocol is not explicitly described; however, there are indicators (description of methods established *a priori,* etc.) that a protocol may have been completed. Or the protocol is provided only via online supplement not accessible to reader.	A protocol is explicitly referenced, but there is no description nor link provided.	A protocol is referenced, and there is a clear description provided. Or there is a link to an accessible online supplement or protocol registry.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant 2 = Somewhat Relevant 2 = Somewhat Relevant 2 = Somewhat Relevant 4 = Highly Relevant	4 = Highly Relevant 2 = Somewhat Relevant 3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant
0.40	**0.60**

EXPERT FEEDBACK

Face Validity

My understanding is a protocol is for scoping or systematic reviews. I have never heard of a formal protocol for an integrative review (IR) "a priori" usually is not an element of an IR

I think is relevant for some reviews but not necessary for all reviews.

Wasn't sure how to rate because this is 2 criteria - protocol specification and protocol registration. Think this should be 2 items. Was the protocol described? Was the protocol registered?

Content Validity

I do not agree with the premise that an IR needs a protocol it is iterative method instead.

REVISIONS & COMMENTS

Item is deleted.

RETAINED WITH MODIFICATIONS: Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?

DESCRIPTION: To identify the gap in the literature and necessity for the review, a preliminary background literature review should be conducted.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clear statement of a preliminary background literature review being conducted nor a reference to the gaps in the literature for this review topic.	There may be references to gaps in the literature for this review topic, but there is no statement of preliminary literature reviews being conducted.	There are references to preliminary literature search(es) being conducted, but there is no clear statement that the preliminary searches were used to identify the gap.	There is a clear statement that a preliminary background literature review was conducted to identify the gap in the literature for this review topic.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 2 = Somewhat Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

My understanding of the introduction or background section is that this step is implied and does not require a clear statement that a preliminary background literature review is being conducted.

Very clearly written

Consider that in the review of the literature, other reviews on the topic were synthesized and the need for this review is justified. ? should there be a rationale for why an integrative review vs. other types of reviews.

Content Validity

REVISIONS & COMMENTS

A statement was added to the item description to address the expert comments about previously existing comprehensive literature reviews or evidence syntheses articles. Regarding the expert comment about including rationale for why an integrative review was chosen over other review types, this does not appear in any of the thirteen methodological sources used as the basis for item generation and is thus not included.

REVISED ITEM: Was there a statement of a preliminary background literature review being conducted as rationale for undertaking the review?

DESCRIPTION: To identify the gap in the literature and necessity for the review, a preliminary background literature review should be conducted. The statement of a preliminary background literature review being conducted may also include a reference to whether prior comprehensive literature reviews or evidence syntheses are available in the literature.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clear statement of a preliminary background literature review being conducted nor a reference to the gaps in the literature for this review topic.	There may be references to gaps in the literature for this review topic, but there is no statement of preliminary literature reviews being conducted.	There are references to preliminary literature search(es) being conducted, but there is no clear statement that the preliminary searches were used to identify the gap.	There is a clear statement that a preliminary background literature review was conducted to identify the gap in the literature for this review topic.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
2 = Somewhat Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
0.80	**1.00**

RETAINED WITH MODIFICATIONS: Were the variables of interest clearly stated?

DESCRIPTION: Variables of interest may be defined as concepts, target population, health care problem, etc. Variables may or may not be defined in terms of PICO(S) elements (population, intervention, comparison, outcome, setting).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of variables of interest, and variables cannot be inferred.	Variables of interest are not explicitly stated but may be inferred.	There is an explicit statement of the variables of interest, but variables are not clearly defined within specific contexts (population, outcomes, etc.).	There is an explicit statement of the variables of interest, and variables are clearly defined within specific contexts (population, outcomes, etc.).

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
2 = Somewhat Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	2 = Somewhat Relevant
4 = Highly Relevant	4 = Highly Relevant
0.80	**0.80**

EXPERT FEEDBACK

Face Validity

My understanding is that a systematic review would use the PICOS format not an IR. This is one of their distinguishing differences. I agree that key concepts such as problem, population and concepts need to be defined but an IR looks at a phenomenon broadly not narrowly as this question suggests.

No - very clearly written

I think the word "variables" is misleading and implies that integrative reviews focus on quantitative studies. Possibly "factors" which could be defined as concepts, variables, etc.

Content Validity

Variables may not be the starting point

REVISIONS & COMMENTS

Regarding the expert comment that integrative reviews may not use the PICOS format, it is noted that the item description is worded such that use of PICO(S) may (or may not) be used. The word “variables” has been replaced with “factors” in response to expert comment and to increase clarity for this item. The expert comment that “Variables may not be the starting point” is slightly unclear but is interpreted to mean that this item may not properly align with the Starting domain; however, it is believed that the clarification that this item refers to “factors,” not solely “variables” as commonly defined in quantitative literature helps to clarify that this item is applicable to the Starting domain.

REVISED ITEM: Were the factors of interest (concepts, variables, etc.) clearly stated?

DESCRIPTION: Factors of interest may be defined as concepts, variables, target population, health care problem, etc. Factors may or may not be defined in terms of PICO(S) elements (population, intervention, comparison, outcome, setting).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of factors of interest, and variables cannot be inferred.	Factors of interest are not explicitly stated but may be inferred.	There is an explicit statement of the factors of interest, but factors are not clearly defined within specific contexts (population, outcomes, etc.).	There is an explicit statement of the factors of interest, and factors are clearly defined within specific contexts (population, outcomes, etc.).

RETAINED WITH MODIFICATIONS: Were operational and conceptual definitions of variables provided?

DESCRIPTION: Operational definitions define variables in terms of how they are measured. Conceptual definitions are the abstract or theoretical meanings of variables.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Neither operational nor conceptual definitions are provided.	It is unclear whether operational and conceptual definitions are provided, or definitions are provided but not for all relevant variables.	Only operational -or- only conceptual definitions are provided.	Both operational and conceptual definitions are provided for all variables.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant 4 = Highly Relevant
1.00	**0.80**

EXPERT FEEDBACK

Face Validity

I would highlight conceptual definitions more than operational since that is not typically a focus of an IR which does not perform a meta-analysis

Again, very clear with no ambiguity.

Think this is highly relevant, but same comment about use of the term "variable"

Could it be that some integrative reviews are more conceptual than empirical...in that case you may not have operational definitions.

Content Validity

Variables may not be the starting point

REVISIONS & COMMENTS

The word “variables” has been replaced with “factors” in response to the expert comment and to increase clarity for this item. The order of operational and conceptual has been reversed in the item to address experts’ comments that inclusion of operational definitions may not be necessary. The “and” operand has been modified to “and/or” to address these comments as well. Again, the expert comment that “Variables may not be the starting point” is slightly unclear but is interpreted to mean that this item may not properly align with the Starting domain; however, it is assumed that the clarification that this item refers to “factors” – not solely data extraction “variables” as commonly defined in quantitative literature – helps to clarify that this item is applicable to the Starting domain.

REVISED ITEM: Were conceptual and/or operational definitions of factors provided?

DESCRIPTION: Conceptual definitions are the abstract or theoretical meanings of factors. Operational definitions define factors in terms of how they are measured. Depending on the research question addressed by the integrative review, supplying operational definitions may not be necessary.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Neither operational nor conceptual definitions are provided.	It is unclear whether operational and conceptual definitions are provided.	Operational and/or conceptual definitions are provided but only for one factor.	Operational and/or conceptual definitions are provided for all factors.

RETAINED WITH MODIFICATIONS: Was the problem stated unambiguously and was it easy to identify?

DESCRIPTION: A problem statement outlines the areas under examination by the review. An easily identifiable problem statement may begin with a statement such as “This review aims to address the problem of…”, or “Previous literature searches highlighted problems surrounding…”, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clearly stated problem.	A problem statement is alluded to but is not clearly stated.	The problem statement is clear but is not easily identifiable.	The problem statement is clear and is easily identifiable.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

How is problem statement different from the review of the literature that identifies gaps.

No [recommendations for this item]

This is a "spot on" criterion, but think it is confusing to state "this review aims to ..." because the next criterion addresses aim. I think criterion 5 and 6 should more clearly differentiate identifying the problem being addressed from the more specific aims of the review.

Content Validity

REVISIONS & COMMENTS

A clarifying statement has been added to the item description to address the expert questions of how a problem statement differs from the preliminary literature review. The exemplary sentence has been modified in response to the expert suggestion to further distinguish the item on problem statements from the following item on the purpose and aim of the review.

REVISED ITEM: Was the problem stated unambiguously and was it easy to identify?

