Identifying Effective Instructional Design Features of Simulation in Nursing Education: A Realist Review

by

Nazlin Hirji

A thesis submitted in conformity with the requirements for the degree of Master of Science
Institute of Health Policy, Management and Evaluation
University of Toronto

© Copyright by Nazlin Hirji 2015
Identifying Effective Instructional Design Features of Simulation in Nursing Education: A Realist Review

Nazlin Hirji

Master of Science
Institute of Health Policy, Management and Evaluation
University of Toronto

2015

Abstract

Simulation is increasingly being used as an instructional strategy, yet evidence-informed guides for how to use simulation effectively in nursing training are lacking. A realist review of the nursing simulation literature was conducted to determine how simulation is currently being used in nursing training. Unfortunately, the majority (69%) of primary studies were excluded due to missing information about the context, mechanism, or outcome of the study. In the remaining 13 studies, four themes of instructional design features emerged including: feedback, debriefing, self-regulated learning, and the intersection between feedback and self-regulated learning. However, the small number of studies permits only cautious recommendations for nursing educators. Improved reporting standards in nursing education journals are imperative so that authors of original studies are prompted to report all relevant information, and so authors of knowledge syntheses are better able to make recommendations about when simulation should be used, for which trainees, under which circumstances, and why.
Acknowledgments

I would like to acknowledge and sincerely thank my supervisors, Dr. Rhonda Cockerill and Dr. Ryan Brydges for sharing their extensive expertise and wisdom with me. I cannot express how much I have learned and how much I appreciate all their support and encouragement.

Special thanks also goes to Dr. Joan Samuels-Dennis and Sandra Cop for sharing their valuable time and knowledge regarding research methodology and nursing simulation, and for their acceptance and openness to this new approach.

Finally, I would like to thank my family and colleagues for their support, patience, encouragement, and mere tolerance (at times) for my love of learning which results in a never-ending stream of articles that must be read and papers that must be written! I couldn’t do it without you!
# Table of Contents

## Table of Contents

Acknowledgments ........................................................................................................ iii

List of Tables ................................................................................................................. vi

List of Figures ................................................................................................................ vii

List of Appendices ......................................................................................................... viii

1 Introduction ................................................................................................................... 1

2 Literature Overview ..................................................................................................... 3

  2.1 Nursing Simulation – Brief History and Current Status ........................................ 3

  2.2 Literature Review - Health Professions Education .............................................. 5

  2.3 Literature Review - Nursing Reviews .................................................................. 8

  2.4 Critique of Previous Reviews of Evidence for Simulation-Based Training .......... 10

  2.5 Opportunities ....................................................................................................... 13

3 Manuscript For Peer-Reviewed Journal ................................................................. 14

  3.1 Introduction ........................................................................................................... 14

    3.1.1 Objectives and Focus of the Review ............................................................. 17

  3.2 Methods ................................................................................................................ 18

    3.2.1 Philosophy Underpinning the Realist Review and Synthesis ................. 18

    3.2.2 The Present Realist Review and Synthesis ............................................. 21

  3.3 Results ................................................................................................................... 26

    3.3.1 Study characteristics .................................................................................. 28

    3.3.2 Overview of Realist Synthesis Findings ................................................. 30

    3.3.3 Realist Theme 1: Feedback ..................................................................... 31
List of Tables

Table 1: Key Study Features
List of Figures

Figure 1: Trial Flow of Included and Excluded Studies
List of Appendices

Appendix 1: Data Extraction Form 1

Appendix 2: Data Extraction Form 2

Appendix 3: Data Extraction Form 3

Appendix 4: List of Excluded Studies (missing C-M-O data)
1 Introduction

The study and use of simulation is proliferating in health-related research, education and practice. Broadly defined, simulation is “the technique of constructing and running a model of a real system in order to study the behaviour of that system, without disrupting the environment of the real system” (Bobillier, Kahan, & Probst, 1976, as cited in Pritsker, 1979). As we learn more, we recognize opportunities to further explore, learn and refine how simulation is being used. As a nursing leader and educator, I (primary author – NH) wanted to understand the state of the nursing simulation literature and practice. I wanted to explore what is currently known about nursing simulation and how simulation is presently being used in nursing training. As such, my study question became “What simulation design features are most effective in enhancing nursing trainees' learning and clinical care outcomes?”

Given the breadth of literature, the complexity of nursing training, and the variety of nursing simulation review articles that have already been published, I recognized that this question could not be sufficiently answered using a traditional review approach. Instead, I wanted to examine the literature using a different lens, and offer the nursing community something distinct. Thus, this thesis was undertaken using a relatively new knowledge synthesis method, called realist review, with the aim of providing a thorough report on the current, cumulative state of the literature on nursing simulation, and to identify gaps and opportunities for research. The realist review methodology facilitates the identification of connections between contexts, mechanisms, and outcomes (more on these below) to a greater depth than traditional reviews. The underlying premise of this methodology is to help make sense of the literature and practice in order to understand “what works, for whom, under what circumstances”. An explanation of the methodology and the details of the process will be described below.

This thesis is organized in a “peer-reviewed publication” format. It is structured with an extended literature review in Chapter 2, outlining the state of the nursing simulation literature, and provides a critique, identifying the opportunities that can be addressed through this review. The manuscript for consideration of publication is provided in full in Chapter 3, and thus some duplication will be apparent. Next, an expansion of the implications is provided in Chapter 4, with a detailed discussion including recommendations for educators and researchers. Finally,
Chapter 5 outlines my final conclusions. Appendices include examples of the data extraction forms and a list of the studies that were excluded from the synthesis.

I worked closely with RB and also involved two additional experts (JSD and SC) in the data extraction process. Thus I have tried to delineate where we worked as a team, where only RB and I worked together, and where I worked independently.
2 Literature Overview

The term “simulation” has become commonplace in the vernacular of many disciplines in the health sciences (Cook et al., 2011). The use of simulation is increasing in education and practice and thus the literature that explores, researches, and analyzes simulation is also expanding (Cook et al., 2012). In 1976, Bobillier, Kahan and Probst defined simulation as “the technique of constructing and running a model of a real system in order to study the behaviour of that system, without disrupting the environment of the real system.” (as cited in Pritsker, 1979). Within the various health professions, the term simulation has evolved, been refined and taken on its own nuanced definitions. For the purposes of this review, we used the broad definition of ‘technology-enhanced simulation’ provided by Cook et al. (2011): “Simulation technologies encompass diverse products including computer-based virtual reality simulators, high-fidelity and static mannequins, plastic models, live animals, inert animal products, and human cadavers.” Based on this definition, we excluded studies involving human patient actors (standardized patients) or simulation for non-education purposes such as procedural planning, disease modeling, or evaluating the outcomes of clinical interventions.

As the body of evidence for simulation has accumulated, advanced and matured, literature reviews have been conducted to summarize and organize the data. While each review in nursing simulation has augmented the knowledge in the field individually, when analyzed together, gaps and opportunities become evident. As a result, this thesis was undertaken using a relatively new knowledge synthesis method, called realist review, with the aim of providing a thorough report on the current, cumulative state of the literature on nursing simulation, and to identify gaps and opportunities for research.

2.1 Nursing Simulation – Brief History and Current Status

While references to the use of “a mechanical dummy” in nursing training date as far back as 1847, the use of medium and high-fidelity simulators in undergraduate nursing programs began in the late 1990s, and accelerated in the mid-2000s as faculty realized the potential benefits of simulation, including allowing students to practice skills, critical thinking and clinical decision-making in a safe environment (Hayden, Smiley, Alexander, Kardong-Edgren, & Jeffries, 2014). With the availability of high-fidelity simulators, and as the use of simulation increased, this
instructional design method began to be viewed as the solution to the challenge of providing high-quality clinical experiences facing nursing schools. Educators have traditionally sought clinical placements in various practice settings to give nursing students the opportunity to demonstrate, practice and refine their technical, interpersonal, and critical thinking skills. However, access to sufficient numbers of high-quality clinical experiences has been steadily decreasing. In particular, access to valuable clinical placements has been limited due to faculty shortages, longer practice-setting orientations, restrictions on what nursing students can do during the placement to safeguard patient safety, and as the number of nursing programs competing for placements has grown (Hayden et al., 2014).

