Journal of Clinical Epidemiology 129 (2021) 176−187

Journal of Clinical Epidemiology

REPLICATION RESEARCH SERIES

Replication Research Series-Paper 1 : A concept analysis and meta-narrative review established a comprehensive theoretical definition of replication research to improve its use

Brigitte Vachon[a,*], Janet A. Curran[b], Sathya Karunananthan[c], Jamie Brehaut[c,d], Ian D. Graham[c,d], David Moher[c,d], Anne E. Sales[d], Sharon E. Straus[e,f], Michele Fiander[c], P. Alison Paprica[g], Jeremy M. Grimshaw[c,d]

[a]*Faculty of Medicine, Occupational Therapy Program, School of Rehabilitation, Université de Montréal, 7077 Avenue du Parc, Montreal, Quebec, Canada, H3C 3J7*
[b]*School of Nursing, Faculty of Health, Dalhousie University, 5869 University Avenue, Halifax, Nova Scotia, Canada, B3H 4R2*
[c]*Centre for Practice-Changing Research, Ottawa Hospital Research Institute, 501 Smyth Road, Box 711, Ottawa, Ontario, Canada, K1H 8L6*
[d]*Department of Learning Health Sciences, Clinical Epidemiology Program, University of Ottawa, 600 Peter Morand Crescent, Ottawa, Ontario, Canada, K1G 5Z3*
[e]*Medical School, University of Michigan, 209 Victor Vaughan Building, 1111 E. Catherine St, Ann Arbor, MI 48109-2054*
[f]*Department of Geriatric Medicine, University of Toronto, Medical Sciences Building, Toronto, 1 King's College Circle, Toronto, M5S 1A8, Ontario, Canada*
[g]*Institute of Health Policy, Management and Evaluation University of Toronto Health Sciences Building, 155 College Street, Suite 425, Toronto, Ontario, Canada, M5T 3M6*

Accepted 9 July 2020; Published online 16 July 2020

Abstract

Objectives: The aim of this study is to clarify the concept of replication research to improve its appropriate use by researchers, editors, research funders, and decision makers.

Study Design and Setting: We combined concept analysis and metanarrative review methods to synthetize knowledge on replication research from various scientific fields. We used multiple search strategies to identify the relevant literature published before April 2018. We summarized the data by seeking commonalities and differences in underlying conceptual and theoretical assumptions in the literature.

Results: A total of 153 articles from various disciplines were included. The analysis led to the identification of three major definitions of replication: the repetition of a previous study, the extension of a previous study, and the road-testing of a theory. Attributes, conditions required to conduct replication studies, concerns related to the interpretation of replication studies, and diverse replication research typologies were synthesized, combined, and analyzed. Based on this metanarrative review, a comprehensive theoretical definition of replication research was formulated.

Conclusion: This study can support the adoption of a shared understanding and recognition of the indispensable nature of replication research for the sound development of knowledge in all research fields.

Keywords: Replication; Reproducibility; Validity; Reliability; Concept analysis; Metanarrative review

1. Introduction

Replication of research is essential for the advancement of science to establish credibility of quantitative research findings and understanding their likely generalizability [1,2]. Single studies are rarely a sufficient foundation on which to base practice or policy decisions [3−5]. Often the results of highly cited original clinical research have been refuted over time [6,7]. Yet, researchers continue to conduct investigations, which are not optimally informed by previous research and thus only partially contribute to the development of knowledge [8]. Typically, requests for research funding encourage researchers to evaluate new strategies even though a lack of replication may exist for previous promising strategies in a field [8−10].

Replication is often understood to mean the exact duplication of a quantitative study [8]. However, it can also

Funding: This work was supported by the Canadian Institutes of Health Research, Knowledge Synthesis Grant, Fall 2010 Competition [KST-108047].

Declaration of interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

* Corresponding author. Faculty of Medicine, Université de Montréal C.P, 6128, Succursale Centre-ville, Montréal, Quebec, Canada, H3C 3J7. Tel.: (+1) (514) 343-2094.

E-mail address: brigitte.vachon@umontreal.ca (B. Vachon).

https://doi.org/10.1016/j.jclinepi.2020.07.006

involve complete or partial reproduction of the study design and methods with a new group or in another context to determine whether similar findings will be obtained [11]. Different types of replication have been proposed that vary in their purpose and methods. However, in a recent literature review, Gomez et al. point to the lack of a generally accepted classification or agreement on terms for the different types of replication within and across disciplines [12]. The use of multiple terms and definitions is confusing and may hinder scientific discussion on the value of replication [13]. Advancing health research and improving the applicability of research findings require the development of clear recommendations for replication [14,15], based on a common understanding of the concept and types of replication.

In this study, we contribute to the critical discourse on replication through a metanarrative review and concept analysis of the literature on replication. Our objectives are to define, clarify, and conceptualize the concept of replication in quantitative research through the analysis of the literature from health sciences, social sciences, psychology, education, business, mathematics, and other basic sciences. This report provides a broad perspective of replication, presents definitions and uses of the concept across research fields, and identifies key attributes, conditions, and consequences of replication. Finally, through our review and concept analysis, we arrive at a theoretical definition for replication research which can support adoption of a shared understanding and recognition of the indispensable nature of this kind of activity for the sound development of knowledge in all research fields.

2. Materials and methods

Concept analysis is a way to deconstruct a term to understand it better and to develop a sound definition of a concept by examining its structure and function. We used the iterative eight-step approach proposed by Walker and Avant [16,17] (Table 1) as a guide to examine the literature and organize the information found in a comprehensive way.

For the purpose of enhancing the rigor of the review process, we used the systematic metanarrative review methodology to plan and conduct our search strategy, data extraction, and analysis processes [18,19]. Metanarrative review is a method proposed to synthesis the literature from a complex body of evidence [20]. Because it is used to highlight the contrasting and complementary ways in which researchers have studied or described a given topic, it could easily be combined with the concept analysis approach. It looks historically at how particular research traditions have unfolded over time and shaped the kind of questions being asked and the methods used to answer them. The main objective of a metanarrative review is to shed new light on a complex topic area by drawing from multiple perspectives [18,20]. This process follows the principle of interpretivist analysis by seeking commonalities in conceptual assumptions, highlighting differences and drawing inferences from these commonalities and differences [20]. More details on both approaches can be found in the protocol in Appendix 1.