DESCRIPTION: A problem statement outlines the areas under examination by the review. Where the statement of a preliminary search describes the action which was taken to identify the need for the integrative review, the problem statement is an explicit statement of the

problem the review seeks to address. An easily identifiable problem statement may begin with a statement such as “This review addresses the problem of…”, or “Previous literature reviews highlighted problems surrounding…”, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clearly stated problem.	A problem statement is alluded to but is not clearly stated.	The problem statement is clear but is not easily identifiable.	The problem statement is clear and is easily identifiable.

RETAINED AS IS: Was there an explicit statement of the purpose or aim of the review?

DESCRIPTION: The purpose and aim of the review differs from the problem statement in that the purpose/aim are actionable items and/or address the problem in some way. An explicit statement of the purpose/aim may commonly be represented by a phrase such as “The purpose/aim of this review is…”.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of the purpose or aim of the review, and it cannot be inferred from other context.	There is no explicit statement of the purpose or aim of the review, but it may be inferred from other context.	The purpose or aim is clear but is not explicitly stated.	The purpose or aim is clear and is explicitly stated.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	3 = Quite Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

No

See comment for #5 [This is a "spot on" criterion, but think it is confusing to state "this review aims to ..." because the next criterion addresses aim. I think criterion 5 and 6 should more clearly differentiate identifying the problem being addressed from the more specific aims of the review.]

Content Validity

REVISIONS & COMMENTS

The exemplary sentence in the previous item addressing problem statements has been modified in response to the expert suggestion to further distinguish that item on problem statements from this item on the purpose and aim of the review. This item has not been modified.

RETAINED AS IS: Was there an explicit statement of the research question(s) the review addresses?

DESCRIPTION: The research question(s) of the review differ from the problem statement and the purpose/aim in that the research question(s) are written in question form. An explicit statement of the research questions may commonly be represented by a phrase such as “The research questions guiding this review are…”.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of the research question(s) guiding the review, and it cannot be inferred from other context.	There is no explicit statement of the research question(s) guiding the review, but it may be inferred from other context.	The research question(s) guiding the review are clear but are not explicitly stated in question format.	The research question(s) guiding the review are clear and are explicitly stated in question format.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	3 = Quite Relevant
4 = Highly Relevant	4 = Highly Relevant

2 = Somewhat Relevant	2 = Somewhat Relevant
0.80	**0.80**

EXPERT FEEDBACK

Face Validity

Having a purpose, aim and research question seems redundant to me.

Nothing to add.

Given the space constraints of most journal manuscripts, I question whether it is realistic to expect a review to include both aims and research questions. Having reviewed and published multiple synthesis reports, I think it's acceptable to state something along the line of "the aim/purpose of this review is to address the following research questions. Of to state "the synthesis addressed the following aims".

Content Validity

seems redundant to #6

REVISIONS & COMMENTS

Although this item meets the statistical criteria for inclusion in the revised tool, it is noted that numerous experts commented on the redundant nature of this item as it relates to the previous two items; item is retained for further testing.

RETAINED WITH MODIFICATIONS: Was the sampling frame clearly stated?

DESCRIPTION: A sampling frame for an integrative review may include (but is not limited to): the type of empirical studies, specific research design(s), inclusion of methodological or theoretical literature/framework, etc. Providing a rationale for choice of sampling frame strengthens the choice of using a sampling frame. Attention should also be paid to the alignment of the choice of sampling frame with the integrative review methodology.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES

There is no clear statement of sampling frame nor is rationale provided. Or, the sampling frame does not align with the integrative review methodology, i.e., limiting to only randomized controlled trials, where a systematic review would be more appropriate.	There is no clear statement of whether a sampling frame was used, nor is rationale provided; however, a sampling frame may be alluded to in the screening (inclusion/exclusion) or data analysis descriptions.	A sampling frame is clearly stated but no rationale is provided.	A sampling frame with rationale is clearly stated. If no sampling frame is used, this is also clearly stated.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant	4 = Highly Relevant
2 = Somewhat Relevant	4 = Highly Relevant
4 = Highly Relevant	3 = Quite Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

this question seems premature at this stage and typically is addressed in the method section of an IR

Nothing more to add.

I think it's important to specify inclusion and exclusion criteria of that is what's meant by sampling frame. I think the following statement is unclear 'Attention should also be paid to the alignment of the choice of sampling frame with the integrative review methodology.' Does methodology mean consistent to using the integrative review approach to knowledge synthesis or does it refer to how the methods used in the review undertaken? Why does "yes" indicate that it's OK not use a sampling as long as you say you didn't?

Content Validity

REVISIONS & COMMENTS

The expert comment about this item better aligning with the method section is noted; however, the use of a sampling frame is best determined during the starting phase of the review as the decision to use a sampling frame will directly impact the choices made during the search component of the integrative review process. Greater clarification on the statement regarding alignment with integrative review methodology has been added to the item description. Based on expert feedback, this item has also been reworded to indicate that it may not be applicable, and the rating scale has been modified to reflect a N/A option.

REVISED ITEM: If applicable to the purpose and/or type of literature included, was the sampling frame clearly stated?

DESCRIPTION: A sampling frame for an integrative review may include (but is not limited to): the type of empirical studies, specific research design(s), inclusion of methodological or theoretical literature/framework, etc. Providing a rationale for choice of sampling frame strengthens the choice of using a sampling frame. Attention should also be paid to the alignment of the choice of sampling frame with the integrative review methodology. For example, the choice to use a sampling frame which limits the retrieved sources to only quantitative results does not align with the choice to use an integrative review methodology, which implies inclusion of varying result types (empirical [quantitative, qualitative, mixed methods], methodological, or theoretical).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES	N/A
There is no clear statement of sampling frame nor is rationale provided. Or, the sampling frame does not align with the integrative review methodology, i.e., limiting to only randomized controlled trials,	There is no clear statement of whether a sampling frame was used, nor is rationale provided; however, a sampling frame may be alluded to in the screening (inclusion/exclusion) or data analysis descriptions.	A sampling frame is clearly stated but no rationale is provided.	A sampling frame with rationale is clearly stated.	The use of a sampling frame is not applicable to the review purpose and/or type of literature included.

where a systematic review would be more appropriate.				

DELETED: Is a theoretical or conceptual framework used to guide the review?

DESCRIPTION: A theoretical or conceptual framework (also called theoretical or conceptual models) are established descriptions of relationships between concepts and are often represented by visualizations or diagrams. The use of a theoretical or conceptual framework to guide the review is one of the elements that may distinguish integrative reviews from other review types (systematic, scoping, etc.).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no statement of a theoretical or conceptual model being used to guide the review.	A theoretical or conceptual framework may be inferred but is not clearly stated. It is also unclear how it is guiding or shaping the review.	A theoretical or conceptual framework is clearly stated, but it is unclear how it is guiding or shaping the review.	A theoretical or conceptual framework is clearly stated and is explicitly used throughout the review (screening, data extraction, synthesis, etc.).

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
1 = Not Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 2 = Somewhat Relevant	1 = Not Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 4 = Highly Relevant
0.40	**0.40**

EXPERT FEEDBACK

Face Validity

Theoretical model is not always necessary in an IR. Consider rewording stating that" If a theoretical or conceptual model is used....

Suggest deleting this one. Integrative reviews can be more exploratory. They don't have to be theory building. You may want to consider differentiating two levels of criteria - Basic (relevant to all integrative reviews) and optional (applicable to some but not all - would depend of review aims/questions)

While integrative reviews may have a conceptual framework, i don't think that they all need one. maybe have this as appropriate.

This is an important element of an IR but often not addressed.

No [recommendations for this item]

Content Validity

Not all integrative reviews will require a theoretical framework

Review may be exploratory and not based on a conceptual framework

REVISIONS & COMMENTS

Item is deleted.

SEARCHING THEME: 5 items (Previously 5 items)

RETAINED WITH MODIFICATIONS: Is a description of a comprehensive search provided?

DESCRIPTION: A comprehensive search consists of a minimum of two to three search strategies (databases, gray literature, forward/backward searching, etc.) and a minimum of two databases, with three to six databases being preferred.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit description of a search process, nor can it be inferred from the text.	There is a limited description of a search process which only includes: • 1 search strategy **AND**	There is an explicit description of a search process which includes: • 2-3 search strategies	There is an explicit description of a search process which includes: • 2-3 search strategies

	• 1+ database(s)	**AND** • 1-2 databases	**AND** • 3-6+ databases

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 2 = Somewhat Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

what about grey literature?

I think a description of the search strategy is important, but suggest not linking to specific number of strategies and data bases. Could just say search strategy is strengthened by the use of multiple data bases and strategies

Content Validity

REVISIONS & COMMENTS

In response to the expert question about gray literature, gray literature is included in the item description, and the inclusion of non-database search strategies, including gray literature, is addressed more comprehensively by another item under the Searching domain. The rating scale criteria have been simplified in response to expert comments.