In response to the limited clinical opportunities for their trainees, nursing schools began exploring the possibility of replacing clinical placements with simulation experiences. Regulatory bodies including Boards of Nursing in the U.S.A. lacked the data to make an informed decision. Consequently, Hayden et al. (2014) recently undertook a large-scale, randomized, controlled study to compare traditional clinical experiences to replacement simulation experiences across the entire nursing curriculum. Students from ten nursing programs (five associate degree and five bachelor’s degree), ranging from community colleges to large universities, representing rural and metropolitan communities, were randomized into one of three study groups: (i) students who had traditional clinical experiences, (ii) students who had 25% of their clinical hours replaced by simulation, and (iii) students who had 50% of their clinical hours replaced by simulation. For the simulations, the researchers undertook a broad consultative process to develop a standardized simulation curriculum, by surveying nurse educators and clinicians about key concepts, student behaviours, patient conditions and health topics students should know. The participating nursing schools then validated and ranked the key concepts and patient conditions according to their own programs. Based on the final identified topics (including safety, communication and patient and family education), the research team assembled a bank of simulation scenarios provided by publishers / vendors, or acquired through donations (upon request) for topics not commercially available. All scenarios conformed to the NLN / Jeffries Simulation Framework (i.e. they had clear learning objectives, scenarios included problem-solving components, the level of fidelity was appropriate for the learning objectives, and structured debriefing followed each scenario, using the Debriefing for Meaningful Learning® Method), and could be chosen for use by the participating schools, based on course learning
outcomes (Hayden et al., 2014). Schools were not required to use any specific scenarios but when offering simulation, only scenarios from the ones provided could be used. To ensure consistency, each school was provided with mannequin programming files for each scenario. Unfortunately, the authors did not provide specific details about the simulations, including how they were conducted by each school, in each program, and within each course. The students and graduates were assessed by preceptors, instructors and hiring managers on a variety of measures including clinical competence, nursing knowledge, readiness-to-practice, in addition to how well their needs were met during clinical placements and during simulation experiences. Hayden et al. (2014) found no differences in any of the measured areas, and concluded that there is “substantial evidence” that comparable outcomes can be achieved when up to 50% of traditional clinical hours are replaced by high-quality simulation.

Given the persuasive nature of this evidence and the significant decision to potentially replace practical, real-life clinical experiences with simulation, it is imperative for nursing faculty to know how to design simulation training using evidence-informed instructional methods. The presence or absence of simulation experiences alone will not guarantee technical, interpersonal and critical thinking skills. Thus, further exploration of the simulation-based training approach, it’s educational design, and how learning is measured requires rigorous study of what works, for whom and under what circumstances (Pawson, Greenhalgh, Harvey, & Walshe, 2004).

2.2 Literature Review - Health Professions Education

With the widespread use of anything new, comes the desire and obligation to investigate its effectiveness. The same is true for the use of simulation. As a novel teaching approach, it was necessary to explore the actual efficacy of simulation. Thus within the health professions education literature, numerous studies and reviews have been conducted across multiple health professions comparing simulation to no intervention. The unequivocal result is that simulation is more effective than no intervention (Cook et al., 2011; Ilgen, Sherbino, & Cook, 2013; Lorello, Cook, Johnson, & Brydges, 2014). This issue has been explored at length, including within clinical specialties such as emergency medicine (Ilgen et al., 2013) and anaesthesiology (Lorello et al., 2014), as well as specific skills such as cardiac auscultation (McKinney, Cook, Wood, & Hatala, 2012), and laparoscopic surgery (Zendejas, Brydges, Hamstra, & Cook, 2013). The
results are consistent and thus further such studies will likely not add anything new to the body of simulation evidence.

Logically, the next area to investigate is how simulation compares to other instructional methods. Again, through a systematic review and meta-analytical process, Cook et al. (2012) have demonstrated that in comparison to other instructional modalities, technology-enhanced simulation training is associated with small to moderate positive effects. Perhaps more importantly, subgroup analysis suggests that strong instructional design is in fact partially responsible for the observed effects of improved knowledge, skills and behaviours (Cook et al., 2012). Thus, the next steps for studying simulation-based training appears to be to ask questions, such as: when is simulation more, less or not effective at all? Can less costly types of simulation be used in certain circumstances, and if so, when? And in cases where simulation will be effective, what is the best way to use it? Building on the work of Issenberg et al. (2005), Cook et al. (2013) present twelve features of effective simulation. The evidence indicates that the following eleven instructional design features, when they are available consistently, are associated with positive outcomes in comparison to when they are not available: range of difficulty, repetitive practice, distributed practice, cognitive interactivity, multiple learning strategies, individualized learning, mastery learning, feedback, greater time spent learning, clinical variation and curriculum integration. A twelfth design feature (group instruction) was associated with unfavourable results.

Fidelity, often thought of as the degree of realism in a simulation, is a particularly popular topic in much of the healthcare simulation literature. Two recent papers have problematized how the concept of fidelity has been examined in the literature (Norman, Dore, & Grierson, 2012; Hamstra, Brydges, Hatala, Zendejas, & Cook, 2014). As Hamstra et al. (2014) explain, consensus on a definition of fidelity is lacking as it varies depending on the simulation context, who is using the simulator, and how it is being used. Norman et al. (2012) explain the difference between engineering fidelity (whether the simulation looks realistic) versus psychological fidelity (whether specific behaviours are required during the simulation), suggesting that the two types of fidelity can be separated and that psychological fidelity may be more important in facilitating learning and transfer of knowledge. Taking it a step further, Hamstra et al. (2014) recommend abandoning the use of the term fidelity completely and instead replacing the term structural fidelity (physical realism) with physical resemblance, and replacing the term functional
fidelity (how closely the simulation resembles the clinical behaviour) with functional task alignment. The new terminology outlined in either of these papers has yet to permeate the entire literature, though the sentiment clearly marks a need to use the term ‘fidelity’ more carefully and thoughtfully.

Additionally, Norman et al. (2012) explore the common assumptions regarding the use of simulation in education and present caveats for those assumptions. The assumptions and associated limitations include: (i) as demonstrated by Cook et al. (2011), simulator usage in education results in meaningful learning, however the studies demonstrating those findings often only compare simulator training to no training, (ii) skills acquired through simulation can be transferred to real patients, however again, the control group received no formal training in such studies, (iii) there is better transfer of learning to real-life situations when the simulators are more authentic, however there is a lack of clarity about what specifically needs to be in place to ensure the simulation is more authentic, (iv) similarly, having the simulation resemble real-life situations is critical to the transfer of learning, but they question whether the authenticity must include both engineering fidelity and psychological fidelity, and finally (v), there is an assumption that more complex skills require more complex simulators, however, they found little supporting evidence and thus they encourage educators to consider the cost versus benefit of complex, high-fidelity simulators. Based on these assumptions, Norman et al. (2012) compare low-fidelity simulation with high-fidelity simulation and find that for auscultation skills, basic motor skills and complex crisis management skills, while high-fidelity simulation shows improvements in performance and transfer to real patient settings in comparison to no training or traditional training, the gains are more modest and not statistically significant in most cases, when compared to low-fidelity simulation.

Similarly, in order to better achieve the desired learning outcomes, Hamstra et al. (2014) recommend a shift in emphasis from physical resemblance to functional task alignment in simulation education, suggesting that the physical properties of the simulator are not the most important priority. Instead, the priority should be to select simulators and simulation scenarios that best enable the educator to meet their intended learning outcomes by focusing on the desired clinical behaviours and being selective and purposeful in the instructional design of the learning context in order to achieve them.
2.3 Literature Review - Nursing Reviews

The nursing education simulation literature has paralleled the broader health professions literature. Specifically, reviews have demonstrated that simulation is a useful teaching/learning tool (Cant & Cooper, 2010), that simulation as a teaching tool improves student learning (Harder, 2010), and that simulation is useful for knowledge, skills, safety and confidence (Norman, 2012). Weaver (2010) demonstrated that high-fidelity patient simulation benefits nursing students in knowledge, perceived value, realism, and learner satisfaction, while Yuan, Williams, Fang, & Ye, (2012) found that high-fidelity simulation enhances scores on knowledge and skill exams.

However, again, mixed results are evident, and the authors provide numerous caveats and discuss broad categories of issues. Firstly, a number of authors explain that poor study design plagues the nursing simulation research evidence and has diluted the effect sizes. Specifically, the individual studies have methodological deficiencies including small sample sizes, and a lack of high-quality randomized control trials. (Cant, & Cooper, 2010; Harder, 2010; Kardong-Edgren, Adamson, & Fitzgerald, 2010; Yuan, Williams, & Fang, 2011; Yuan et al., 2012).

Secondly, researchers’ inconsistent use of numerous measurement instruments has been listed as a concern. Kardong-Edgren et al. (2010) explain that there are too many different assessment tools being used and that most of them do not have established validity and reliability evidence supporting their use. They further suggest that simulation-based assessment tools should measure effects in the cognitive, affective and psychomotor learning domains. Kardong-Edgren et al. (2010) as well as Yuan et al. (2011) state that self-reported measures such as satisfaction and confidence are not objective and thus they recommend not using them. In fact, Kardong-Edgren et al. (2010) demonstrate that an inverse relationship between self-perceived confidence and actual skills may exist and thus using confidence as an outcome measure is problematic. Similarly, Harder (2010) and Yuan et al. (2011) claim that some of the assessment tools currently being used were designed for low-fidelity scenarios or for assessing students completing clinical placements and thus do not apply broadly to the spectrum of simulation-based training activities. Harder (2010) identifies the need to develop an assessment tool that is
scaffolded to adapt to the complexity of the simulation as well as the level of the expected learning outcomes.