Table 1. Steps of the concept analysis method developed by Walker and Avant [16,17]

1. Select a concept that lacks clarity and is relevant to practice
2. Determine the aims and purpose of the analysis
3. Identify all uses of the concept in different contexts and disciplines
4. Determine the defining attributes that can differentiate the specific concept from a similar one
5. Identify cases that describe and delimitate clearly what is the concept
6. Identify borderline and contrary cases that describe and delimitate what is not the concept
7. Identify antecedents (events or incidents that must be in place for the concept to occur) and consequences (outcomes that occur as a result of the concept)
8. Identify empirical referents or observable phenomena demonstrating the occurrence of the concept

We conducted a preliminary search to identify five ''model'' articles on replication from across different disciplines [1,2,8,21,22] and used the reference lists of these articles to identify major works on replication (most commonly cited when defining replication) and to identify synonyms, surrogate terms, and bibliographic indexing terms for different bibliographic databases. An information scientist (M.F.) developed the search strategy (see Appendix 2 for the search strategy) for multiple relevant databases (including PsycINFO 1848-(Ovid), CINAHL 1980-(EBSCO), Medline 1948-(Ovid), EMBASE 1980-(Ovid), ERIC, and WorldCat). The relevance of each article was independently assessed by two investigators (J.C., S.K., and B.V.). An article was considered relevant to the analysis if it 1) provided a clear definition of the concept (replication research or a surrogate term) and 2) described different assumptions related to its use and importance in research. We excluded replication studies that did not discuss the concept of replication, as well as viral replication studies, replication of statistical model articles, and articles published in languages other than English and French. While screening these articles, relevant citations were retrieved and screened using the snowball method. Finally, team members were asked about any germinal articles. Search results were discussed at a team consensus meeting. Articles that generated disagreement among the screeners were presented to the group to clarify inclusion and exclusion criteria. In April 2018, to update the results, a forward search of the five model articles and the most cited articles [23,24] was conducted to include relevant articles published between 2014 and May 2018 and make sure we have included new perspectives on replication that may have emerged since our initial search.

Table 2. Mapping results

Field	# Included	Pub dates	Type of articles					
			Concept	Lit. review	Editorial	Methods	Study rep	Opinion
Business	16	1990–1999 (n = 1) 2000–2009 (n = 11) 2013–2018 (n = 4)	10	-	1	1	4	-
Health	43	1960–1969 (n = 1) 1980–1989 (n = 8) 1990–1999 (n = 7) 2000–2009 (n = 24) 2010–2012 (n = 2) 2013–2018 (n = 1)	17	4	8	7	7	-
Psychology	39	1950–1959 (n = 2) 1960–1969 (n = 1) 1970–1979 (n = 3) 1980–1989 (n = 7) 1990–1999 (n = 4) 2000–2009 (n = 5) 2010–2012 (n = 2) 2013–2018 (n = 15)	20	1	7	2	5	4
Education	14	1980–1989 (n = 1) 1990–1999 (n = 2) 2000–2009 (n = 4) 2013–2018 (n = 7)	6	1	-	1	5	1
Science/math	6	1990–1999 (n = 1) 2000–2009 (n = 3) 2013–2018 (n = 3)	2	2	-	-	2	-
Social science	31	1960–1969 (n = 1) 1970–1979 (n = 4) 1980–1989 (n = 5) 1990–1999 (n = 13) 2000–2009 (n = 7) 2013–2018 (n = 1)	26	-	1	2	2	-
Other	4	1990–1999 (n = 1) 2000–2009 (n = 1) 2010–2012 (n = 2)	1	-	1	-	2	-

The mapping phase was performed by B.V. and J.C. Each article was characterized by its discipline, year of publication, journal title, and type of publication (research report, editorial, opinion, conceptual article, literature review, and methods article). The disciplinary field was identified by looking at the titles of the journal and the article.

A deductive coding framework was developed to guide the data extraction process [17]. Data extraction was performed by one reviewer (J.C., B.V., E.B. (see acknowledgments) or S.K.) using the NVivo software. Random checks were performed by a second reviewer (J.C. or B.V.) on 10% of included articles in each field. Coding disagreements were discussed and resolved through consensus between reviewers. Data were summarized by two authors (J.C. and B.V.) using paradigm bridging (seeking commonalities in underlying conceptual and theoretical assumptions found in the literature) and paradigm bracketing (highlighting differences or conflicting data from research to generate higher-order insights) to describe the different meanings of replication within different research fields [20]. Similarities and contrasts within and between fields were noted. After the updated search, tables were updated to include replication definitions and replication types described in the more recent articles.

The research team involved various stakeholders including researchers, research funders, journal editors, and government policy makers. They were consulted for methodological decisions and provided ongoing feedback on the analysis. Three consensus meetings were organized to validate with the team: 1) the search strategies, 2) data extraction and analysis procedures, and 3) synthesis of the results obtained before the updated search.

3. Findings

We included 153 papers in this concept analysis. The review flow diagram is presented in Appendix 3. All articles were sorted in accordance with disciplinary fields and publication type (Table 2). The health field contributed the greatest number of articles (43 of 153) followed by psychology (39 of 153) and social science (31 of 153). The