REVISED ITEM: Is a description of a comprehensive search provided?

DESCRIPTION: A comprehensive search consists of a minimum of two to three search strategies (databases, gray or non-academic literature, ancestry/descendancy [forward/backward] searching, hand searching, networking, searching research registries, etc.). Additionally, a comprehensive search is strengthened by searching three or more databases.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit description of a search process, nor	The description of the search process is limited. Or, only	There is an explicit description of a comprehensive	There is an explicit description of a comprehensive

can it be inferred from the text.	one search strategy (only searching databases, for example) is stated.	search process which includes: • 2 or more search strategies	search process which includes: • 2 or more search strategies Additionally, one of the search strategies includes an explicit statement of: • 2 or more databases

---

RETAINED WITH MODIFICATIONS: Were detailed descriptions of the database search processes provided?

DESCRIPTION: A detailed description of a database search will include (at a minimum): database name(s), keywords, whether subject headings were used, any applied limits (year, language), etc. A more comprehensive description will include (at a minimum): database name AND vendor, keywords AND variations, which controlled vocabularies were used, list of all limiters WITH rationale, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no description of the database search process. There may be a statement such as “Two databases were searched…”, but the database names, keywords, etc. are not provided.	There is a minimal description of the database search process provided, which may include: • database name(s) OR • keywords OR • whether or not subject headings were used OR • any applied limits, etc.	The following are explicitly stated: • database name(s) AND • keywords AND • whether or not subject headings were used AND • any applied limits, etc.	The following are explicitly stated: • database name(s) AND vendor AND • keywords AND variations AND • which controlled vocabularies were used AND • list of all limiters WITH rationale, etc.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant

2 = Somewhat Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

Not sure if this tool captures the collaboration of a reviewer and librarian but this is an important element for this section to be thorough. Key words are often not comprehensive it would be important to link this to the variables/theoretical framework

In general, I think it's highly relevant to provide an overview of the data base search process, but it would be unlikely to have adequate space in an article to include everything referenced in "yes". What is meant by vendor? In general I think it would be better (more realistic) to have a "yes", "no" and "can't tell" as options. Authors can always indicate that more detailed information is available from them or could include in supplemental material when allowed by journal

Content Validity

REVISIONS & COMMENTS

Regarding the expert comment about librarian collaboration: although librarian collaboration does appear in one of the methodological sources, that source was somewhat weaker in comparison to more seminal methodological sources. Because this was only raised by one expert, the addition of an item addressing librarian collaboration is not warranted for inclusion. Because the item on use of a theoretical/conceptual framework did not receive enough significant ratings to warrant inclusion in the tool, the comment on connecting keywords to the theoretical framework is disregarded. Since it is possible to include all the elements indicating a Yes response in 1-2 sentences, the full spectrum of requirements is retained. Additional detail on what is meant by “vendor” for the database has been added to the item description.

REVISED ITEM: Were detailed descriptions of the database search processes provided?

DESCRIPTION: A detailed description of a database search will include (at a minimum): database name(s), keywords, whether subject headings were used, any applied limits (year, language), etc. A more comprehensive description will include (at a minimum): database name AND vendor (i.e. OVID MEDLINE; EBSCOhost MEDLINE with Full Text, etc.), keywords AND variations, which controlled vocabularies were used, list of all limiters WITH rationale, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no description of the database search process. There may be a statement such as "Two databases were searched…", but the database names, keywords, etc. are not provided.	There is a minimal description of the database search process provided, which may include: • database name(s) OR • keywords OR • whether or not subject headings were used OR • any applied limits, etc.	The following are explicitly stated: • database name(s) AND • keywords AND • whether or not subject headings were used AND • any applied limits, etc.	The following are explicitly stated: • database name(s) AND vendor AND • keywords AND variations AND • which controlled vocabularies were used AND • list of all limiters WITH rationale, etc.

RETAINED WITH MODIFICATIONS: Was a reproducible line-by-line search strategy (or a sequence of terms for simpler interfaces) provided for at least one database?

DESCRIPTION: A reproducible line-by-line search strategy includes sufficient detail that a reader may recreate the search process and results in a manner that the search results are replicated.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no search strategy provided. There may be a	A reproducible search strategy is not explicitly	All reproducible search strategies are explicitly described;	All reproducible search strategies are explicitly described,

statement such as “A comprehensive search was conducted …”, but keywords or terms, etc. are not provided sufficiently that a search could be replicated.	described; however, there may be enough details provided that the search could possibly be replicated. Or the reproducible search strategy is provided only via online supplement not accessible to reader.	however, none are provided in the manuscript text, but reproducible line-by-line search strategies may be referenced and/or supplied only as online supplements.	and at least one is provided in the manuscript text (not as an online supplement). Additional reproducible line-by-line search strategies may be referenced and/or supplied as online supplements.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant	3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

This is something that may be available on request or in an appendix. not usually seen in a published manuscript.

same comment as above - in general a description of the search strategy needs to be provided, but the level of detail needed to get "full credit" for a "yes" may be unrealistic in the context of space constraints of a manuscript.

Content Validity

REVISIONS & COMMENTS

REVISED ITEM: Are reproducible line-by-line search strategies (or a sequence of terms for simpler interfaces) provided?

DESCRIPTION: A reproducible line-by-line search strategy includes sufficient detail that a reader may recreate the search process and results in a manner that the search results are replicated. The line-by-line search strategies may be supplied as online-only supplements.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no search strategy provided. There may be a statement such as "A comprehensive search was conducted …", but keywords or terms, etc. are not provided sufficiently that a search could be replicated.	A reproducible search strategy is not explicitly described; however, there may be enough details provided that the search could possibly be replicated.	Reproducible line-by-line search strategies are provided for **ONE** database with sufficient detail that a reader may recreate the search process.	Reproducible line-by-line search strategies are provided for **ALL** databases with sufficient detail that a reader may recreate the search processes.

RETAINED WITH MODIFICATIONS: Were other recommended approaches to searching the literature used?

DESCRIPTION: Non-database search approaches may include searching grey literature, ancestry/descendancy (forward/backward) searching, hand searching, networking, or searching research registries.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
No non-database search approaches are described.	Non-database approaches may be inferred but are unclear or not described with sufficient detail.	Only one non-database search approach is used and/or minimal or insufficient detail is provided.	One or more non-database search approaches are used and described with sufficient detail for replication.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

it looks like the word gray or grey are used in this tool...I would suggest one spelling only to be used

Do not agree about the grey literature because of the difficulties with QA

Same concern as mentioned above. I think it's relevant to indicate what non database sources were used, but the four options are cumbersome. Throughout difference between "quite" and "highly" relevant is unclear.

What about searching for non-academic literature?

Content Validity

REVISIONS & COMMENTS

The spelling of gray has been standardized throughout all items, descriptions, and rating scales. The rating scale for this item has been modified to provide greater distinction between Yes and Partial Yes. Non-Academic literature has been added as an example in the description. In response to the expert comment that gray literature presents issues with quality assessment, it is noted that the inclusion of gray literature is one of many examples provided; the item is assessing whether other recommended approaches to searching the literature were used and is stated as such.

REVISED ITEM: Were other recommended approaches to searching the literature used?

DESCRIPTION: Non-database search approaches may include searching gray or non-academic literature, ancestry/descendancy (forward/backward) searching, hand searching, networking, or searching research registries.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
No non-database search approaches are described.	Non-database approaches may be inferred but are unclear and not described with sufficient detail.	~~Only one~~ A non-database search approach is explicitly stated but minimal or insufficient detail is provided so that the	~~One or more~~ Non-database search approaches are explicitly stated and described with sufficient detail for replication.

		process is not replicable.	

RETAINED WITH MODIFICATIONS: If applicable to the purpose and type of literature included, was publication bias addressed?

DESCRIPTION: The use of gray literature as a non-database search approach may be one way of addressing publication bias. An explicit statement in the review acknowledging how publication bias was (or was not) addressed may strengthen the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Publication bias is applicable to the purpose and type of literature included but is not addressed by the review.	Publication bias is not applicable to the purpose and type of literature included, which is explicitly acknowledged by the review.	There is a not an explicit statement of how publication bias was acknowledged by the review, but there is evidence that gray literature or sources outside of traditional publication avenues were retrieved.	There is an explicit statement of how publication bias was acknowledged by the review.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant 4 = Highly Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant
0.80	**0.80**

EXPERT FEEDBACK

Face Validity

Sounds like this is an optional criterion based on the purpose of the review, though I am not sure how one would determine if publication bias was an issue if the non-published literature wasn't addressed. Authors typically address this in the "limitations" section stating it may have been an issue. I think this is sufficient.