The next group of findings the nursing simulation authors identify align with Cook et al.’s (2012) twelve instructional design features listed above. Namely, Cant and Cooper (2010) state that while simulation may be better than other teaching methods, to optimize its use depends on the context, topic and method. They expand to suggest that briefing, simulation and then debriefing are core components of successful simulations. Two reviews suggest that feedback may be the most important factor that influences learning (Cant and Cooper, 2010; Leigh, 2008). Cant and Cooper (2010) also suggest that exposure to simulation is more important than the specific role the students plays, while Leigh (2008) identifies the learner role as something to be further explored. However, in contrast to Cook et al.’s (2012) findings, Cant and Cooper (2010) found that the quantity of simulations the learner is exposed to does not necessarily result in superior learning. Finally, Cant and Cooper (2010) support the finding that to increase effectiveness, the simulation must parallel reality and that the simulation must be embedded into the broader learning curriculum.

Important recommendations from these nursing simulation reviews include the need: for interprofessional simulation research based on scenarios that replicate the true clinical environment (Harder, 2010; Kardong-Edgren et al., 2010), to differentiate the fidelity capability of a simulator from how it is actually used as an instructional method during specific simulation scenarios when conducting studies (Harder, 2010), to investigate the cost associated with the various forms of simulation in order to truly understand the cost and benefit of simulation in comparison to current methods (Weaver, 2010), to use teaching and learning, and nursing theory to design and interpret findings in research studies (Kaakinen, & Arwood, 2009; Rourke, Schmidt & Garga, 2010; Waldner & Olson, 2007), and finally, to critically investigate the transfer of simulation skills into the clinical setting (Cant, & Cooper, 2010; Norman, 2012; Weaver, 2010; Yuan et al., 2011).
2.4 Critique of Previous Reviews of Evidence for Simulation-Based Training

As discussed above, the literature on the use of simulation in health care professionals’ education has been evolving rapidly over time. As is to be expected with any body of evidence, the contributions have been valuable, but there have also been gaps in the reviews that have been conducted within the nursing-specific simulation literature. These gaps can be organized into a few categories, including methodology, results and conclusions.

With regards to methodology, when examining the group of nursing simulation literature reviews, a number of commonalities surface. Firstly, many of the reviews have been conducted by a single individual (Leigh, 2008; Harder, 2010; Norman, 2012; Weaver, 2011; Stroup, 2014; Shearer, 2013). In order to conduct an effective review, the work needs to be done by more than one person so that tasks such as selecting studies for eligibility and data extraction can be done by each of the reviewers independently, making it more likely that errors will be caught (Higgins, & Green, 2011). Furthermore, searches should be as extensive as possible in order to ensure the search is sensitive and captures a large number of relevant articles (Higgins, & Green, 2011). The process can be significantly enhanced by collaborating with a librarian / information professional who has expertise in conducting comprehensive, systematic, literature searches.

Unfortunately, the search terms and inclusion criteria in some of the nursing reviews were narrow. Examples include: (i) a study that included “systematic review” in it’s title, though the authors only used the search terms “simulation outcomes measurement” and “nursing education”, (ii) a literature review that limited the search years to between 2010 and 2013, and (iii) a literature review about successful debriefing that only included the search terms “simulation”, “debriefing” and “research”, limited to “meta-analysis”, “randomized controlled trial”, “review”, “comparative study” and “controlled clinical trial” (Norman, 2012; Skrable & Fitzsimons, 2014; Dufrene & Young, 2014). Only one review indicated that support from a library information specialist was accessed (Jansson, Kaariainen, & Kyngas, 2013). Similarly, some of the reviews limited the studies to nursing only (Leigh, 2008; Kaakinen, & Arwood, 2009; Weaver, 2011; Yuan et al., 2011; Norman, 2012; Foronda, Liu, & Bauman, 2013; Shearer, 2013; Dufrene & Young, 2014; Skrable & Fitzsimons, 2014; Stroup, 2014). However, as demonstrated above, and acknowledged by others (Paige & Morin, 2013; Franklin, Leighton,
Cantrell, & Rutherford-Hemming, 2015), there is much that can be learned from the other health professions and thus exclusion leads to a limited pool of knowledge and duplication of efforts. Furthermore, when conducting a literature review, there are published guidelines and standards available to reviewers to facilitate a more systematic process. One such example is Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA); an evidence-based minimum set of items for reporting (Moher, Liberati, Tetzlaff, & Altman, 2009). Through the use of PRISMA, the reviewers are directed to specify inclusion and exclusion (eligibility) criteria prior to examining the studies. For the nursing reviews, the use of such a systematic review process was not always evident (Hyland, & Hawkins, 2009; Kaakinen, & Arwood, 2009; Yuan et al., 2011; Foronda et al., 2013; Skrable & Fitzsimons, 2014). Another interesting phenomenon that limits the body of nursing simulation evidence is the focus on “fidelity”, and specifically the particular significance paid only to high-fidelity simulation (Leigh, 2008; Harder, 2010; Hyland, & Hawkins, 2009; Foronda et al., 2013; Dufrene & Young, 2014; Skrable & Fitzsimons, 2014; Jansson et al., 2013; Norman, 2012; Rourke et al., 2010; Shearer, 2013; Weaver, 2011; Yuan et al., 2011; Yuan et al., 2012). This singular view has negated the potential value of low-fidelity simulation, particularly in certain contexts, and has also blurred the lines between two distinct factors – the simulator and the simulation, the equipment and the intervention, the context and the mechanism. As noted by Harder (2010), the use of a mannequin capable of high-fidelity functions, does not guarantee that high-fidelity simulation is occurring.

The final methodological issue concerns the data analysis and contributes to a gap in the results obtained as well as the conclusions that can be drawn by the researchers. As the number of studies accumulate regarding a research topic, over time, there becomes enough information across numerous studies to pool the data and conduct meta-analyses. Meta-analysis is a useful tool to develop strong conclusions from the data and reduce the risk of bias and errors in interpretation of single studies and sets of data (Higgins, & Green, 2011). Unfortunately, this pooling and analysis technique was not used in the majority of the nursing reviews (Cant, & Cooper, 2010; Dufrene & Young, 2014; Foronda et al., 2013; Harder, 2010; Hyland, & Hawkins, 2009; Jansson et al., 2013; Leigh, 2008; Norman, 2012; Paige & Morin, 2013; Shearer, 2013; Skrable & Fitzsimons, 2014; Stroup, 2014; Weaver, 2011). As such, the power of the results are diminished due to the lack of pooling. When pooling is not done, the results often have to be presented in a narrative format. While narrative reviews have descriptive benefits, the inferences
and conclusions that can be drawn are severely limited as evidenced in the nursing simulation reviews (Green, Johnson, & Adams, 2006).

There are additional weaknesses with the conclusions that were drawn in the nursing simulation reviews. Specifically, given the methodological limitations described above, the inferences and judgments seem to reach beyond what the evidence demonstrates. This may be due to the single reviewer weakness described above. In fact, the conclusions are very subjective and far-removed from what the actual evidence conclusively indicates. Finally, as some authors have demonstrated (Kaakininen, & Arwood, 2009; Groom, Henderson, & Sittner, 2014; Rourke et al., 2010; Waldner & Olson, 2007; Franklin et al., 2015), there is a lack of grounding the evidence in any theoretical framework (Cant, & Cooper, 2010; Foronda et al., 2013; Harder, 2010; Hyland, & Hawkins, 2009; Jansson et al., 2013; Norman, 2012; Paige & Morin, 2013; Skrable & Fitzsimons, 2014; Stroup, 2014; Weaver, 2011; Yuan et al., 2011; Yuan et al., 2012). The lack of research theory leads to the methodological limitations described above, and the lack of educational and nursing theory makes it difficult to situate the findings within a larger context to facilitate broader understanding.

Franklin et al. (2015) recently published an invited article offering advice to novice simulation researchers to help them avoid common errors in conducting and reporting research studies. Their ideas address several of the critiques indicated above. Some of their suggestions include that the conceptual basis of the study should be grounded in a theoretical framework, the researchers should use rigorous research methods with reliable and valid evaluation instruments that match the outcomes, and the researchers must pay attention to intervention fidelity (i.e. validate scenarios with participants who are similar to the target population, prepare a protocol and detailed script for all simulation actors, and schedule meetings and rehearsals for the research team to ensure successful implementation of logistics). Franklin et al. (2015) end with four recommendations to novice simulation researchers including to review literature within and beyond their specific disciplines because there are learnings that transcend disciplines, to collaborate with other simulation researchers for the statistical, generalizability and educational value, to adopt and use consistent terminology regarding simulation, and to disseminate both the positive and negative outcomes of simulation research with the support of a seasoned mentor.
2.5 Opportunities

As demonstrated above, there is enough literature proving that simulation has a place in health professions education; it is better than no intervention and it is sometimes better than other instructional methods. The questions now requiring exploration include “What works? For whom? Under what circumstances? And why?” (Pawson et al., 2004). What these questions are really seeking to explore is the broad concept of instructional design using the realist review methodology (Pawson et al., 2004). Given that simulation is already being used in many nursing programs, and the idea of replacing clinical experiences with simulation experiences is being explored, the realist review methodology will enable us to clarify which instructional design features nursing educators are using in their practice, what the evidence is for the effectiveness of those design features, and to provide nursing simulation educators with pragmatic recommendations about how they can use simulation as an effective teaching and learning tool in their curricula.