Table 3. Examples of replication definitions found in different research fields

Business	Replication is traditionally defined as the duplication of a previously published empirical study to determine whether the findings of that study are repeatable [30]. A replication is defined as a duplication of a previously published empirical study that is concerned with assessing whether similar findings can be obtained on repeating the study [36].
Education	Replication is a method of validating research in the sciences and is considered a criterion for the acceptance of new theories and knowledge. In the context of testing hypotheses through empirical research, replication requires experiments to be repeated in such a way that investigators can be reasonably certain that the results of two experiments are comparable and, therefore, that they are measuring the same phenomenon [37]. Replication is the repetition of an experiment following, as closely as possible, the original experiment [29].
Health	Replication is a multifaceted, creative, significant methodology which provides an objective, credible means of evaluating the results and findings of nursing studies [2]. Replication research is the means of establishing the credibility of findings from a previous study. It involves reproducing the study design and methods with a new group of individuals to determine whether similar findings will be obtained [8]. Replication is the deliberate repetition of research procedures in a second investigation for the purpose of determining if earlier results can be repeated [22].
Social science	Replication should be the ''response'' that one researcher gives to another in following up which questions the original researcher claims to have answered. Replication studies necessarily make deliberate and explicit reference to one or more previous empirical investigations, whose findings or interpretations provide the specific issues of theory or method that are joined in a replication study [24]. Replications are often quite different from previous research; what is important is that they may replicate theoretical hypotheses, not research designs. In contrast to scientific research in general, which necessarily builds upon previous work, replication involves a direct linkage with a specific previous study. Replication, as here defined, cannot occur without an identifiable baseline study [32]. In its simplest form, replication involves doing the same study again to retest the same hypotheses, or explore the same issues, in the same way. The term, however, also has been used to refer to cases where the repeated study is conducted on a different population of participants; the researcher uses a different method of data collection or uses a different method of data analysis, and so on [33]. Replication is a methodological tool based on a repetition procedure that is involved in establishing a fact, truth, or piece of knowledge [35].
Psychology	Replication refers to a conscious and systematic repeat of an original study. That is, the general function is to re-examine explicitly a given study in part or in whole. It is not an investigation that simply acknowledges previous work [21]. To replicate an experiment means to duplicate all the conditions of experiment A in an experiment B and to achieve identical results in both experiments [38]. When a study is methodologically or conceptually similar to an earlier one, then the latter study is considered a replication [39].

most common type of included articles was conceptual (82 of 153).

3.1. Historical perspective

Included articles were published between 1926 and 2018, came from multiple disciplines, and consistently advocated for more replication studies. Some of them documented the lack of replication studies in their fields of research, lack of incentives to conduct this type of research, and the limited appeal replication has in the research community. Interest in research replication increased since 2000 in the fields of business, health, and social science. In psychology, several articles discussing replication were published in the past 10 years. This coincides with the ''replication crisis'', the growing recognition of the replication problem by researchers since 2010 [25,26]. Nevertheless, interests and concerns of more recent articles still focused on the need for more replication and lack of importance granted in all fields to this research activity [27,28].

Three key articles by Lykken (psychology) [23], La Sorte (psychology) [21], and Finifter (social sciences) [24] published between 1968 and 1975 provided definitions and typologies of replication that have had, based on their citation records, a lasting influence across disciplines. All three adopted a broad perspective of how replication should be defined and conducted.

3.2. Definition and uses of the concept in the literature

The first step in concept analysis is to identify as many uses of the concept by extensive reading in as many different sources as possible without limiting or ignoring some uses of the concept [16]. Multiple definitions of replications were reported (Table 3). Although research traditions in the natural and social sciences are very different, we found no distinct differences in definitions across disciplines. Some defined replication narrowly as the duplication or repetition of an original study [9,22,29,30,31]. Other definitions were broader and inclusive of different approaches [2,8,24,32,33,34] and suggesting that replication studies could be conducted using different methods, with different populations and in different contexts to the index study [8,35]. The analysis of the literature led to

Table 4. Illustration of how replication can be used for three different major purposes

Index study	Replication studies
In a randomized controlled trial, cognitive behavioral therapy was found to improve depression, anxiety, and substance use in a group of young adult men with concurrent depression and substance misuse. Duration and number of treatment sessions received appeared to influence the results.	Repetition To test the reproducibility of findings, this study can be repeated by recruiting another sample of young adult men from the same region, repeating exactly the same study methods and providing the same intervention by the same trained professionals. Extension To test the generalizability of findings, this study can be extended by recruiting another sample of young men and women from multiple regions with low and heavy substance misuse. Road-testing of a theory The theory underlying the index study intervention can be tested using treatment adaptations for different users with low or high substance misuse.

the identification of three major different uses of the concept that vary in accordance with the purpose of the replication study and the similarity between the replication and the index study. To illustrate these differences, an example is provided in Table 4.

3.2.1. Replication as the repetition of a previous study

Replication in its simplest form involves doing the same study again to test the reproducibility of findings (degree of agreement between results of experiments) or to check how results may have been affected by internal validity threats [2,8,21,22,24,27,30,34,35,36,40–46]. Reliability of findings is established when researchers are able to perform exactly the same experiment, under the same conditions and generate similar results. They can either perform an intrastudy replication [30,44,47], where replication is integrated within the same study (repeat the analysis using the same data set or recruiting two sample of participants simultaneously) or an interstudy replication, where the same research procedures are repeated but after the index study is completed [44,47].

Replication can also be used to check the ways in which threats to internal validity may have influenced the results obtained in the index study [33,37]. For example, it can help assess for type I or type II errors in the index study, assess how results were affected by problems of definition and particular measurement and analysis techniques, and assess how far the results of a study are generalizable within the same population and the same context (the degree to which inferences from a study can be generalized [48]) [12,22,31,49,50]. By repeating as closely as possible an index study, one can compare several effect estimates to evaluate the variation between these estimates and identify potential sources of errors such as the influence of possible confounding variables not assessed in the index study, differences between groups at baseline, or the possible influence of data collection strategies [33,51,52].

3.2.2. Replication as the extension of a previous study

Replication can also refer to a study that only has a certain level of similarity with the original experiment. In this case, researchers purposively make minor or important changes to the study to evaluate possible generalization and extension of previous research findings [1,2,12,31,33,34,35,43,46,47,53–58]. Examples would be changes in either the manipulated or measured variables, investigating the influence of additional variables, repeating the study using different populations, contexts, geographical areas, time periods, or using any combination of the aforementioned changes [59]. As argued by Rosenthal [60], the more changes in this kind of replication, the greater the likelihood of benefit to the external validity of the original finding, if results support the finding.

3.2.3. Replication as the road-testing of a theory

Some authors extend their definition of replication to a study that deliberately avoids imitation of the original investigators' methodology, sampling procedures, measurement techniques, and data analysis methods [1,2,8,9,13,23,31,35,50,52]. This kind of replication can verify earlier research but also broadens the scope of the findings and tests whether the original findings can stand as general principles [2,8,23]. This can provide support for theory and conceptual framework development and validation. The rationale for this approach to replication is that if the proposed theoretical relationship exists between two or more variables, then it should be demonstrable regardless of the method used in a particular study [8,9].