Delete [this item]

Content Validity

This seems most relevant to a meta analysis, and not so much an integrative review

REVISIONS & COMMENTS

In response to the expert comment that this is criteria is optional, the ability to rate the item as “Not Applicable” has been added to the rating scale. The description has been revised based on expert feedback to clarify that the item is meant to verify whether the *potential* for publication bias was *acknowledged* and that this statement is often found in the limitations section. While the issue of publication bias is significant to a meta-analysis as one expert noted, the concept of publication bias is also relevant to an integrative review, as the inclusion of a diverse sample of retrieved sources will strengthen the review.

REVISED ITEM: If applicable to the purpose and type of literature included, was the potential for publication bias acknowledged?

DESCRIPTION: An explicit statement acknowledging publication bias may strengthen the review. The use of gray literature as a non-database search approach may be one way of addressing publication bias. This may include a statement that publication bias was (or was not) addressed by the review and is often found in the limitations section.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES	N/A
Publication bias is applicable to the purpose and type of literature included but is not addressed by the review.	Publication bias is not applicable to the purpose and type of literature included, which is explicitly acknowledged by the review.	There is a not an explicit statement of how publication bias was acknowledged by the review, but there is evidence that gray literature or sources outside of traditional publication avenues were retrieved.	There is an explicit statement of how publication bias was acknowledged by the review.	Addressing publication bias is not applicable to the review purpose and/or type of literature included.

DIFFERENTIATION THEME – 4 items (Previously 5 items)

RETAINED AS IS: Were the inclusion/exclusion criteria for the eligible sources clearly stated?

DESCRIPTION: Specific eligibility criteria for inclusion and exclusion are clearly stated. Inclusion and exclusion criteria should be aligned with the purpose and aim of the review. Ideally, rationale will be provided for all limits, particularly those such as excluding non-English sources or limiting to a recent span of years.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Inclusion/exclusion criteria for the eligible sources are not clearly stated.	Inclusion/exclusion criteria are provided but may not be clearly aligned with the review purpose. Or, rationale for eligibility criteria may not be clearly stated.	Inclusion/exclusion criteria which are clearly aligned with the review purpose are explicitly stated. Rationale for eligibility criteria is not clearly stated.	Inclusion/exclusion criteria which are clearly aligned with the review purpose are explicitly stated. Rationale for eligibility criteria is stated.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Description of criterion is clear, but continue to find it confusing to be rating both the criterion and the rating scale. They are two different things.

Content Validity

REVISIONS & COMMENTS

Sole comment appears to be referring to the design of the expert rating survey, not the item under question for the appraisal tool. Item retained as is with no modifications.

DELETED: Was purposive sampling of the search results used?

DESCRIPTION: Purposive sampling is the deliberate selection of a sample of the retrieved sources, as opposed to evaluating the entirety of retrieved sources. Purposive sampling differs from the use of a sampling frame for the entire review, but purposive sampling is also strengthened by explicitly stating rationale.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Purposive sampling was not used. All retrieved sources were included.	There is no explicit statement that purposive sampling was used, but the number of results and sources cannot be reconciled.	Purposive sampling was used but no rationale was provided.	Purposive sampling was used, and an explicit statement of rationale was provided.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
1 = Not Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 3 = Quite Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 2 = Somewhat Relevant
0.40	**0.60**

EXPERT FEEDBACK

Face Validity

If a comprehensive search was conducted, all papers meeting inclusion criteria should be included in the review

Therapeutically relevant, but unrealistic given editors' limitations on page counts.

Omit

It's confusing as this is written...would a no be a lower score...but may not be relevant to have purposive sampling.

Content Validity

Not sure what is meant by purposive sampling. There are multiple approaches to purposive sampling.

Not always relevant in integrative reviews. may be an exhaustive search.

REVISIONS & COMMENTS

Item is deleted.

RETAINED AS IS: Were the search results screened for relevance using a pre-specified set of eligibility criteria?

DESCRIPTION: The pre-specified eligibility criteria for inclusion/exclusion are used to screen retrieved sources. A strong review will include a statement of both title/abstract and full text relevance screening.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Retrieved sources were not screened for inclusion/exclusion using eligibility requirements.	Insufficient detail is provided to determine whether retrieved sources are screened. Or, there is a statement of screening but no statement of eligibility criteria.	Retrieved sources are screened for inclusion/exclusion, but there is not an explicit statement that screening occurred in two phases (title/abstract and full text). Or, the set of eligibility requirements was not pre-specified.	Retrieved sources are screened for inclusion/exclusion in two phases – title/abstract and full text – using a pre-specified set of eligibility requirements.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Description is excellent.

Content Validity

REVISIONS & COMMENTS

Item retained as is with no modifications.

RETAINED WITH MODIFICATIONS: Is a flowchart of search results provided?

DESCRIPTION: A flowchart of search results presents the number of retrieved results, deduplicated results, results remaining after title/abstract and full text screening, and the final number of eligible sources. The most comprehensive flowcharts include detailed breakdowns of retrieved results by database/source and number of results excluded with reasons.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no flowchart of search results provided.	Flowchart of search results is not readily available or is provided only via online supplement not accessible to reader.	A flowchart with the following elements is provided: • number of retrieved results • deduplicated results • number of results remaining after title/abstract screening • number of results remaining after full text screening • final number of eligible sources	A flowchart with the following elements is provided: • number of retrieved results by database/source • deduplicated results • number of results remaining after title/abstract screening • number of results remaining after full text screening by reason for exclusion • final number of eligible sources

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Difference between partial yes and yes is very minor. should make it easier for reviewer to figure out what the difference is.

Content Validity

REVISIONS & COMMENTS

Based on expert comments, the matrix descriptions of Yes and Partial Yes have been modified to be more distinctive. Item is retained with modifications to description.

REVISED ITEM: Is a flowchart of search results provided?

DESCRIPTION: A flowchart of search results presents the number of retrieved results, deduplicated results, results remaining after title/abstract and full text screening, and the final number of eligible sources. The most comprehensive flowcharts include detailed breakdowns of retrieved results by database/source and number of results excluded with reasons.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no flowchart of search results provided.	Flowchart of search results is not readily available or is provided only via online supplement not accessible to reader.	A flowchart is provided but is not comprehensive and only contains minimal elements, such as: • initial number of results • final number of eligible sources, etc.	A flowchart is provided with comprehensive elements, including: • number of retrieved results by database/source • deduplicated results • number of results remaining after title/abstract screening • number of results remaining after full text screening by reason for exclusion • final number of eligible sources

RETAINED WITH MODIFICATIONS: Were the retrieved sources divided into subgroups?

DESCRIPTION: Examples of subgroups may include (but are not limited to): types of evidence, chronology, setting, sample characteristics, etc. The use of subgroups is strengthened by explicitly stating rationale for which group type(s) were chosen.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Retrieved sources were not divided into subgroups.	Retrieved sources were not divided into subgroups, where subgroups would have been appropriate. Or, subgroups were not appropriate for the types of sources retrieved.	Retrieved sources were divided into subgroups but rationale for the group type(s) is not explicitly stated.	Retrieved sources were divided into subgroups, and the rationale for the group type(s) is explicitly stated.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant 4 = Highly Relevant 3 = Quite Relevant 1 = Not Relevant 4 = Highly Relevant	3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 4 = Highly Relevant
0.80	**0.80**

EXPERT FEEDBACK

Face Validity

This criteria addresses aspects of the analysis. I would expect authors to present their analytic approach, but don't think criteria should mandate.

Delete

I do think this is important as it sets integrative review as different from some other review methods.

Content Validity

Identification of subgroups could be an outcome of the analysis

REVISIONS & COMMENTS

Item is modified to acknowledge expert feedback that use of subgroups as part of analysis may not be applicable for all reviews. Inclusion of statement of applicability and option for a rating of N/A is added.

REVISED ITEM: If applicable to the purpose and/or type of literature included, were the retrieved sources divided into subgroups?

DESCRIPTION: Examples of subgroups may include (but are not limited to): types of evidence, chronology, setting, sample characteristics, etc. The use of subgroups is strengthened by explicitly stating rationale for which group type(s) were chosen. If not applicable to the purpose of the review and/or type of literature included, select N/A.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES	N/A
Retrieved sources were not divided into subgroups.	Retrieved sources were not divided into subgroups, where subgroups would have been appropriate. Or, subgroups were not appropriate for the types of sources retrieved.	Retrieved sources were divided into subgroups but rationale for the group type(s) is not explicitly stated.	Retrieved sources were divided into subgroups, and the rationale for the group type(s) is explicitly stated.	Dividing resources into subgroups is not applicable to the review purpose and/or type of literature included.