Moving beyond previous systematic reviews of simulation in specific clinical specialties or skills (e.g., Ilgen et al., 2013; Lorello et al., 2014; McKinney et al., 2012; Zendejas et al., 2013), I believe that the realist review methodology holds promise in allowing us to synthesize the literature in novel ways that may produce better clarification on how best to use simulation in nursing education. In particular, my aim is to inform the nursing academic world more deeply about their use of simulation, specifically regarding when simulation should be used, for which students, under which circumstances, and why. More broadly, other health professions may also find some of the results useful within their own educational programs.
3  Manuscript For Peer-Reviewed Journal Submission

3.1  Introduction

The use of simulation in education and practice in the health professions, including nursing, has increased significantly over the past few decades (Cook et al., 2011; Franklin et al., 2015). Simulation has gained popularity for a number of reasons including: The Institute of Medicine report (Institute of Medicine, 2011) and The American Association of Colleges of Nursing (AACN) (American Association of Colleges of Nursing, 2008) recommendations to use simulation in nursing education (Weaver, 2011; Foronda et al., 2013; Skrable & Fitzsimons, 2014; Dufrene & Young, 2014), the continued limited access to clinical placements (Waldner & Olson, 2007; Weaver, 2011; Harder, 2010; Yuan et al., 2012), employers’ expectations that graduates are workplace ready and capable of caring for the complex patients they will be assigned (Leigh, 2008; Hyland, & Hawkins, 2009; Harder, 2010; Stroup, 2014), the advances in simulation technology and the associated complex clinical scenarios that can now be experienced in a simulation centre (Stroup, 2014), and the consistent, rigorous, and risk-free learning environment that simulation centres can provide (Leigh, 2008; Weaver, 2011; Yuan et al., 2012; Stroup, 2014). As such, simulation is being used ubiquitously to train practicing nurses and nursing students globally.

As the use of simulation increases, the literature that explores, researches, and analyzes its use also grows. (Cook et al., 2011; Franklin et al., 2015). Within the advancing body of simulation literature, reviews have been conducted to summarize and organize the data, offering valuable information. Within the nursing literature specifically, simulation has been shown to be a valid teaching and learning strategy for learners’ knowledge and skills (Cant & Cooper, 2010; Weaver, 2011; Harder, 2010; Yuan et al., 2012; Norman, 2012; Foronda et al., 2013; Skrable & Fitzsimons, 2014; Stroup, 2014), critical thinking ability (Cant & Cooper, 2010; Skrable & Fitzsimons, 2014; Stroup, 2014), and their self-reported satisfaction and confidence (Leigh, 2008; Skrable & Fitzsimons, 2014; Stroup, 2014; Cant & Cooper, 2010; Harder, 2010; Norman, 2012; Weaver, 2011; Yuan et al., 2012). In accomplishing these positive outcomes, instructional design has been identified as an important consideration (Cant, & Cooper, 2010; Leigh, 2008; Harder, 2010; Kaakinen, & Arwood, 2009; Kardong-Edgren et al., 2010; Weaver, 2011; Yuan et
al., 2011; Hyland, & Hawkins, 2009). For example, Cant and Cooper (2010) identified core components of successful simulation including: student orientation to a physical environment that parallels reality, curriculum-based scenarios incorporating best practice guidelines, and the presence of academic support during the simulation: a 3-step process that includes briefing, simulation and debriefing. The nursing literature contains a number of knowledge syntheses and literature reviews, yet specific recommendations about which instructional design features are most effective are currently lacking.

While each review in nursing simulation has augmented the knowledge in the field individually, when analyzed together, opportunities for further contribution also become evident. Methodologically, many of the existing reviews of the nursing simulation literature have been conducted by a single individual (Leigh, 2008; Harder, 2010; Norman, 2012; Weaver, 2011; Stroup, 2014; Shearer, 2013), the search terms have been narrow (Norman, 2012; Skrable & Fitzsimons, 2014; Dufrene & Young, 2014) and only one specifically mentions being supported by a librarian or information specialist (Jansson et al., 2013). Furthermore, many studies (Hyland, & Hawkins, 2009; Kaakinen, & Arwood, 2009; Yuan et al., 2011; Foronda et al., 2013; Skrable & Fitzsimons, 2014) do not report using a systematic review process such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009), do not explicitly report the theoretical framework used to ground the review (Cant, & Cooper, 2010; Foronda et al., 2013; Harder, 2010; Hyland, & Hawkins, 2009; Jansson et al., 2013; Norman, 2012; Paige & Morin, 2013; Skrable & Fitzsimons, 2014; Stroup, 2014; Weaver, 2011; Yuan et al., 2011; Yuan et al., & Ye, 2012), and have not capitalized on the opportunity to pool the cumulative data and conduct meta-analyses (Cant, & Cooper, 2010; Dufrene & Young, 2014; Foronda et al., 2013; Harder, 2010; Hyland, & Hawkins, 2009; Jansson et al., 2013; Leigh, 2008; Norman, 2012; Paige & Morin, 2013; Shearer, 2013; Skrable & Fitzsimons, 2014; Stroup, 2014; Weaver, 2011). These methodological issues have limited the conclusions and recommendations the reviewers have been able to make. Additionally, a consistent definition of simulation has not been adopted which has led to a very narrow conceptual understanding of what constitutes simulation (Groom et al., 2014). Consequently, the nursing simulation literature appears to have a bias towards “high-fidelity simulation” (HFS), often neglecting the literature on all other types of simulation (Leigh, 2008; Harder, 2010; Hyland, & Hawkins, 2009; Foronda et al., 2013; Dufrene & Young, 2014; Skrable & Fitzsimons, 2014; Jansson et al., 2013; Norman,
This singular view has negated the potential value of low-fidelity simulation, particularly in certain contexts, and has also blurred the lines between two distinct factors – the simulator and the simulation, the equipment and the intervention, the context and the mechanism. As noted by Harder (2010), the use of a mannequin capable of high-fidelity functions, does not guarantee that high-fidelity simulation is occurring.

Further knowledge can be gained by broadening a literature search to include simulation evidence across all health professions. Within the broader healthcare education literature, numerous studies have compared simulation to no intervention, with the unequivocal result being that simulation is more effective than no intervention (i.e. typical control group) (Cook et al., 2011). Furthermore, Cook et al. (2012) conducted a review and meta-analysis of studies that compared technology-enhanced simulation training to other instructional modalities (e.g., small group work, didactic lectures) and found that simulation training is often associated with higher learning outcomes than other instructional modalities. However, they also conducted subgroup analyses demonstrating larger effect sizes when the simulation intervention incorporated stronger instructional design, and smaller effect sizes when the comparison intervention incorporated stronger instructional design. Hence, instructional design once again stood out as a key theme to consider in simulation training. In a third systematic review that built on the work of Issenberg et al., (2005), Cook et al., (2013) evaluated twelve instructional design features of effective simulation and produced quantitative evidence supporting the use of several of those features in simulation training. In addition to this rich database, I believe a more focused and systematic review of the nursing simulation literature is necessary, especially in light of the limitations and associated opportunities identified in our review of previous nursing simulation reviews. While the existing evidence clearly demonstrates the value of simulation, and that certain instructional design features improve learning outcomes for the broader population of healthcare trainees, the existing literature does not specifically guide nursing educators on how and when to use simulation optimally. As such, I chose the realist synthesis methodology as the framework to conduct our review.
3.1.1 Objectives and Focus of the Review

Realist synthesis is a relatively new method for conducting systematic reviews of complex systems and multi-method, multi-disciplinary bodies of literature. Rather than for evaluating the efficacy of an intervention, the method is designed for reviewing evidence regarding complex social interventions, endeavouring to explain how and why they work, or do not work, in particular contexts (Pawson, Greenhalgh, Harvey, & Walshe, 2005). This is done through an iterative, interpretive, theory-driven narrative summary process using the findings from primary studies (Wong, Greenhalgh, Westhorp, & Pawson, 2012). The goal is to make explicit the assumptions (what the creator terms the ‘program theory’) about how an intervention is expected to work and then systematically gather the evidence to test and refine those assumptions.

The supposition behind the realist research process is that an intervention alters the context of a situation, which triggers mechanisms that elicit specific intended and unintended outcomes (Wong et al., 2012). The effect of interventions and their resulting mechanisms vary depending on the context. Thus the outcome of the intervention will also vary.

Pawson et al. (2005) identify four possible purposes for seeking the explanatory function of a realist review, which include: reviewing for ‘program’ theory integrity, reviewing to adjudicate between rival ‘program’ theories, reviewing the same theory in comparative settings, or reviewing official expectations against actual practice. The process for conducting a realist review involves five steps which align with the conventional Cochrane review formula, but with more sub-steps: defining the scope of the review; searching for evidence; appraising the evidence and extracting data; synthesizing the findings and drawing conclusions; and disseminating, implementing and evaluating (Pawson et al., 2005). This review process complements traditional methods of systematic reviews, which are more effective with simpler interventions such as clinical treatments (Pawson et al., 2005). As such, I believe that the realist review approach is ideal for the current study question.