3.3. Related concepts

We identified in the literature related concepts that appeared distinct from replication: repetition, duplication, and reproduction. Finifter [24] warns that interpreting replication as merely repetition or duplication is simplistic and these terms should not be considered equivalent. Reproduction is considered a new experiment using different experimental methods to test the same hypotheses as the baseline experiment. This form of triangulation can be used to corroborate the findings of an index study or to verify if the results obtained in an experiment were influenced by methods and apparatus [61].

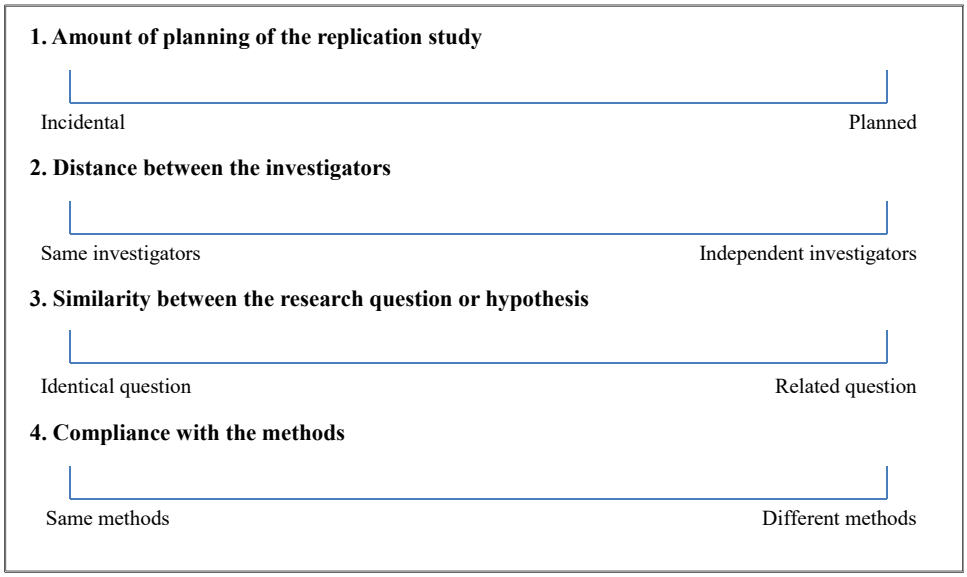

Fig. 1. Attributes of replication studies.

3.4. Attributes of replication studies

The aim of this third step in concept analysis is to extract from the literature a list of defining characteristics that gives broad insight into the concept and differentiates it from other phenomena [16]. We identified different attributes of replication, which allow us to characterize and compare the design and conduct of various replication studies for one or more purposes. Each of the attributes can be conceptualized on a continuum and varies depending on whether the replication study aims to be a repetition of the index study, an extension, or the road-testing of a theory (Figure 1).

The *amount of planning* is the extent to which the replication study is intentionally repeating the methods or testing the results of a previous study. Replication can be planned, deliberately and explicitly referencing an index study and developing a research protocol that is methodologically or conceptually similar [45]. Replication studies can also be incidental or ad hoc because they were not conceptualized as replications per se and may vary from the index study in multiple known and unknown ways [62]. These unknown and unplanned differences can make the comparison and interpretation of the results between the index and incidental replication studies more difficult [46]. Da Silva and colleagues (2014) mentioned that ‘‘empirical studies that address similar questions or hypotheses, but without explicit reference to one or more previous studies that can be considered the original study, should not be considered replications’’ [63]. Deliberate planning of replication studies is considered required to improve the development of highly credible research results [33,35,64].

The *distance between investigators* refers to the independence between researchers conducting the index and the replication study [65]. According to Rosenthal [34,60], the replication of a study that obtains similar results is more convincing if it is maximally separated from the first study along such dimensions as time, physical distance, personal attributes of the experimenters, experimenters’ expectancy, and experimenters’ degree of personal contact with each other. Some authors report that replications by different investigators are worth more than replications by the same investigator [66] because they eliminate the potential influence of vested interests or bias about the interpretation of data collected [67]. As mentioned by Ioannidis [64], replication studies conducted by the original researchers are certainly important and can be encouraged as a first step but they are not sufficient to establish convincing reliability of findings.

The *similarity between the research questions* relates to whether the replication study is investigating the exact same question as the original study or how they differ, thus determining for what purpose the replication study is conducted [12].

The *amount of compliance with the methods of the original study*[12,45,65] relates to whether the replication study uses the same or different methods to the original study A to create new means to verify the findings of a previous study, to address inadequacies in the original study, or to evaluate the influence the research procedures may have had on the findings.

3.5. Antecedents of replication studies

The next step in concept analysis involves identifying events or circumstances that are necessary for the occurrence of the concept [16]. Replication should be conducted when 1) appropriate and theoretically relevant research