EXTRACTING THEME – 4 items (Previously 4 items)

RETAINED AS IS: Were the relevant data extracted from all sources using a pre-determined extraction instrument (for all subgroups, if applicable)?

DESCRIPTION: Relevant data should be defined by the authors in relation to the problem, purpose, and research questions guiding the review. An explicit statement of designing the extraction instrument before extraction begins will strengthen the review. If applicable, a mention of data extracted for all subgroups should be present.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
The relevant data were not extracted from all sources and/or were not extracted for all subgroups (if applicable).	It is unclear if all relevant data were extracted from all sources for all subgroups (if applicable), and there may also be no evidence that a pre-determined extraction instrument was used.	The relevant data were extracted from all sources for all subgroups (if applicable), but there is no evidence that a pre-determined extraction instrument was used.	The relevant data were extracted from all sources for all subgroups (if applicable), and there is evidence that a pre-determined extraction instrument was used.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant	3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant 4 = Highly Relevant
1.00	**0.80**

EXPERT FEEDBACK

Face Validity

I think a standardized extraction template is relevant, but also think there should be an option for revising based on the incoming data. I think the subgroup aspects is irrelevant.

Content Validity

Template may need to be modified based on the ongoing analysis

REVISIONS & COMMENTS

Regarding the comment that there should be an “option for revising based on the incoming data,” this would be best be addressed as part of a pilot test of data extraction and is likely not applicable to the point of the review process that this item references. Regarding the inclusion of a reference to subgroups, this was only noted by one expert and is clearly indicated as optional (“if applicable”) by the item as it is written. Item is retained with no modifications.

RETAINED WITH MODIFICATIONS: Is the data extraction process explicit, unbiased, and reproducible?

DESCRIPTION: Like the information retrieval/search process, explicit detail should be presented about the data extraction process so that it could be reproduced, with an emphasis on minimizing transcription error. An unbiased data extraction process would involve more than one reviewer extracting and/or verifying data.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
The data extraction process is not described explicitly so that it is reproducible.	The data extraction process was completed by more than one person but is not sufficiently described so that it is reproducible.	The data extraction process was completed by only one person and is described explicitly so that it is reproducible.	The data extraction process was completed by more than one person and is described explicitly so that it is reproducible.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
2 = Somewhat Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
0.80	**1.00**

EXPERT FEEDBACK

Face Validity

I would suggest adding under "yes" that authors addressed what the process Is if disagreement occurs among reviewers during this phase.

Think this is relevant, but don't know what is meant by transcription error. What transcription?

Content Validity

REVISIONS & COMMENTS

Revised to incorporate expert comments about process for resolving disagreements and removal of reference to “transcription”.

REVISED ITEM: Is the data extraction process explicit, unbiased, and reproducible?

DESCRIPTION: Like the information retrieval/search process, explicit detail should be presented about the data extraction process so that it could be reproduced, with an emphasis on minimizing ~~transcription~~ error. An unbiased data extraction process would involve more than one reviewer extracting and/or verifying data and a process for resolving disagreements.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
The data extraction process is not described explicitly so that it is reproducible.	The data extraction process was completed by more than one person but is not sufficiently described so that it is reproducible.	The data extraction process was completed by only one person and is described explicitly so that it is reproducible.	The data extraction process was completed by more than one person and is described explicitly so that it is reproducible. The process for resolution of disagreements is described.

RETAINED AS IS: Were data analyzed using a review matrix?

DESCRIPTION: A review matrix arranges the extracted data from all sources into columns and rows for a visual representation of data and as a means of arranging data for further analysis. If using subgroups, several matrices may be presented as opposed to one single matrix.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Data were not analyzed using a review matrix.	Data matrix is not readily available or is provided only via online supplement not accessible to reader.	Data were analyzed using a review matrix, but not all extracted data (or subgroups, if applicable) are present.	Data were analyzed using a review matrix, and all extracted data (including subgroups, if applicable) are present.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY

4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Agree. It's important that authors state they developed a template for across-study review - likely one sentence in the manuscript

May require more than one matrix...but I don't think that will change how you phrase this criteria.

Content Validity

REVISIONS & COMMENTS

Item retained with no modifications.

RETAINED AS IS: Is a data display assembling the data from retrieved sources provided?

DESCRIPTION: Examples of data displays may include (but are not limited to): matrices, graphs, charts, networks, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no visualization of data extracted from retrieved sources.	Data display is not readily available or is provided only via online supplement not accessible to reader.	Only partial data extracted from retrieved sources are displayed visually.	All extracted data from all subgroups (if applicable) are provided via visual data display.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 4 = Highly Relevant	3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 4 = Highly Relevant

0.80	**1.00**

EXPERT FEEDBACK

Face Validity

Authors may have done this, but it wouldn't have to be discussed in detail in the manuscript. Typically what is presented in the results section is a table or summary display of the results of the analysis.

This may require supplemental tables.

Content Validity

REVISIONS & COMMENTS

Item retained with no modifications.

VERIFYING THEME – 5 items (Previously 6 items)

RETAINED AS IS: Was a systematic analytic method explicitly identified?

DESCRIPTION: There is an explicit statement of the systematic analytic method (such as constant comparison, content analysis, or thematic analysis, etc.) used for data analysis.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
A systematic analytic method is not named and is not vaguely described.	Analysis is only vaguely described or may be inferred through context but is not explicitly identified by name.	A systematic analytic method process is thoroughly described but is not explicitly named.	A systematic analytic method is explicitly identified by name (constant comparison, content analysis, thematic analysis, etc.), and the process is thoroughly described.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	3 = Quite Relevant

3 = Quite Relevant	2 = Somewhat Relevant
1.00	**0.80**

EXPERT FEEDBACK

Face Validity

While I think this is very relevant...I'm not sure that it requires a specific method of constant comparison or content analysis...that works only with narrative or qualitative analysis...sometimes with integrative reviews you will be creating matrices to compare data from one study to another...why conflicting findings....but you are using the table for pattern recognition (eg. all the significant studies had samples of certain age, or using certain instrument to measure outcome variable). so I think you need a systematic analytic method...but could also include use of tables and matrices.

Content Validity

stated this previously [?]

REVISIONS & COMMENTS

Item retained with no modifications.

RETAINED AS IS: Were commonalities and differences identified?

DESCRIPTION: Identification of commonalties and differences most commonly refers to an examination and comparison between retrieved sources but may also refer to identifying commonalties and differences between the retrieved sources and other extant materials, such as background literature or statistics.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Commonalities and differences are not identified and cannot be inferred from the text.	Commonalities and differences are identified may be inferred from the text but are not explicitly described.	Commonalities and differences are identified and present in the text but are not explicitly described.	Commonalities and differences are identified and explicitly described.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 2 = Somewhat Relevant
0.80	**0.80**

EXPERT FEEDBACK

Face Validity

I'm confused by this. Where would you expect to see commonalities and differences discussed…in the results or the discussion? Seems to me the focus should be in the results section.

Content Validity

also stated this previously [?]

REVISIONS & COMMENTS

Regarding the comment that commonalities and differences should be discussed in a particular section: unlike reporting guidelines, this appraisal tool is not designed to specify that methodological elements appear in certain sections of the manuscript, only that the specified elements are present in the manuscript so that they are able to be appraised. Item retained with no modifications.

RETAINED AS IS: Was conflicting evidence addressed?

DESCRIPTION: Addressing conflicting evidence differs from looking at identified differences in that an examination of conflicts seeks to determine a reason or explanation for the differences.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Conflicting evidence is clearly present but is not addressed.	It cannot be determined if conflicting evidence is present. Or, conflicting evidence	Conflicting evidence is examined with possible explanations for differences able to	Conflicting evidence is examined with possible explanations for

	is present and discussed but not adequately.	be inferred but not explicitly stated.	differences explicitly stated.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

This is a suggestion for the tool to use the IMRAD format to guide the questions for users. I now realize you are addressing elements of a discussion section of a review.

How is this different from what is stated in previous item?

Content Validity

REVISIONS & COMMENTS

Regarding the suggestion to use IMRAD format, IMRAD is applicable to reporting guidelines; this is an appraisal tool. The description of the item sufficiently addresses how addressing conflicting evidence differs from identifying commonalities and differences. Item retained with no modifications.

RETAINED WITH MODIFICATIONS: Was the quality of retrieved sources addressed in a meaningful way?