Our purpose for conducting this review is to compare official expectations regarding the use of simulation in nursing education against actual practice reported in the literature. As explained above, the assumption or “theory” is that simulation helps educators to overcome challenges accessing sufficient numbers of clinical placements, to more fully prepare graduates to manage
complex patients, and to provide a consistent, rigorous, and risk-free learning environment for nursing trainees in comparison to clinical placements. We aimed to clarify which instructional design features nursing educators are using in their practice to attend to these assumptions, along with the associated evidence for the effectiveness of each design feature. This clarification is needed because many of the published nursing simulation reviews recommend further research in numerous domains without specificity or prioritization (Leigh, 2008; Hyland, & Hawkins, 2009; Kaakin, & Arwood, 2009; Cant, & Cooper, 2010; Harder, 2010; Weaver, 2011; Yuan et al., 2011; Yuan et al., 2012; Norman, 2012; Foronda et al., 2013; Dufrene & Young, 2014; Skrable & Fitzsimons, 2014; Paige & Morin, 2013; Jansson et al., 2013; Groom et al., 2014; Hallmark, Thomas, & Gant, 2014; Stroup, 2014; Shearer, 2013). When these recommendations are examined further, each lends itself well to the realist methodology because the authors identify specific questions – what works (e.g., low, medium or high-fidelity simulation, simulation instructional design characteristics), for whom (e.g., novice nurses, associate degree nurses, critical care nurses), under what circumstances (e.g., fundamental nursing skills courses, continuing education, clinical practice). In conducting this review, we endeavoured to answer these questions, and to synthesize the literature in a way that allowed us to produce specific recommendations for nursing educators on how to use simulation as an effective educational tool. As such, our review uses the realist synthesis methodology as the foundation for asking the question: “What simulation design features are most effective in enhancing nursing trainees' learning and clinical care outcomes?”

3.2 Methods

3.2.1 Philosophy Underpinning the Realist Review and Synthesis

Realist synthesis is a new method of conducting a theory-based systematic review of a body of evidence. Rather than for simple interventions, the realist review methodology is particularly beneficial when examining complex social programs or interventions such as the design, implementation, management and regulation of entire services (Pawson et al., 2004). A realist review is not designed to conclude whether a specific intervention is effective or not. It will not provide simple answers to complex questions. Instead, by examining the assumptions or ‘theory’ underlying why an intervention is implemented in a certain way, the goal is to provide a deeper understanding about an intervention and how to make it work most effectively (Pawson et al.,
2005). Realist reviews explore why an intervention works, and more specifically, why the same intervention may be successful in one situation but not in another. As such, the ‘theory’ is the unit of research, not the specific intervention; reviewers start with the ‘program theory’ and then systematically look to the literature to test and refine the theory (Wong, Westhorp, Pawson, & Greenhalgh, 2013a).

Realist research is based on the realist philosophy that suggests that our social reality cannot fully be measured and explained directly because it is processed through a variety of perspectives and lenses, but it can be explained indirectly by understanding the interactions between contexts, mechanisms and outcomes through the use of theories. More specifically, an intervention will alter a context, which will then stimulate mechanisms that will produce both anticipated and unforeseen outcomes (Wong et al., 2013b). The realist review methodology explains that a specific intervention will not always produce the same result; the outcome will depend on the context, which can vary. By exploring the context – mechanism – outcome (C-M-O) relationship, realist researchers propose theoretical explanations or assumptions about why specific interventions are successful or not (Wong, Greenhalgh, Westhorp, Buckingham & Pawson, 2013a). In contrast to formal theories, in realist research, a ‘program’ theory is an explanation or assumption about what a program or intervention is expected to do and how. Integral to the theory are the relational properties of the context, mechanism and outcome (Wong et al., 2013b).

The “context” in which an intervention occurs can include a variety of considerations including broad social or geographical features, resources or other features affecting the implementation of an intervention, the demographics or population profile of the participants, or the conditions within which the participants are functioning. Stated simply, the context is any condition that triggers and / or modifies the behaviour of a mechanism. The context is the “back-drop” of the research intervention (Wong et al., 2013b). Mechanisms are underlying entities, processes or structures that operate in particular contexts to generate outcomes of interest. Mechanisms are agents of change that generate outcomes (Wong et al., 2013b).

Pawson et al. (2005) identify four possible purposes for seeking the explanatory function of a realist review, which include: reviewing for ‘program’ theory integrity, reviewing to adjudicate between rival ‘program’ theories, reviewing the same theory in comparative settings, or reviewing official expectations against actual practice. This review process complements yet
contrasts traditional methods of systematic reviews, which are more effective for simpler interventions such as clinical treatments (Pawson et al., 2005).

The method for conducting a realist review broadly aligns with the method for conducting other types of syntheses, however a realist review is more heterogeneous and iterative, and less prescriptive (Pawson et al., 2004). When scoping a realist review, the question will involve a complex intervention or service. As stated above, there are four possible purposes for conducting a review: to test theory integrity, for theory adjudication, for comparison, or for reality testing. At this point, the reviewers will identify the review question, refine the purpose of the review, and find and articulate the ‘program’ theories. In contrast to a “formal” or “substantive” theory, a ‘program’ theory is the assumption of what the program or intervention is expected to do (Wong et al., 2013b). Once the theories have been identified, the reviewers then search for the evidence, in a systematic way, against the theories. The reviewers use purposive sampling, from diverse evidence sources and the “snowballing” search technique (pursuing references of references) until reaching saturation (Pawson et al., 2004). The next step is to appraise the primary studies and extract the data. Of note is that realist research rejects the hierarchical evaluation of evidence. Because, as explained above, a realist review is useful in examining complex social systems, the study questions do not lend themselves well to randomized controlled trials. Since the unit of research is the theory, not the intervention or individual study, the information gleaned from each study is assessed rather than the quality of the study design (Pawson et al., 2004). Wong et al. (2013a) explain that in a realist review, there may be different pieces of information that support the theory-building process in different ways. As such, deciding and selecting pieces of data from within the document must be based on the review question, rather than the overall study quality. In realist reviews, rather than a single form containing a standard list of questions, the reviewers may use different data extraction forms during the various phases and assimilate information through notes and annotation. However, while it is described linearly here, in practice, the two phases of searching and appraising the evidence and extracting and synthesizing the findings are iterative and fluid; the reviewers go back and forth depending on what they find (Pawson et al., 2004). Next, the reviewers synthesize the findings and draw conclusions. The final step is to make recommendations and disseminate findings, often directed at policy-makers and decision-makers. Rather than resulting in immediate changes, the goal of a realist review is to help decision-makers make sense of the evidence and contribute to a long-
term sustained shift in thinking (Pawson et al., 2004). As such, I believe that the realist review approach is ideal for the current study question.

3.2.2 The Present Realist Review and Synthesis

3.2.2.1 Purpose

Our purpose for conducting this review is to compare official expectations regarding the use of simulation in nursing education against actual practice reported in the literature. Simulation is being proposed as a replacement for real-life clinical placements as a strategy to overcome challenges accessing sufficient numbers of clinical placements, to more fully prepare graduates to manage complex patients, and to provide a consistent, rigorous, and risk-free learning environment for trainees. We aimed to clarify which instructional design features nursing educators are using in their practice to attend to these assumptions, along with the associated evidence for the effectiveness of each design feature.

3.2.2.2 Defining the Scope

As indicated above, there is a plethora of evidence regarding nursing simulation education and a number of reviews have also been published. However the gaps and opportunities justified conducting another review. But rather than simply re-synthesizing the studies, albeit with an attempt to resolve the limitations of the previous reviews, our goal was to synthesize the literature differently. I wanted to be able to provide nursing simulation educators with pragmatic recommendations about how they could use simulation as an effective teaching and learning tool.

To guide my decision-making, I used the Kastner et al. (2012) scoping review protocol that broadly and systematically identifies, and defines the various types of knowledge synthesis methods. Given my goal, I narrowed down the options to three methods: a narrative review, a mixed studies review or a realist review.

I decided against a narrative review because while it is useful for describing the history or the development of a specific issue, narrative reviews lack a systematic process and thus are viewed as problematic because they are inherently selective and emphasize some results over others, creating biased results (Green et al., 2006; Higgins, & Green, 2011). A narrative review would
also not provide insight into the specific questions that now need to be addressed in the simulation literature. I also decided against a mixed studies review because it would not provide the depth of understanding I wanted to achieve about the nursing simulation literature.

Instead, I chose to conduct a realist review because of its applicability to my study purpose and question. As explained above, realist reviews are effective when examining complex interventions or programs such as the use of simulation in nursing education. Given the volume of literature available about simulation in health professions education and specifically simulation in nursing education, and the plethora of ways and topics in which it is being used, my goal was to help educators make sense of the evidence and provide practical recommendations about how simulation is used in nursing education. Since a realist review synthesizes data using an explanatory rather than judgemental focus, such a method will support the achievement of this goal. As the literature review demonstrated, it has been established that simulation works. A realist review will now enable educators to understand what is it about simulation that works, for which students, and under what conditions, in order to help them achieve their course and program learning outcomes.