Table 5. Analysis of replication typologies and types found in the literature

Type	Definition	Purpose	Fields/references
Replication as the repetition of a previous study			
Intrastudy replication (within an index study)			
Pseudo or internal replication	A type of replication study where data are collected at the same time as those for the original study. The simultaneous confirmation of the study is built into the original design.	Useful when exact or almost exact replications are considered to be impossible, impractical, or undesirable to conduct. Controls for historical bias as data are collected for both studies simultaneously.	Psychology [21,43,51] Social science [24,75] Health [1,8,31] Other [52]
Checking of analysis or restricted secondary analysis or verification or verifiability	A type of replication study using exactly the same procedures used in a past study to analyze the latter's data set.	Useful to check whether investigators of the original study have committed any errors in the process of analyzing the data.	Social science [13,33,71] Business [75]
Reanalysis of data or model comparisons or complete secondary analysis or robustness	A type of replication study using different procedures to reanalyze the data of a previous study.	Useful to assess whether similar findings can be obtained using a different methodology with the same data or a subsample of the data.	Social science [13,33,71] Business [76] Other [53]
Assumed replication	A type of replication study where the influence of extraneous factors (such as gender, education, social support, etc.) is explicitly incorporated in the design of the study.	Useful to examine the reproducibility of results and to establish the boundary conditions on the effects of a particular independent variable.	Business [44,47]
Interstudy replication (across different studies)			
Exact replication or identical replication or literal replication or real replication or pure replication or narrow replication or duplication or direct replication or method reproducibility or reproduction	A type of replication study where the procedures and conditions of the initial study are repeated and where all the conditions of the original research are maintained: same observer, same population, same procedures, same measurements, same context, and the same time.	Useful to determine if the findings of an original study are repeatable.	Social science [25,32,33,35,54,75,77] Health [1,2,8,9,31,50,58] Psychology [10,23,27,39,43,51,57,65,78–82] Education [37,67,80,83,84] Business [28,47,76,83] Science/math [12]
Statistical replication	A type of exact or identical replication study used to reduce sampling error to a small enough level to draw correct statistical conclusions.	Useful to increase the statistical power of a study in comparison with the index study	Psychology [76] Business [85]
Retest replication or virtual replication or close replication or repeatability	A type of replication study involving the repetition of an original study with few, if any, significant changes in the research design. The study is conducted under similar conditions, following the methods as closely as possible.	Useful to check the reliability of the original research, to uncover inconsistencies and errors in procedure and data analysis, to assess the extent to which the original results hold up against chance composition of the sample, hidden third factors, or other forms of methodological artifact.	Psychology [21,82] Social science [25,32,71,75] Health [1,31] Education [29,86] Business [42]
Operational replication	A type of replication study duplicating the exact sampling procedures and experimental conditions of an original study, but not the methods of analysis and/or the measurement techniques.	Useful to test the reliability of the results of the experimental conditions and testing the measurement techniques.	Psychology [23] Health [2,9] Business [28] Education [80,83] Science/math [12]
Replication as the extension of a previous study			
Replication with extension or conceptual replication or differentiated replication Partial replication	A type of replication altering various features of the original study via a combination of new and previously used methods and measures.	Useful to confirm the generalizability or external validity of the research. Checking if the results hold for a different population, in a different setting, or for a different modality	Social science [35,54] Psychology [10,23,27,39,43,65,74,79–82] Business [29,43,45,48] Education [29,67,80,83,84,86] Science/math [12]

(*Continued*)

Table 5. Continued

Replication as the extension of a previous study			
Empirical generalization or extension	A type of replication study repeating a past study on a different population.	Useful to test how far the results of the study are generalizable to another population.	Social science [33] Business [76]
Approximate replication	A type of replication study duplicating the methods of the initial study allowing for the unavoidable variations in conditions such as the location, time, participants, and investigators.	Useful to determine whether the original findings hold up despite modest changes in research conditions.	Health [8,31] Education [37]
Independent replication or systematic replication or result reproducibility	A type of replication study introducing significant modifications in the design of the original study. There is no intent on the part of the replicators to follow the design or methods of the earlier investigators; rather, the second investigative team begins with a similar problem statement but formulates new means to verify the first investigators' findings.	Useful to verify an empirical generalization that goes beyond those of reliability and confirmation, not only to validate previous work but also to extend the findings and tests the limits of generalizability.	Psychology [21,55,57] Social science [24,35,54] Health [1,31,56,58] Education [37]
Replication as the road-testing of a theory			
Conceptual extension or quasi-replication	A type of replication involving using procedures different from those of the original study such as measuring constructs, structuring the relationships among constructs, and analyzing data differently. In spite of these differences, the replication is based on the same theory as the original study.	Useful to revise a theory.	Social science [33] Business [28]
Construct replication or constructive replication or theoretical replication or generalization and extension or metareplication	A type of replication beginning with a similar hypothesis but using new methods of measurement and design. Researchers deliberately avoid imitation of the original investigators' methodology and develop their own design, sampling procedures, measurement techniques, and data analysis methods.	Useful to examine the feasibility of fitting empirical findings into a general theoretical framework, to assess the explanations that the original researcher gave to account for the results obtained in his or her study.	Health [1,2,8,9,31,50] Social Sscience [13,35,71] Education [37,80] Psychology [82] Other [52]

questions remain insufficiently answered and require further investigation, 2) controversies about the research results remain unresolved, 3) weaknesses in the original study were found that could be overcome by changing the design or procedures of the study, 4) important theories are only supported by a few small or low-quality studies, and 5) generalizability of the findings is limited because results may be only applicable to a specific population or setting [37]. To conduct any kind of replication study, it is recommended that investigators planning a replication study should be knowledgeable of the research subject [40], make sure they have enough information about the index study and a thorough understanding of the rationale underpinning the conduct of the index study [2,23,40,68], and define beforehand the relationship between the original and the replication study [21]. If researchers are replicating a study for the purpose of testing the reproducibility of results or testing their empirical generalizability, more specific conditions should be fulfilled; the original study or information available from researchers must provide sufficient detail to allow replication including the following: 1) a precise report of the methods, 2) a detailed description of the intervention (characteristics of those delivering the intervention, recipients, setting, mode of delivery, intensity, duration, and

adherence to the delivery protocol), and 3) a complete report of the results [2,27,36,37,40,42,68,69]. Furthermore, investigators should be willing and open to provide unpublished and complementary data and information to ensure appropriate replication of their study [36,68,70].

3.6. Consequences of replication studies

As part of the concept analysis, we also identified the consequences of the occurrence of the concept. Consequences of replication studies are related to the interpretation of the results (successful or unsuccessful replication).Results of replication studies can either confirm the results of the index study or fail to confirm these results [67,71,72]. This outcome can be influenced by the amount of similarity or differences between the attributes of the index and replication studies [71]. It may also be difficult to determine if the replication was successful or not.

When a replication study very similar to the initial study finds the same results, it adds little new evidence but it strengthens confidence in the results by demonstrating that the findings are either reproducible or valid in exact or similar conditions [72]. The probability of chance explaining the outcome is reduced [73], and the narrow interpretation of results from the index study has been strengthened. However, if the replication study repeats the errors from the original study, confirmation of the results will not necessarily be informative to the development of knowledge. The results may also be specific to the method, rather than a general truth [52]. If the replication study introduces greater differences from the initial study through modifications to the research question or methods, more information may be gained about the validity and generalizability of index study results and this ‘‘strengthens broad interpretation’’ [71].