DESCRIPTION: Quality of retrieved sources must be addressed in a meaningful way, with attention paid to aligning assessment of quality with the methodology or research design of identified sources. Ideally, the methods and tools used to assess quality are explicitly stated. Thorough assessment of quality will also include an examination of risk of bias and study reliability, if applicable.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Quality of retrieved sources was not addressed.	Quality of retrieved sources was addressed but not in a meaningful way. Alignment of tools with research designs of sources may be inadequate or may not be addressed.	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to assess quality. Risk of bias and study reliability were not assessed.	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to assess quality. Risk of bias and study reliability are also assessed.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Yes -- relevant. Authors need to report their approach, indicate if reports were excluded based on quality, and briefly describe how quality was addressed in the analysis.

This is always a challenge in integrative reviews...multiple methods. some integrative reviews can also include theory papers. how to evaluate quality in these. I do think for descriptive purposes, it is helpful to know for each section or type of study, what was overall quality.

Content Validity

REVISIONS & COMMENTS

The question of whether identified sources were excluded based on quality was addressed by a separate question that was deleted due to insufficient content validity. The description of this item is retained with modifications to address the deleted question.

REVISED ITEM: Was the quality of retrieved sources addressed in a meaningful way?

DESCRIPTION: Quality of retrieved sources must be addressed in a meaningful way, with attention paid to aligning assessment of quality with the methodology or research design of identified sources. Ideally, there is an explicit statement of the methods and tools used to assess quality and how quality assessments were addressed (i.e. were any sources removed or interpreted differently due to quality, etc.). Thorough assessment of quality will also include an examination of risk of bias and study reliability, if applicable.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Quality of retrieved sources was not addressed.	Quality of retrieved sources was addressed but not in a meaningful way. Alignment of tools with research designs of sources may be inadequate or may not be addressed. There may not be a statement of how sources were handled based on quality appraisal.	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to assess quality. Risk of bias and study reliability were not assessed. There may not be a statement of how sources were handled based on quality appraisal.	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to assess quality. Risk of bias and study reliability are also assessed. How sources were handled based on quality appraisal is stated.

RETAINED AS IS: Was quality verified by two independent reviewers?

DESCRIPTION: The use of two independent raters is preferable as this ensures accuracy and reduces bias.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no evidence that quality was verified as part of the review.	It cannot be determined if quality was verified. Or, there is evidence that quality was	There is evidence that quality was verified by one reviewer.	There is evidence that quality was verified by two reviewers. Inter-rater agreement

	verified but the number of reviewers is not stated.		calculations may have also been provided.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant	3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Not always necessary but should be stated.

Sure -- but all authors have to say is that they did it.

Content Validity

REVISIONS & COMMENTS

Item is retained as is with no modifications.

DELETED: Were any articles excluded based on quality appraisal?

DESCRIPTION: Based on the outcome of quality appraisal, low quality articles may be excluded. Rationale for inclusion or exclusion of low quality should be explicitly stated. If low quality articles are excluded, it is ideal that a detailed log of excluded sources is provided with rationale explicitly stated.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Quality appraisal was not completed.	Quality appraisal was completed AND there were articles of low quality, but no articles were excluded based on quality appraisal nor was rationale for inclusion of low-	Quality appraisal was completed AND there were articles of low quality which were excluded, but there is no detailed log of excluded articles nor rationale provided.	Quality appraisal was completed AND there were articles of low quality which were excluded, and there is a detailed log of excluded articles or rationale provided.

	quality articles provided.		

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant	3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 1 = Not Relevant 2 = Somewhat Relevant
1.00	**0.60**

EXPERT FEEDBACK

Face Validity

All authors have to say is how they reviewed quality and and how quality assessments were addressed in the results.

As mentioned, this is always challenging when you have many types of designs. I don't necessarily think that you need to exclude studies of low quality...rather author needs to point out which results were based on studies of low quality, so may be more tentative/cautious in interpretation.

Content Validity

Authors need to provide a rationale for excluding or not excluding based on relevance.

stated previously [?]

REVISIONS & COMMENTS

Item is deleted.

ENDING THEME – 3 items (Previously 5 items)

DELETED: Were patterns, themes, relationships, or conclusions verified with sources?

DESCRIPTION: Patterns, themes, relationships, or conclusions are clearly stated. Explicit statements of verification with retrieved sources will strengthen the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Patterns, themes, relationships, or conclusions are not identified.	Patterns, themes, relationships, or conclusions may be identified, but it is unclear whether conclusions are verified with sources.	Patterns, themes, relationships, or conclusions are identified, but verification with sources is not explicit but may be easily inferred.	Patterns, themes, relationships, or conclusions are identified and are explicitly verified with sources.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
3 = Quite Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant 2 = Somewhat Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant
0.60	**1.00**

EXPERT FEEDBACK

Face Validity

Not sure what this means - what is verification from retrieved resources. Isn't this part of the extraction process?

While this is important, I don't often see this in published manuscripts. This is more a criteria I associate with qualitative research.

Delete

Content Validity

REVISIONS & COMMENTS

Item is deleted based on insufficient face validity and expert recommendations for deletion.

RETAINED AS IS: Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?

DESCRIPTION: The review is concluded by explicitly synthesizing elements or conclusions into a summation of the topic or phenomenon. Merely listing identified elements is

not an explicit means of synthesis. If applicable, synthesized elements or conclusions are present for all subgroups.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Important elements or conclusions are not synthesized into a summation of the topic or phenomenon.	Important elements or conclusions are explicitly listed or discussed separately but are not clearly synthesized into a summation of the topic or phenomenon.	Important elements or conclusions are presented in a manner that is more consistent with listing than synthesis or summation of the topic or phenomenon. Or, if applicable, synthesized summations are not present for all subgroups.	Important elements or conclusions are explicitly synthesized into a summation of the topic or phenomenon. If applicable, synthesized summations are present for all subgroups.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Content Validity

REVISIONS & COMMENTS

Item is retained with no modifications.

DELETED: Were the conclusions reported in table or diagrammatic form?

DESCRIPTION: Tables or diagrammatic forms of conclusions differ from visualizations of extracted data in that they represent the synthesis or outcomes of the review. Examples of visualization of conclusions may include (but are not limited to): matrices, graphs, charts, diagrams, networks, joint displays, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Conclusions are not reported in table or diagrammatic form.	Table(s) or diagrammatic form(s) of conclusions are not readily available or are provided only via online supplement not accessible to reader.	Conclusions are reported in table or diagrammatic form but not explicitly in a meaningful way that contributes to new knowledge or understanding of the topic or phenomenon.	Conclusions are reported in table or diagrammatic form in a meaningful way that contributes to new knowledge or understanding of the topic or phenomenon.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant 1 = Not Relevant 4 = Highly Relevant	2 = Somewhat Relevant 4 = Highly Relevant 3 = Quite Relevant 1 = Not Relevant 4 = Highly Relevant
0.80	**0.60**

EXPERT FEEDBACK

Face Validity

I think the conclusions would be in narrative form with the summary results in tables

Content Validity

Conclusions described in narrative form are appropriate as well.

Don't think the criteria should mandate the presentation format

REVISIONS & COMMENTS

Item is deleted based on insufficient content validity and expert comments.

RETAINED AS IS: Were implications discussed for: research, practice, education, or policy?

DESCRIPTION: Implications for research, practice, education, or policy are explicitly stated. Reviews may be strengthened by directly aligning implications with both the purpose and synthesized findings of the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Implications for research, practice, education, or policy are not explicitly stated nor may they be inferred from the text.	Implications for research, practice, education, or policy may be inferred but do not clearly align to the purpose of the review or its findings.	Implications for research, practice, education, or policy and alignment with review purpose may be inferred but are not explicitly stated. Or, implications for research, practice, education, or policy are stated but do not clearly align to the purpose of the review or its findings.	Implications for research, practice, education, or policy are explicitly stated and discussed, with clear alignment with purpose and findings of the review.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant 2 = Somewhat Relevant 4 = Highly Relevant 4 = Highly Relevant 3 = Quite Relevant	4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 4 = Highly Relevant 2 = Somewhat Relevant
0.80	**0.80**

EXPERT FEEDBACK

Face Validity

I would add somewhere is these last several questions related to the discussion section is: have the reviewers linked their discussion back to the guiding theoretical framework and background information found in the introduction sections...I do think the the IMRAD elements would help organize this tool better.

I'm feeling mixed about this...is this about the quality of the method...which is what I think you are looking for...or a reporting guideline for integrative reviews? this is aligned more with the latter.

Content Validity

REVISIONS & COMMENTS

Item is retained as is.

RETAINED AS IS: Were methodological or other limitations of the review explicitly stated?