Once decided, we (NH and RB) endeavoured to gain an in-depth understanding of the realist review methodology so as to use it competently. This included reading publications explaining the process (Pawson et al., 2004; Pawson, Greenhalgh, Harvey, & Walshe, 2005; Wong et al., 2013a) as well as reading a variety of published realist reviews to see how other authors had implemented the approach (Wong, Greenhalgh, & Pawson, 2010; Wong, Greenhalgh, Westhorp, & Pawson, 2012; Hewitt, Sims, & Harris, 2014; Hatala et al., 2014). This process itself was iterative; we revised and refined our course of action during the review as we continued to learn about the methodology. I then solicited two additional reviewers to balance the knowledge and expertise of the team - to ensure there were additional members with research expertise (Dr. Joan Samuels-Dennis, Professor) and members with nursing simulation education expertise (Sandra Cop, Professor and Clinical Simulation Program Coordinator). I explained the goal of the review, the proposed review methodology and process, and they both willingly joined the review team. While being informed by the realist review process, our team made subjective decisions about what data constituted context, versus mechanism, versus outcome, we set restrictions on which studies to include based on the context – mechanism - outcome (C-M-O) criteria, and we agreed on how to conduct our review and synthesize the findings.
3.2.2.3 Searching for and Appraising the Evidence

The studies used in this review are a subset of a comprehensive literature search and review that has been previously reported (Cook et al., 2011). With the assistance of an experienced research librarian, a comprehensive search was undertaken with broad inclusion criteria focused on the modality (i.e., simulation, part-task trainers), health care professionals (e.g., nursing, medicine), and education or training. We used a broad definition for simulation-based technologies: “Simulation technologies encompass diverse products including computer-based virtual reality simulators, high-fidelity and static mannequins, plastic models, live animals, inert animal products, and human cadavers.” The search did not include a beginning date, did include all languages, and the last date of the search was May 11, 2011. We screened all titles and abstracts independently and in duplicate to assess study inclusion. In the event of disagreement or if information was lacking in the abstract, we reviewed the full text independently and in duplicate, and conflicts were resolved by consensus. The review was conducted in accordance with PRISMA standards (Moher et al., 2009). No human subjects were involved and thus ethics board approval was not required.

For the realist review process, we followed the RAMESES publication standards for realist syntheses (Wong et al., 2013a). Eligible studies consisted of those that were included via the methods described above, in which > 25% of the participants were nurses or nursing students. From these, NH and RB excluded studies comparing simulation to no intervention (Cook et al., 2011), as well as those comparing simulation to another, non-simulation intervention (Cook et al., 2012) because both types of comparison do not permit analysis of the realist review questions related to the context, mechanism, and outcomes associated with nursing simulation interventions.

3.2.2.4 Extracting and Synthesizing the Findings

The analysis and synthesis process was iterative. Wong et al. (2013a) explain that in a realist review, there may be different pieces of information that support the theory-building process in
different ways. As such, deciding and selecting pieces of data from within the document must be based on the review question, rather than the overall study quality. After applying the inclusion and exclusion criteria, 42 studies comparing a simulation intervention to another simulation intervention were analyzed independently and in duplicate where data about the context, mechanism and outcome of the study were extracted. Our team made judgment calls in extracting the data that corresponded to many of our queries, especially given that most authors provided equivocal information. We met often to discuss and clarify uncertainties to ensure consistency in the analysis and synthesis processes. After context – mechanism - outcome (C-M-O) data had been extracted for each article individually and in duplicate, we reviewed the data gathered for all articles and excluded the studies with insufficient or absent data. Once we completed our exclusion process, the included articles represented those in which authors provided data regarding context, mechanism, and outcome, thus allowing NH and RB to conduct a thematic analysis. In contrast to the typical multi-stage literature search process explained by Pawson et al., (2005) of starting with the ‘program’ theory and then searching the evidence, given the comprehensive search that had already been conducted by Cook et al. (2011), NH and RB analyzed the existing studies and allowed the ‘intervention theories’ to emerge inductively. Based on our purpose of the realist review, which was to identify the instructional design features of nursing simulation training used in actual practice, our iterative method allowed the themes to emerge inductively during our synthesis process.

As indicated above, the analysis and synthesis process conducted by NH and RB was iterative. Initially, our team coded the studies, independently and in duplicate, for a variety of variables including the credential of nursing faculty, the reported program of the nursing students, and the clinical specialty area. After documenting that information, our team extracted additional data regarding the context, mechanisms, and outcomes for each study. During the first iteration, we extracted the following information (see Appendix 1):

- For context, we extracted data on who was there (learner and instructor), what was there (the task and simulator), where they were (setting) and when (formal education or professional training).
- For mechanism, we extracted data on how trainees were learning the task.
For outcome, we extracted data on how well the trainees learned the task.

Following this first iteration, the team realized that our analysis had been rather superficial and used our experience to refine the extraction form and recode the studies as follows (see Appendix 2):

- For context, we extracted data on who was there (RNs, RPNs, undergrads, post-grads, etc.), what was there (equipment, supports, instructors, etc.), how the learning setting was organized, and the location of the learning setting (college, university, hospital, lab, classroom, etc.).

- For mechanism, we extracted data on which learning principles were indicated to have influenced the instructional design, what everyone in the room was doing (i.e., how the activity was experienced by everyone), and what interactions were occurring (i.e., between people, between equipment, between all components).

- For outcomes, we extracted data on whether the expected outcomes were achieved (i.e., whether results aligned with hypotheses), the explicit (author described) outcomes, the implicit or unintended consequences of the activity, how the setting impacted the outcomes, and what (if any) differences in outcomes were based on setting.

After examining the second round of extracted data, our team collectively decided that many authors did not report sufficient information regarding the conceptual framework that guided their study. Furthermore, for the studies that did refer to a framework, few connected the framework to the measured outcomes. As such, we chose to modify the data extraction tool once more to extract the following information (see Appendix 3):

- For context, we extracted data on what was there (equipment, supports, instructors, etc.), how the learning setting was organized, the location of the learning setting (college, university, hospital; lab, classroom, etc.), and which simulator(s) were used (e.g., mannequin, task trainer, etc.).

- For mechanism, the questions were refined to examine which conceptual framework or theory was noted to have influenced the instructional design, how the conceptual framework was utilized (i.e., in study design, data interpretation), and what interactions
were occurring: between participants, between investigators or research assistants, between participants & equipment, and between all components.

- For outcomes, we extracted information on how well the outcome matched the conceptual framework (were the authors testing the theory directly), how the outcomes were measured (e.g., by whom? when?), how the setting impacted the outcomes, and what (if any) differences in outcomes were based on setting.

Given that conducting a realist review was new for three of the four reviewers, and because we had not done this type of work together before, there was a lot of time invested in ensuring we all understood and agreed about the process, ensuring consistency and accuracy in the methodology, and clarifying when anyone had questions. In the beginning, much time was spent gaining a shared definition of context, mechanism and outcome. We consulted the RAMESES standards (Wong et al., 2013a) frequently to understand the process, to clarify questions, and to ensure we followed the methodology. We also reviewed other published realist syntheses to gain further understanding of the process and see how other authors had conducted such reviews (Pawson et al., 2005; Wong et al., 2010; Wong et al., 2012; Hewitt et al., 2014; Hatala et al., 2014). With each iteration of data extraction, we practiced coding articles together, then coded additional articles independently, and compared results. We met regularly to clarify and resolve disagreement.

The data extracted during the third time through the studies are the results that led to the identified themes. The themes emerged based on the foci of the primary studies. That is, when examined individually and together, all the studies fit into clusters or groups based on the explicit or implicit assumptions of the researchers.

### 3.3 Results

Of the 10903 studies identified in the original systematic review (Cook et al., 2011), we found 147 studies that focused on nursing and/or nursing students using any experimental study design. After applying our inclusion criteria, we identified 42 studies for analysis using our context – mechanism - outcome (C-M-O) judgments. Figure 1 summarizes the numbers of studies included and excluded at each stage of our review process.
Figure 1:

Search by Cook et al. (2011) with the support of an experienced research librarian in databases: MEDLINE, EMBASE, CINAHL, PsychINFO, ERIC, Web of Science, and Scopus

10297 (search) + 606 (review of reference lists and journal indexes) = 10903

↓

Excluded 10756 (nursing trainees < 25% of participants) = 147

↓

Excluded 84 (SIM versus no intervention) = 63

↓

Excluded 21 (SIM versus other non-simulation intervention) = 42

↓

Total included in this review = 42 SIM versus SIM
3.3.1 Study characteristics

Table 1 summarizes the key features of all 42 studies we included that met our criteria of comparing two different forms of technology-enhanced simulation, and where > 25% of the participants were nursing trainees. All of the 42 studies were published in English.