When a replication study fails to confirm previous findings, the failure does not necessarily mean a conclusive falsification [72]. When very similar studies achieve different results, this can create doubt on both the index and replication study and ‘‘undermines narrow and broad interpretation’’. It is then not possible to know if the index study results are reliable. The index study results could have been a false positive, or the replication could produce a false negative [26]. It is also too easy to conclude that a failure to replicate a result means that the original evidence was a false positive [26]. When both studies differ in methods or have less similar research questions, it is easier for authors to explain the reasons for the failure of replication because they knowingly altered the initial study design. This ‘‘undermines broad interpretation’’ of the index study results but not ‘‘narrow interpretation’’ because differences between study findings are attributed to the differences between the studies [71]. However, results that do not replicate over time can also lead to improved understanding of a phenomenon and support the revision or strengthening of an intervention or theory and lead to better understanding of the generalizability of findings.

3.7. Model, borderline, and contrary cases of replication studies

The seventh step of the concept analysis methodology involves constructing a model of the phenomena of interest and identifying borderline and contrary cases. A model case shares all the key characteristics of the concept, a contrary case has none of the defining characteristics, and a borderline case might be a metaphoric use of the concept [16].

This review identified several model cases of replication. For example, in the health research, five reviews or conceptual articles on replication presented a typology of replication research. Other replication typologies were found in the fields of business [36,47,59], psychology [21,23,74], and social science [24,33]. To make sense and to organize this knowledge, we compared the multiple replication type definitions and typologies. We identified 32 labels for replication types that we grouped into 14 types in accordance with their definitions (Table 5). Eight types of replication were classified as belonging to the ‘‘repetition of a previous study’’ and were differentiated if they were intrastudy or interstudy replications [47]. Four types of replication were classified under the ‘‘extension of a previous study’’ and make minor or more important changes to the methods and/or research questions of the index study to evaluate the internal and external validity of the previous results. Finally, two types fell under the ‘‘road-testing of a theory’’ definition. They purposefully do not comply with the methods and design of the index study and aim to test a similar hypothesis or to test or revise a theory.

A borderline case of replication would be an unplanned or an ad hoc replication such as two researchers conducting very similar studies without being aware of it.

A contrary case of replication would be ‘‘not replicating’’ or simply acknowledging previous work without re-examining a given study in part or whole or investigating an issue raised from a past study from a different theoretical perspective and with different methodological imperatives. This would constitute new research rather than any kind of replication [21].

3.8. Theoretical definition of replication studies

Based on the definitions, attributes, antecedents, and consequences of replication identified in this concept analysis, our team formulated a theoretical definition of a replication study.

A replication study is a study that is methodologically or conceptually similar to an earlier one. It is a deliberate repetition of an index study in whole, in part, or conceptually. It provides a means to assess the reliability, validity, and/or generalizability of previous findings or theory. A

replication study may not only verify earlier research, but also may broaden the scope of the findings and test whether the original findings can stand as general principles. It is possible to describe how much variation exists between index and replication studies in terms of the amount of planning of the study, the distance between the investigators, and the similarity between the research questions or hypothesis and methods of the index study.

4. Conclusion

Advocating for more replication research should be important to all scientists, editors, funders, and the public. As mentioned by Jones [39], ''science in its ideal form is self-correcting: over time, ideas are corroborated or falsified. To corroborate or falsify any idea, researchers must replicate prior studies.'' Nevertheless, research methods textbooks rarely address the importance, the purposes, and the design of replication studies. There is a lack of consensus in the literature on a definition of ''what is'' and ''what is not'' replication in quantitative research [12].

We found multiple definitions and typologies of replication studies in different research fields. Criteria used to develop these typologies varied, and it was not possible to identify one that was complete or well adapted to guide the development and conduct of replication. Combining, comparing, and critically analyzing the literature across disciplines and authors led us to adopt a broad perspective of replication which included the repetition of an index study, the extension of an index study, and the road-testing of a theory. Other literature reviews on the subject also reported these three main subtypes of replication [12,68,82]. The combination of the concept analysis [16,17] and metanarrative review [18] approaches were useful to frame the relevant aspects and dimensions of the concept and to provide a sound structure to guide data extraction and analysis using an interpretative approach.

A limitation of this synthesis was the difficulty in capturing all the literature on the subject across different research fields. The combination of the snowball and systematic search method was valuable to allow us to identify the relevant literature for this synthesis. Nevertheless, managing the volume of included studies was a challenge. Because of the wide range of possible definitions of replication, we chose inclusion criteria that were more lenient and led us to include some articles that contributed less to our understanding of the concept. Furthermore, because of the amount of the literature used in this synthesis, the appraisal and synthesis phases were lengthy.

The thorough analysis of the literature allowed our team to develop a comprehensive definition of replication in quantitative research and provides a shared language around types and uses of different kinds of replication studies. This review highlighted important relationship between the purposes of replication studies and its attributes.

CRediT authorship contribution statement

Brigitte Vachon: Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Project administration, Funding acquisition. **Janet A. Curran:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - review & editing, Visualization, Project administration, Funding acquisition. **Sathya Karunananthan:** Investigation, Writing - original draft, Writing - review & editing. **Jamie Brehaut:** Conceptualization, Formal analysis, Writing - review & editing, Funding acquisition. **Ian D. Graham:** Conceptualization, Formal analysis, Writing - review & editing, Funding acquisition. **David Moher:** Conceptualization, Formal analysis, Writing - review & editing, Funding acquisition. **Anne E. Sales:** Conceptualization, Formal analysis, Writing - review & editing, Funding acquisition. **Sharon E. Straus:** Conceptualization, Formal analysis, Writing - review & editing, Funding acquisition. **Michele Fiander:** Methodology, Investigation, Writing - review & editing. **P. Alison Paprica:** Formal analysis, Writing - review & editing, Funding acquisition. **Jeremy M. Grimshaw:** Conceptualization, Methodology, Writing - original draft, Supervision, Funding acquisition.