DESCRIPTION: Methodological limitations of the review may include (but are not limited to): only one individual for screening, quality appraisal, thematic analysis, etc.; limiting to English-language or recent years with no rationale; limited access to databases or full text resources for retrieval of sources, etc. Methodological limitations of the review do not include limitations of the retrieved sources, which is more appropriately discussed during quality appraisal.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There are no methodological limitations listed.	Limitations are only mentioned regarding retrieved sources, not methodological limitations of the review itself.	Methodological limitations of the review may be inferred but are not explicitly stated. Or, methodological limitations of the review are present but are not discussed.	All methodological or other limitations of the review are explicitly stated.

EXPERT RATINGS

FACE VALIDITY	CONTENT VALIDITY
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
4 = Highly Relevant	4 = Highly Relevant
1.00	**1.00**

EXPERT FEEDBACK

Face Validity

Content Validity

REVISIONS & COMMENTS

Item is retained as is.

Appendix H: *IRB Exemption for Inter-Rater Reliability and Construct Validity Aims*

University of CINCINNATI
Institutional Review Board
FWA #: 000003152

EXEMPT DETERMINATION

September 20, 2022

Emily Kean
CON Adjunct Faculty

Dear Emily Kean,

Type of Submission:	Initial Study
Title:	Testing of Inter-Rater Reliability and Construct Validity of the Integrative Review Methodology Appraisal Tool (IRMAT)
Investigator:	Emily Kean
IRB ID:	2022-0780
Funding:	Name: Investigator Initiated
Documents Reviewed:	• 2022-09-10 HRP-503 Protocol Submission V2.pdf
Type of Review:	Exempt
Review Category:	• (2)(ii) Tests, surveys, interviews, or observation (low risk)

On **9/20/2022**, the IRB reviewed the above submission and determined that this protocol meets the criteria for exemption from IRB review in accordance with 45 CFR 46.104.

The IRB has determined the following consent requirements:

- The IRB has waived the requirement to obtain DOCUMENTATION of informed consent for all adult participants.

PI Notification
Ongoing IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities impact the exempt determination, please submit a new request to the IRB for a determination.

Note: The exemptions at 45 CFR 46.101(b) do not apply to research involving prisoners, fetuses, pregnant women, or human in vitro fertilization, Subparts B and C. The exemption at 45 CFR 46.101(b)(2), for research involving survey or interview procedures or observation of public behavior, does not apply to research with children, Subpart D, except for research involving observations of public behavior when the investigator(s) do not participate in the activities being observed.

Thank you for your cooperation during the review process.

v 2.12.20

Appendix I: *IRMAT User Guide (V 1.3)*

Integrative Review Methodology Appraisal Tool (IRMAT)

User Guide

What is the IRMAT?

The Integrative Review Methodology Appraisal Tool, or IRMAT ("ear" "mat") is a tool designed to appraise published integrative reviews in the nursing literature.

How was the IRMAT developed?

The IRMAT was developed by generating items from an extensive scoping review of integrative review methodologies in the nursing literature. These items were appraised by a team of expert reviewers for face and content validity and were subsequently psychometrically tested and deemed to demonstrate adequate construct validity and slight to moderate inter-rater reliability.

What can the IRMAT be used for?

The IRMAT has been designed and tested to be used to evaluate integrative reviews published in the nursing discipline. It is not recommended to use the IRMAT with non-nursing disciplines without first testing for validity and reliability with published integrative reviews from those non-nursing disciplines.

What are the requirements?

If the IRMAT is being used to appraise integrative reviews for potential inclusion in evidence synthesis projects or publications, it is advisable to have two raters independently apply the IRMAT to the integrative review article(s) being appraised. Like appraising any research output, it is ideal if raters have knowledge or receive training on the general format of integrative reviews and/or the content areas discussed in the articles being appraised.

How to use the IRMAT

Each item of the IRMAT is presented alongside a description of the item and a rating scale including examples with indicators for each possible rating. All items should be rated on the following scale:

NO | PARTIAL NO (or UNKNOWN) | PARTIAL YES | YES

1. Read the article to be appraised in its entirety.
2. Score each item using the ratings:
   - NO
   - PARTIAL NO (or UNKNOWN)
   - PARTIAL YES
   - YES

   Refer to the item description and rating scale examples for guidance.
3. As indicated in the rating scale examples, a score of "PARTIAL NO or UNKNOWN" may be applied in cases where the item is not present or cannot be determined. Refer to each item's rating scale examples for scoring guidance on cases where the item *is* present but is not explicitly or fully represented.
4. Select items have an option of "N/A" or Not Applicable. Raters should refer to the rating scale examples for those items and use best judgement to determine where "N/A" is appropriate.

How to score

The IRMAT has not been designed nor tested to compute a score in the traditional sense; totaling the item ratings to compute an overall score is discouraged. The IRMAT is best used as a holistic tool, with the scored items used as an overall indicator of the adherence to integrative review methodology of the article being appraised.

Were the factors of interest (concepts, variables, etc.) clearly stated?

DESCRIPTION: Factors of interest may be defined as concepts, variables, target population, health care problem, etc. Factors may or may not be defined in terms of PICO(S) elements (population, intervention, comparison, outcome, setting).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no explicit statement of factors of interest, and variables cannot be inferred.	Factors of interest are not explicitly stated but may be inferred.	There is an explicit statement of the factors of interest, but factors are not clearly defined within specific contexts (population, outcomes, etc.).	There is an explicit statement of the factors of interest, and factors are clearly defined within specific contexts (population, outcomes, etc.).

Was the problem stated unambiguously and was it easy to identify?

DESCRIPTION: A problem statement outlines the areas under examination by the review. Where the statement of a preliminary search describes the action which was taken to identify the need for the integrative review, the problem statement is an explicit statement of the problem the review seeks to address. An easily identifiable problem statement may begin with a statement such as “This review addresses the problem of…”, or “Previous literature reviews highlighted problems surrounding…”, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no clearly stated problem.	A problem statement is alluded to but is not clearly stated.	The problem statement is clear but is not easily identifiable.	The problem statement is clear and is easily identifiable.

Were the inclusion/exclusion criteria for the eligible sources clearly stated?

DESCRIPTION: Specific eligibility criteria for inclusion and exclusion are clearly stated. Inclusion and exclusion criteria should be aligned with the purpose and aim of the review. Ideally, rationale will be provided for all limits, particularly those such as excluding non-English sources or limiting to a recent span of years.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Inclusion/exclusion criteria for the eligible sources are not clearly stated.	Inclusion/exclusion criteria are provided but may not be clearly aligned with the review purpose. Or, rationale for eligibility criteria may not be clearly stated.	Inclusion/exclusion criteria which are clearly aligned with the review purpose are explicitly stated. Rationale for eligibility criteria is not clearly stated.	Inclusion/exclusion criteria which are clearly aligned with the review purpose are explicitly stated. Rationale for eligibility criteria is stated.

If applicable to the purpose and/or type of literature included, was the sampling frame clearly stated?

DESCRIPTION: A sampling frame for an integrative review may include (but is not limited to): the type of empirical studies, specific research design(s), inclusion of methodological or theoretical literature/framework, etc. Providing a rationale for choice of sampling frame strengthens the choice of using a sampling frame. Attention should also be paid to the alignment of the choice of sampling frame with the integrative review methodology. For example, the choice to use a sampling frame which limits the retrieved sources to only quantitative results does not align with the choice to use an integrative review methodology, which implies inclusion of varying result types (empirical [quantitative, qualitative, mixed methods], methodological, or theoretical).

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES	N/A
There is no clear statement of sampling frame nor is rationale provided. Or, the sampling frame does not align with the integrative review methodology, i.**e.,** limiting to only randomized controlled trials, where a systematic review would be more appropriate.	There is no clear statement of whether a sampling frame was used, nor is rationale provided; however, a sampling frame may be alluded to in the screening (inclusion/exclusion) or data analysis descriptions.	A sampling frame is clearly stated but no rationale is provided.	A sampling frame with rationale is clearly stated.	The use of a sampling frame is not applicable to the review purpose and/or type of literature included.

Were detailed descriptions of the database search processes provided?

DESCRIPTION: A detailed description of a database search will include (at a minimum): database name(s), keywords, whether subject headings were used, any applied limits (year, language), etc. A more comprehensive description will include (at a minimum): database name AND vendor (i.e. OVID MEDLINE; EBSCOhost MEDLINE with Full Text, etc.), keywords AND variations, which controlled vocabularies were used, list of all limiters WITH rationale, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no description of the database search process. There may be a statement such as "Two databases were searched…", but the database names, keywords, etc. are not provided.	There is a minimal description of the database search process provided, which may include: • database name(s) **OR** • keywords **OR** • whether or not subject headings were used **OR** • any applied limits, etc.	The following are explicitly stated: • database name(s) **AND** • keywords **AND** • whether or not subject headings were used **AND** • any applied limits, etc.	The following are explicitly stated: • database name(s) *AND* vendor **AND** • keywords *AND* variations **AND** • which controlled vocabularies were used **AND** • list of all limiters *WITH* rationale, etc.