More than half (n=2329; 54%) of the trainees were nursing students (in diploma, degree and post-graduate programs). Other trainees included: medical students, practicing nurses, physicians in post-graduate training, physicians in practice, pre-hospital providers, and “other” (ambiguous / mixed populations). The major clinical area of training (either in practice or in academic settings) under study was the broadly coded “fundamental nursing skills” (n=14) (e.g., vital signs, nasogastric and urinary catheter insertion); other clinical areas of training in three or fewer studies included paediatrics, surgery, anesthesia, critical care, internal medicine / subspecialty, and emergency medicine. A total of 15 studies did not provide enough detail for us to code a specific clinical area of training.

The majority of outcomes being assessed were subjective (n=3636; 84%). The most frequent outcomes measured were process skills (e.g., advanced cardiac life support skills on a simulator; 29 studies, n=3204; 50%), and the second most frequent were reaction / self-reported satisfaction (17 studies, n=1425; 22%). Other outcomes measured in the remaining studies included knowledge (e.g., MCQ exam scores; 9 studies), product skills (e.g., applying a wound dressing appropriately; 7 studies), time skills (e.g., time for nasogastric tube insertion; 5 studies), provider behaviours (e.g., rating of venous catheter insertion on real patients; 2 studies), and patient effects (e.g., rate of procedural success on patients; 2 studies).
Table 1 summarizes key study features.

<table>
<thead>
<tr>
<th>Study Characteristic</th>
<th>Category</th>
<th>No. of studies (No. of participants*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants**</td>
<td>Nursing Students</td>
<td>24 (2329)</td>
</tr>
<tr>
<td></td>
<td>Nurses</td>
<td>14 (993)</td>
</tr>
<tr>
<td></td>
<td>Medical students</td>
<td>4 (191)</td>
</tr>
<tr>
<td></td>
<td>Physicians in post-graduate training</td>
<td>3 (79)</td>
</tr>
<tr>
<td></td>
<td>Physicians in practice</td>
<td>4 (231)</td>
</tr>
<tr>
<td></td>
<td>Pre-hospital providers</td>
<td>2 (27)</td>
</tr>
<tr>
<td></td>
<td>Other/ambiguous/mixed populations</td>
<td>9 (464)</td>
</tr>
<tr>
<td>Nurses degree designation**</td>
<td>Degree (i.e., BN, RN)</td>
<td>2 (84)</td>
</tr>
<tr>
<td></td>
<td>Other/ambiguous/mixed</td>
<td>3 (246)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>9 (663)</td>
</tr>
<tr>
<td>Nursing students’ degree program**</td>
<td>Degree (i.e., BN, RN)</td>
<td>10 (821)</td>
</tr>
<tr>
<td></td>
<td>Diploma (i.e., RPN, LPN)</td>
<td>2 (65)</td>
</tr>
<tr>
<td></td>
<td>Graduate (e.g., masters, NP, CNP, CNS, nurse anesthetist)</td>
<td>5 (92)</td>
</tr>
<tr>
<td></td>
<td>Other/ambiguous/mixed</td>
<td>7 (781)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>3 (757)</td>
</tr>
<tr>
<td>Area of clinical training**</td>
<td>Fundamental nursing skills (e.g., vital signs, nasogastric and urinary catheter insertion)</td>
<td>14 (2032)</td>
</tr>
<tr>
<td></td>
<td>Pediatrics</td>
<td>3 (137)</td>
</tr>
<tr>
<td></td>
<td>Surgery</td>
<td>2 (154)</td>
</tr>
<tr>
<td></td>
<td>Anesthesia</td>
<td>2 (127)</td>
</tr>
<tr>
<td></td>
<td>Critical Care</td>
<td>1 (120)</td>
</tr>
<tr>
<td></td>
<td>Internal medicine / subspecialty</td>
<td>1 (30)</td>
</tr>
<tr>
<td></td>
<td>Emergency medicine</td>
<td>1 (49)</td>
</tr>
<tr>
<td></td>
<td>Other/ambiguous/mixed</td>
<td>15 (1456)</td>
</tr>
<tr>
<td></td>
<td>Not reported</td>
<td>3 (209)</td>
</tr>
<tr>
<td>Outcome assessment</td>
<td>Subjective</td>
<td>34 (3636)</td>
</tr>
<tr>
<td></td>
<td>Objective</td>
<td>8 (678)</td>
</tr>
<tr>
<td>Outcomes measured**</td>
<td>Reaction/Self-report satisfaction</td>
<td>17 (1425)</td>
</tr>
<tr>
<td></td>
<td>Knowledge</td>
<td>9 (569)</td>
</tr>
<tr>
<td></td>
<td>Skill: time</td>
<td>5 (274)</td>
</tr>
<tr>
<td></td>
<td>Skill: process</td>
<td>29 (3204)</td>
</tr>
<tr>
<td></td>
<td>Skill: product</td>
<td>7 (609)</td>
</tr>
<tr>
<td></td>
<td>Provide behaviour</td>
<td>2 (161)</td>
</tr>
<tr>
<td></td>
<td>Patient effects</td>
<td>2 (148)</td>
</tr>
</tbody>
</table>

*Numbers reflect the number enrolled.

**The number of studies and trainees in some subgroups (summing across rows or columns) may sum to more than the number for all studies because several studies included >1 comparison arm, >1 trainee group, fit within >1 clinical topic, or reported multiple outcomes.
3.3.2 Overview of Realist Synthesis Findings

Of the 42 studies, we (NH, RB, JSD, and SC) excluded 29 from the synthesis due to insufficient information regarding either the context, mechanism or outcome (see Appendix 4 for list of references). Between context, mechanism, and outcome, while context was the most well reported, some common data points missing from the context description included: the type of nurse trainee (e.g., RNs versus RPNs, undergrads versus post-grads), the setting in which the simulation occurred (e.g., hospital versus academic simulation lab versus off-site simulation centre), the content or topic being taught (e.g., procedural skills versus critical thinking), and/or the type of simulation equipment being used (e.g., authors report using a ‘mannequin’ without describing the company and model).

For mechanism, missing information included: a description of the educational process utilized (e.g., students learned IV skills on different simulators but there was no description on how or what they were taught), which educational mechanism guided the intervention (e.g., feedback versus none, self-regulated learning versus instructor-regulated learning) and a hypothesis about how the intervention would affect the outcome (e.g., how and why would a high-fidelity mannequin affect the measured outcomes in comparison to a low-fidelity mannequin?).

For outcomes, we excluded articles due to some common gaps including: a lack of alignment between the study question and the assessment tool used (e.g., evaluating the effectiveness of an intervention based on performance-based hypotheses, yet measuring participants’ enjoyment and confidence), issues with confounding variables (e.g., intervention group trained on same mannequin used for testing but control group did not) and inconsistent assessment methods between control and intervention groups (e.g., groups assessed on different equipment).

Over several meetings, NH and RB discussed and analyzed the 13 studies that we did include in our realist review, a process that resulted in four broad themes related to instructional design: feedback, debriefing, self-regulated learning (SRL), and the intersection between feedback and SRL. Thirteen studies are reported below with one study (Dine et al., 2008) being reported under two themes - both feedback and debriefing.
3.3.3 Realist Theme 1: Feedback

McGaghie et al. (2010) report feedback to be the most important and frequently cited instructional design feature for effective simulation-based medical education. Feedback offers trainees information about how they are performing, so they can progressively improve (Hatala et al., 2014). Feedback as an instructional design feature has been studied extensively, and Hatala et al. (2014) recently conducted a meta-analysis of simulation-based procedural skills training studies, demonstrating that feedback has moderate to large benefits for training outcomes. Two key findings of that review are that terminal feedback (i.e., provided after the simulation scenario has ended) is likely better than concurrent feedback (i.e., provided during the simulation scenario) for long-term retention in novice learners, and that multiple sources of feedback (e.g., from an instructor and the simulator) is likely better than a single source of feedback.

We included four studies on feedback in the present review (Dine et al., 2008; Grady et al., 2008; Rodgers, Securro, & Pauley, 2009; Domuracki, Moule, Owen, Kostandoff, & Plummer, 2009). We found heterogeneity in the study contexts, with two studies focused on nursing students, in academic simulation centres, using various simulators to learn skills such as nasogastric tube and urinary catheter insertion (Grady et al., 2008), and advanced cardiovascular life support (Rodgers et al., 2009). Two other studies focused on practicing nurses in academic hospitals learning cardiopulmonary resuscitation skills using different mannequins (Dine et al., 2008), and on a combination of medical students, nursing students, and practicing nurses in an academic hospital learning to apply cricoid pressure on the CPT-Compact cricoid pressure simulator (Domuracki et al., 2009).