Acknowledgments

The authors would like to acknowledge the contribution of different stakeholders who collaborated in the different phases of protocol development, data collection, and analyses (Martin Eccles, Denis Roy, Ann Mc Kibbon, and Patrick McGrath) and their research assistants Emma Bell for her support during the data extraction phase and Amanda Dresch during the mapping phase.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.07.006.

References

[1] Beck CT. Replication strategies for nursing research. Image J Nurs Sch 1994;26:191−4.
[2] Connelly CE. Replication research in nursing. Int J Nurs Stud 1986;23:71−7.
[3] Martin PA. More replication studies needed. Appl Nurs Res 1995;8:102−3.
[4] Grimshaw J. A Knowledge Synthesis Chapter. Ottawa: Canadian Health Research Institute; 2010.
[5] Mulrow CD. Rationale for systematic reviews. BMJ 1994;309:597−9.
[6] Ioannidis JP. Large scale evidence and replication: insights from rheumatology and beyond. Ann Rheum Dis 2005;64:345−6.
[7] Niven DJ, McCormick TJ, Straus SE, Hemmelgarn BR, Jeffs L, Barnes TRM, et al. Reproducibility of clinical research in critical care: a scoping review. BMC Med 2018;16:26.

[8] Klein JG, Brown GT, Lysyk M. Replication research: a purposeful occupation worth repeating. Can J Occup Ther 2000;67:155−61.
[9] Gould D. Using replication studies to enhance nursing research. Nurs Stand 2002;16:33−6.
[10] LeBel E. A new replication norm for psychology. Collabra Psychol 2015;1. Article 4.
[11] Mann WC. Replication. Am J Occup Ther 1987;41:127−9.
[12] Gómez OS, Juristo N, Vegas S. Understanding replication of experiments in software engineering: a classification. Inf Softw Technol 2014;56:1033−48.
[13] van IJzendoorn MH. A process model of replication studies: on the relation between different types of replication. In: van der Veer R, Van IIJzendoorn MH, Valsiner J, editors. Reconstructing the mind: Replicability in research on human development. Ablez, NJ: Norwood; 1994:57−70.
[14] Glasgow RE, Emmons KM. How can we increase translation of research into practice? Types of evidence needed. Annu Rev Public Health 2007;28:413−33.
[15] Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. Health Aff (Millwood) 2005;24:138−50.
[16] Walker LO, Avant KC. Strategies for Theory Construction in Nursing. Upper Saddle: Pearson Prentice Hall; 2019.
[17] Walker LO, Avant KC. Strategies for Theory Construction in Nursing. Vol. 4. Upper Saddle: Pearson Prentice Hall; 2005.
[18] Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist synthesis: An introduction. ESRC Research Methods Programme. Manchester: University of Manchester; 2004.
[19] Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. Soc Sci Med 2005; 61:417−30.
[20] Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: meta-narrative reviews. J Adv Nurs 2013;69:987−1004.
[21] La Sorte MA. Replication as a verification technique in survey research: a paradigm. Sociological Q 1972;13:218−27.
[22] Fahs PS, Morgan LL, Kalman M. A call for replication. J Nurs Scholarsh 2003;35:67−72.
[23] Lykken DT. Statistical significance in psychological research. Psychol Bull 1968;70:151−9.
[24] Finifter BM. Replication and extension of social research through secondary analysis. Soc Sci Inf 1975;14:119−53.
[25] Diener E, Biswas-Diener R. The replication crisis in psychology. In: Diener E, editor. NOBA Project, Psychology as a Science.
[26] Nosek BA, Aarts AA, Anderson CJ, Anderson JE, Kappes HB. Estimating the reproducibility of psychological science. Science 2015; 349. aac4716-aac.
[27] Earp BD, Trafimow D. Replication, falsification, and the crisis of confidence in social psychology. Front Psychol 2015;6:621.
[28] Hensel PG. Supporting replication research in management journals: qualitative analysis of editorials published between 1970 and 2015. Eur Manag J 2019;37:45−57.
[29] Grubisic A, Stankov S, Rosic M, Zitko B. Controlled experiment replication in evaluation of e-learning system's educational influence. Comput Educ 2009;53:591−602.
[30] Singh K, Ang SH, Leong SM. Increasing Replication for Knowledge Accumulation in Strategy Research. J Manag 2003;29:533−49.
[31] Haller KB, Reynolds MA. Using research in practice: a case for replication in nursing–Part two. West J Nurs Res 1986;8:249−52.
[32] Bahr HM, Caplow T, Chadwick BA, Middletown III. Problems of replication, longitudinal measurement, and triangulation. Annu Rev Soc 1983;9:243−64.
[33] Tsang EWK, Kwan KM. Replication and theory development in organizational science: a critical realist perspective. Acad Manag Rev 1999;24:759−80.
[34] Rosenthal R. Some issues in the replication of social sicence research. Labour Econ 1997;2:121−3.
[35] Schmidt S. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. Rev Gen Psychol 2009;13:90−100.
[36] Evanschitzky H, Baumgarth C, Hubbard R, Armstrong JS. Replication research's disturbing trend. J Business Res 2007;60:411−5.
[37] Language Teaching Review, P.. Replication studies in language learning and teaching: questions and answers. Lang Teach 2008;41: 1−14.
[38] Potts A. The fallacy of replication in psi research. Parapsychol Rev 1980;11:14−5.
[39] Jones KS, Derby PL, Schmidlin EA. An investigation of the prevalence of replication research in human factors. Hum Factors 2010; 52:586−95.
[40] Kugler C, Fischer S, Russell CL. Preparing a replication study. Prog Transplant 2006;16:15−6.
[41] Herrnson PS. Replication, verification, secondary analysis, and data collection in political science. Polit Sci Polit 1995;28:452−5.
[42] Lehmann S, Bengart P. Replications hardly possible: reporting practice in top-tier marketing journals. J Model Manag 2016;11:427−45.
[43] Zwaan RA, Etz A, Lucas RE, Donnellan MB. Improving social and behavioral science by making replication mainstream: a response to commentaries. Behav Brain Sci 2018;41:e157.
[44] Kwon ES, Shan Y, Lee JS, Reid LN. Inter-study and intra-study replications in leading marketing journals: a longitudinal analysis. Eur J Marketing 2017;51:257−78.
[45] Brandt MJ, IJzerman H, Dijksterhuis A, Farach FJ, Geller J, Giner-Sorolla R, et al. The replication recipe: what makes for a convincing replication? J Exp Soc Psychol 2014;50:217−24.
[46] Coyne MD, Cook BG, Therrien WJ. Recommendations for replication research in special education: a framework of systematic, conceptual replications. Remedial Spec Educ 2016;37:244−53.
[47] Easley RW, Madden CS, Dunn MG. Conducting marketing science: the role of replication in the research process. J Business Res 2000;48:83−92.
[48] Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference/William R. Shedish, Thomas D. Cook, Donald T. Campbell. Boston: Houghton Mifflin; 2002.
[49] Ross LM, Hall BA, Heater SL. Why are occupational therapists not doing more replication research? Am J Occup Ther 1998;52: 234−5.
[50] Blomquist KB. Editorial: replication of research. Res Nurs Health 1986;9:193−4.
[51] Drotar D. Editorial: a call for replications of research in pediatric psychology and guidance for authors. J Pediatr Psychol 2010;35:801−5.
[52] Johnson DH. The importance of replication in wildlife research. J Wildl Manag 2002;66:919−32.
[53] Evanschitzky H, Armstrong JS. Replications of forecasting research. Int J Forecast 2010;26:4−8.
[54] Hendrick C, Replications, Strict Replications, and Conceptual Replications. Are they important? J Soc Behav Pers 1990;5:41−9.
[55] Kratochwill TR, Stoiber KC, Gutkin TB. Empirically supported intervention in school psychology: the role of negative results in outcome research. Psychol Sch 2000;37:399−413.
[56] Onghena P, Edgington ES. Customization of pain treatments: single-case design and analysis. Clin J pain 2005;21:56−68. discussion 9-72.
[57] Iso-Ahola SE. Reproducibility in psychological science: when do psychological phenomena exist? Front Psychol 2017;8:879.
[58] Goodman SN, Fanelli D, Ioannidis JP. What does research reproducibility mean? Sci Transl Med 2016;8. 341ps12-ps12.
[59] Hubbard R, Vetter DE. An empirical comparison of published replication research in accounting, economics, finance, management, and marketing. J Business Res 1996;35:153−64.
[60] Rosenthal R. Replication in behavioural research. In: Neuliep JW, editor. Handbook of replication research in the behavioral and social sciences. Newbury Park, CA: Sage; 1990:1−30.