Are reproducible line-by-line search strategies (or a sequence of terms for simpler interfaces) provided?

DESCRIPTION: A reproducible line-by-line search strategy includes sufficient detail that a reader may recreate the search process and results in a manner that the search results are replicated. The line-by-line search strategies may be supplied as online-only supplements

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no search strategy provided. There may be a statement such as "A comprehensive search was conducted …", but keywords or terms, etc. are not provided sufficiently that a search could be replicated.	A reproducible search strategy is not explicitly described; however, there may be enough details provided that the search could possibly be replicated.	Reproducible line-by-line search strategies are provided for **ONE** database with sufficient detail that a reader may recreate the search process.	Reproducible line-by-line search strategies are provided for **ALL** databases with sufficient detail that a reader may recreate the search processes.

Were the search results screened for relevance using a pre-specified set of eligibility criteria?

DESCRIPTION: The pre-specified eligibility criteria for inclusion/exclusion are used to screen retrieved sources. A strong review will include a statement of both title/abstract and full text relevance screening.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Retrieved sources were not screened for inclusion/exclusion using eligibility requirements.	Insufficient detail is provided to determine whether retrieved sources are screened. Or, there is a statement of screening but no statement of eligibility criteria.	Retrieved sources are screened for inclusion/exclusion, but there is not an explicit statement that screening occurred in two phases (title/abstract and full text). Or, the set of eligibility requirements was not pre-specified.	Retrieved sources are screened for inclusion/exclusion in two phases – title/abstract and full text – using a pre-specified set of eligibility requirements.

Is the data extraction process explicit, unbiased, and reproducible?

DESCRIPTION: Like the information retrieval/search process, explicit detail should be presented about the data extraction process so that it could be reproduced, with an emphasis on minimizing error. An unbiased data extraction process would involve more than one reviewer extracting and/or verifying data and a process for resolving disagreements.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
The data extraction process is not described explicitly so that it is reproducible.	The data extraction process was completed by more than one person but is not sufficiently described so that it is reproducible.	The data extraction process was completed by only one person and is described explicitly so that it is reproducible.	The data extraction process was completed by more than one person and is described explicitly so that it is reproducible. The process for resolution of disagreements is described.

Were data analyzed using a review matrix?

DESCRIPTION: A review matrix arranges the extracted data from all sources into columns and rows for a visual representation of data and as a means of arranging data for further analysis. If using subgroups, several matrices may be presented as opposed to one single matrix.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Data were not analyzed using a review matrix.	Data matrix is not readily available or is provided only via online supplement not accessible to reader.	Data were analyzed using a review matrix, but not all extracted data (or subgroups, if applicable) are present.	Data were analyzed using a review matrix, and all extracted data (including subgroups, if applicable) are present.

Is a data display assembling the data from retrieved sources provided?

DESCRIPTION: Examples of data displays may include (but are not limited to): matrices, graphs, charts, networks, etc.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no visualization of data extracted from retrieved sources.	Data display is not readily available or is provided only via online supplement not accessible to reader.	Only partial data extracted from retrieved sources are displayed visually.	All extracted data from all subgroups (if applicable) are provided via visual data display.

Was a systematic analytic method explicitly identified?

DESCRIPTION: There is an explicit statement of the systematic analytic method (such as constant comparison, content analysis, or thematic analysis, etc.) used for data analysis.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
A systematic analytic method is not named and is not vaguely described.	Analysis is only vaguely described or may be inferred through context but is not explicitly identified by name.	A systematic analytic method process is thoroughly described but is not explicitly named.	A systematic analytic method is explicitly identified by name (constant comparison, content analysis, thematic analysis, etc.), and the process is thoroughly described.

Were commonalities and differences identified?

DESCRIPTION: Identification of commonalties and differences most commonly refers to an examination and comparison between retrieved sources but may also refer to identifying commonalties and differences between the retrieved sources and other extant materials, such as background literature or statistics.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Commonalities and differences are not identified and cannot be inferred from the text.	Commonalities and differences are identified may be inferred from the text but are not explicitly described.	Commonalities and differences are identified and present in the text but are not explicitly described.	Commonalities and differences are identified and explicitly described.

Was conflicting evidence addressed?

DESCRIPTION: Addressing conflicting evidence differs from looking at identified differences in that an examination of conflicts seeks to determine a reason or explanation for the differences.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Conflicting evidence is clearly present but is not addressed.	It cannot be determined if conflicting evidence is present. Or, conflicting evidence is present and discussed but not adequately.	Conflicting evidence is examined with possible explanations for differences able to be inferred but not explicitly stated.	Conflicting evidence is examined with possible explanations for differences explicitly stated.

Was the quality of retrieved sources addressed in a meaningful way?

DESCRIPTION: Quality of retrieved sources must be addressed in a meaningful way, with attention paid to aligning assessment of quality with the methodology or research design of identified sources. Ideally, there is an explicit statement of the methods and tools used to assess quality and how quality assessments were addressed (i.e. were any sources removed or interpreted differently due to quality, etc.). Thorough assessment of quality will also include an examination of risk of bias and study reliability, if applicable.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Quality of retrieved sources was not addressed.	Quality of retrieved sources was addressed but not in a meaningful way. Alignment of tools with research designs of sources may be inadequate or may not be addressed. There may not be a statement of how sources were handled based on quality appraisal.	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to assess quality. Risk of bias and study reliability were not assessed. There may not be a statement of how sources were handled based on quality appraisal.	Quality of retrieved sources was addressed in a meaningful way. If numerous research designs were included, appropriate tools are used to assess quality. Risk of bias and study reliability are also assessed. How sources were handled based on quality appraisal is stated.

Was quality verified by two independent reviewers?

DESCRIPTION: The use of two independent raters is preferable as this ensures accuracy and reduces bias.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There is no evidence that quality was verified as part of the review.	It cannot be determined if quality was verified. Or, there is evidence that quality was verified but the number of reviewers is not stated.	There is evidence that quality was verified by one reviewer.	There is evidence that quality was verified by two reviewers. Inter-rater agreement calculations may have also been provided.

Were important elements or conclusions (for all subgroups, if applicable) synthesized into a summation of the topic or phenomenon?

DESCRIPTION: The review is concluded by explicitly synthesizing elements or conclusions into a summation of the topic or phenomenon. Merely listing identified elements is not an explicit means of synthesis. If applicable, synthesized elements or conclusions are present for all subgroups.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Important elements or conclusions are not synthesized into a summation of the topic or phenomenon.	Important elements or conclusions are explicitly listed or discussed separately but are not clearly synthesized into a summation of the topic or phenomenon.	Important elements or conclusions are presented in a manner that is more consistent with listing than synthesis or summation of the topic or phenomenon. Or, if applicable, synthesized summations are not present for all subgroups.	Important elements or conclusions are explicitly synthesized into a summation of the topic or phenomenon. If applicable, synthesized summations are present for all subgroups.

Were methodological or other limitations of the review explicitly stated?

DESCRIPTION: Methodological limitations of the review may include (but are not limited to): only one individual for screening, quality appraisal, thematic analysis, etc.; limiting to English-language or recent years with no rationale; limited access to databases or full text resources for retrieval of sources, etc. Methodological limitations of the review do not include limitations of the retrieved sources, which is more appropriately discussed during quality appraisal.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
There are no methodological limitations listed.	Limitations are only mentioned regarding retrieved sources, not methodological limitations of the review itself.	Methodological limitations of the review may be inferred but are not explicitly stated. Or, methodological limitations of the review are present but are not discussed.	All methodological or other limitations of the review are explicitly stated.

Were implications discussed for: research, practice, education, or policy?

DESCRIPTION: Implications for research, practice, education, or policy are explicitly stated. Reviews may be strengthened by directly aligning implications with both the purpose and synthesized findings of the review.

NO	PARTIAL NO or UNKNOWN	PARTIAL YES	YES
Implications for research, practice, education, or policy are not explicitly stated nor may they be inferred from the text.	Implications for research, practice, education, or policy may be inferred but do not clearly align to the purpose of the review or its findings.	Implications and alignment with review purpose may be inferred but are not explicitly stated. Or, implications are stated but do not clearly align to the purpose of the review or its findings.	Implications for research, practice, education, or policy are explicitly stated and discussed, with clear alignment with purpose and findings of the review.