[61] Juristo N, Gómez OS. Replication of software engineering experiments. In: Meyer B, Nordio M, editors. Empirical Software Engineering and Verification. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012:60−88.
[62] Valentine JC, Biglan A, Boruch RF, Castro FG, Collins LM, Flay BR, et al. Replication in prevention sciences. Prev Sci 2011;12:103−17.
[63] Da Silva FQ, Suassuna M, França ACC, Grubb AM, Gouveia TB, Monteiro CV, et al. Replication of empirical studies in software engineering research: a systematic mapping study. Empir Softw Eng 2014;19:501−57.
[64] Ioannidis JP. Why science is not necessarily self-correcting. Perspect Psychol Sci 2012;7:645−54.
[65] LeBel EP, McCarthy RJ, Earp BD, Elson M, Vanpaemel W. A unified framework to quantify the credibility of scientific findings. Adv Methods Pract Psychol Sci 2018;1:389−402.
[66] Lamal PA. On the importance of replication. J Soc Behav Pers 1990; 5:31−5.
[67] Makel MC, Plucker JA. An introduction to replication research in gifted education: shiny and new is not the same as useful. Gift Child Q 2015;59:157−64.
[68] Marsden E, Morgan-Short K, Thompson S, Abugaber D. Replication in second language research: narrative and systematic reviews and recommendations for the field. Lang Learn 2018;68:321−91.
[69] Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implement Sci 2009;4:40.
[70] Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. Promoting an open research culture. Science 2015;348:1422−5.
[71] Freese J, Peterson D. Replication in social science. Annu Rev Soc 2017;43:147−65.
[72] Travers JC, Cook BG, Therrien WJ, Coyne MD. Replication research and special education. Remedial Spec Educ 2016;37:195−204.
[73] Gasparikova-Krasnec M, Ging N. Experimental replication and professional cooperation. Am Psychol 1987;42:266−7.
[74] Hunter JE. Re-inquiries: the desperate need for replications. J Consumer Res 2001;28:149−58.
[75] Dow MM. Subsample replication in cross-cultural surveys. Am Anthropol 1987;89:125−33.
[76] Clemens MA. The meaning of failed replications: a review and proposal. J Econ Surv 2017;31:326−42.
[77] Mulkay M. The scientist talks back: a one-act play, with a moral, about replication in science and reflexivity in sociology. Soc Stud Sci 1984;14:265−8.
[78] Heyman RE, Hunt AN. Repliction in observational couples research: a commentary. J Marriage Fam 2007;69:81.
[79] Laws KR. Psychology, replication & beyond. BMC Psychol 2016;4:30.
[80] Makel MC, Plucker JA. Facts are more important than novelty: replication in the education sciences. Educ Res 2014;43:304−16.
[81] Finkel EJ, Eastwick PW, Reis HT. Replicability and other features of a high-quality science: toward a balanced and empirical approach. J Pers Soc Psychol 2017;113:244−53.
[82] Hüffmeier J, Mazei J, Schultze T. Reconceptualizing replication as a sequence of different studies: a replication typology. J Exp Soc Psychol 2016;66:81−92.
[83] Ahadi A, Hellas A, Ihantola P, Korhonen A, Petersen A. Replication in computing education research: researcher attitudes and experiences. In: Proceedings of the 16th Koli Calling International Conference on Computing Education Research. ACM; 2016:2−11.
[84] Cook BG, Collins LW, Cook SC, Cook L. A replication by any other name: a systematic review of replicative intervention studies. Remedial Spec Educ 2016;37:223−34.
[85] Hamermesh DS. Viewpoint: replication in economics. Can J Econ 2007;40:715−33.
[86] Polio C, Gass S. Replication and reporting: a commentary. Stud Second Lang Acquis 1997;19:499−508.