All four studies compared no feedback to various versions of more or less feedback, suggesting that the researchers expected more feedback to result in better skill performance. Feedback sources were high-fidelity mannequins that were fully responsive including the ability to gag, cough, respond to “yes” or “no” questions, and say “ouch” (Grady et al., 2008), a cricoid pressure simulator that showed the force being applied (Domuracki et al., 2009), a CPR-sensing defibrillator that reported CPR quality (Dine et al., 2008), and SimMan, a high-fidelity mannequin that had pulses, chest movement and could respond to questions (Rodgers et al., 2009). The studies demonstrated that some feedback is better than none, resulting in improved accuracy of skill performance for depth of CPR compressions and amount of cricoid pressure.
exerted (Dine et al., 2008; Domuracki et al., 2009). The studies also showed that more feedback (linked by authors to ‘high-fidelity’ mannequins) is better than less feedback (linked by authors to ‘low-fidelity’ mannequins), resulting in more positive student attitudes and higher performance of nasogastric tube and urinary catheter insertion skills (Grady et al., 2008), and knowledge improvement and skill performance for Advanced Cardiovascular Life Support skills (Rodgers et al., 2009).

Given the heterogeneity in the contexts and types of interventions in which feedback has been studied in the nursing simulation literature, we are limited in our efforts to directly answer the inquiries associated with realist syntheses. When we relate our results to the findings of Hatala et al.’s (2014), there is consistency in the message that more feedback is usually superior to less feedback. There are more sophisticated theories associated with feedback and learning (see Hatala et al. 2014 for examples), however, and thus the nursing simulation literature has much to learn about how best to harness this instructional design feature.

### 3.3.4 Realist Theme 2: Debriefing

Debriefing has been defined as “…a discussion between two or more individuals in which aspects of performance are explored and analyzed with the aim of gaining insights that impact the quality of future clinical practice.” (Cheng et al., 2014). Numerous approaches and conceptual frameworks have been proposed for debriefing across the healthcare simulation literature, yet we included just two studies on this topic (Dine et al., 2008; Grant, Moss, Epps, & Watts, 2010). For those two studies, the contexts were largely inconsistent, with participants including nursing and nurse anaesthesia students learning how to provide care for patients with pulmonary and cardiac conditions (Grant et al., 2010) and practicing nurses learning to provide cardiopulmonary resuscitation (Dine et al., 2008). Training took place in either a university simulation lab using an unspecified high-fidelity simulator (Grant et al., 2010) or an academic medical centre using Laerdal’s Resusci-Anne and the MRx-QCPR CPR-sensing monitor and defibrillator (Dine et al., 2008).

While the studies evaluated different mechanisms of debriefing, both focused on comparing combinations of methods to ‘one-shot’ approaches. In one, the authors compared debriefing combined with feedback to either feedback alone or instructor-led debriefing alone (Dine et al.,
2008). The other compared instructor-led debriefing plus the opportunity to view a videotape of the preceding simulation scenario to instructor-led debriefing alone. Hence, both studies appeared to think that debriefing would be most beneficial when it incorporates multiple sources of information, rather than a singular source. Such a ‘theory’ of the mechanisms of debriefing has been studied in the broader healthcare simulation literature (Cheng et al., 2014). Using outcomes related to knowledge (Dine et al., 2008) and both technical and non-technical skills (Dine et al. 2008, Grant et al. 2010) that were measured at the end of the practice session (i.e., on an immediate post-test), both studies demonstrated superiority of the combined approaches over the singular approaches.

Having just two studies does not allow us to answer the realist questions of “what works, for whom, under what circumstances?” effectively. We can relate our results, however, to the broader literature on debriefing in simulation, where we find agreement and disagreement. That is, the nursing literature aligns with findings that combining feedback (from an expert) with debriefing is a superior mechanism, whereas the benefit of adding video to the debriefing process has been challenged (see Cheng et al., 2014; Lorello et al., 2014). The low number of studies suggests much is to be learned about how to optimize the context – mechanism - outcome (C-M-O) linkages when designing debriefing for the various learner populations and content areas available in the nursing literature. Moreover, neither author group provided explicit details about how they designed or implemented the debriefing, suggesting these studies could not be replicated, which is a major limitation of these works.

3.3.5 Realist Theme 3: Self-Regulated Learning (SRL)

Self-regulated learning (SRL) has been defined as “…learning by an individual who, with or without supervision, actively modulates affect, cognition and behaviour using supports or scaffolds designed to facilitate the achievement of desired learning goals.” (Brydges et al., 2015). While there are several theoretical frameworks describing the underlying skills associated with SRL (e.g., goal-setting, help-seeking, perseverance, etc.), the healthcare simulation literature has mostly focused on SRL as learning that happens when an instructor is not providing direct supervision (Brydges et al., 2015).
We included four studies that emphasized the ‘self’ component of SRL (LeFlore, & Anderson, 2009; LeFlore, Anderson, Michael, Engle, & Anderson, 2007; Swanson, et al., 2010; Brydges, Carnahan, Rose, & Dubrowski, 2010). In terms of participant population, the study contexts consistently focused on student learning including interprofessional students (LeFlore, & Anderson, 2009), nurse practitioner students (LeFlore et al., 2007), and nursing students (Swanson et al., 2010; Brydges et al., 2010). All training took place in academic simulation centres (LeFlore, & Anderson, 2009; LeFlore et al., 2007; Swanson, et al., 2010; Brydges et al., 2010). Participants practiced skills including pediatric respiratory distress (LeFlore, & Anderson, 2009; LeFlore et al., 2007), cardiovascular care (Swanson et al., 2010), and intravenous catheterization (Brydges et al., 2010). Simulators used in the training included an unspecified simulator (Swanson et al., 2010), Laerdal SimBaby (LeFlore, & Anderson, 2009; LeFlore et al., 2007), and an inanimate plastic arm, Laerdal Virtual IV and Laerdal SimMan (Brydges et al., 2010).

We coded a consistent mechanism in three of the four studies, where the authors expected that interventions in which the instructor modelled the desired skill would be superior to interventions in which the learners practiced entirely independently (LeFlore, & Anderson, 2009; LeFlore et al., 2007; Swanson et al., 2010). In those studies, the instructor-led groups resulted in moderate to large benefits for skills related to the processes of caring for the simulated patient. Notably, all authors used immediate post-tests, and none measured longer-term skill retention. In the fourth study, the authors compared different types of SRL conditions and found that trainees benefit when they are given self-control over how they sequence their learning, and moreover, that if trainees are empowered to set their own performance standards, they attain a similar level of technical skill as when external standards are set by experts for the trainees (Brydges et al., 2010).

Three of the studies reported using Bandura’s social learning theory or Zimmerman’s related social-cognitive theory (Brydges et al., 2010; LeFlore, & Anderson, 2009; LeFlore et al., 2007), and the fourth study referenced social engagement theory (Swanson et al., 2010). While most of the studies suggested that instructor modelling is superior to allowing trainees to learn entirely independently, none of those three studies used a comparison group that included a more active self-regulated learner (as was done in Brydges et al., 2010). Given the consistency in many contextual variables, including the participant population, study in the simulation centre setting,
and the types of skills measured (often technical or behavioural process skills), the data support a cautious recommendation that students who are entirely novice should not be left to learn on their own without appropriate educational supports built into the learning environment (Brydges & Butler, 2012).

3.3.6 Realist Theme 4: Intersection Between Feedback and SRL

An emerging area of inquiry in health professions education research is on the intersection between how instructors provide feedback and what trainees do with the information to enhance their learning. One series of studies, for example, has emphasized that trainees filter feedback by comparing it to their impressions of themselves, and the alignment between their self-judgments and the instructor’s feedback has a strong influence on whether and how they use that feedback (Eva et al., 2012). Hence, there is a clear link between how trainees self-regulate their learning and the ways that feedback can influence that process.

We found four studies that investigated the intersection of SRL and feedback (de Vries, Schelvis, Rustemeijer, & Bierens, 2008; Kardong-Edgren, Oermann, Odom-Maryon, & Ha, 2010; Oermann et al., 2010; Orde, Celenza, & Pinder, 2010). Three studies included nursing students (Kardong-Edgren et al., 2010; Oermann et al., 2010) or a combination of medical students, nursing students and post-graduate nursing students (Orde et al., 2010), whereas the fourth study focused on practicing nurses (de Vries et al., 2008). Two studies took place in the hospital setting (de Vries et al., 2008; Orde et al., 2010) and two took place in academic simulation centres (Kardong-Edgren et al., 2010; Oermann et al., 2010). Participants learned how to use automated external defibrillators (de Vries et al., 2008), to perform CPR (Kardong-Edgren et al., 2010; Oermann et al., 2010), and to conduct laryngeal mask insertion (Orde et al., 2010). Training utilized an unspecified training AED, mannequin and instructional poster (de Vries et al., 2008), and Laerdal Resusci-Anne Skillmeter (Orde, Celenza, & Pinder, 2010) or Resusci-Anne with voice assist (Kardong-Edgren et al., 2010; Oermann et al., 2010).

The educational mechanisms proposed for how feedback influences trainees’ SRL focused mainly on the source of feedback. Three studies compared participants who received feedback from instructors in traditional course formats to those who received feedback as they self-regulated their practice from either voice-assisted mannequin simulators (Kardong-Edgren et al.,
2010; Oermann et al., 2010), or an AED trainer, a mannequin and a training poster (de Vries et al., 2008). The two studies using the voice-assisted mannequin simulators found that the immediate, concurrent feedback they provided resulted in more accurate CPR performance than the ‘sporadic’ feedback given by an instructor. The third study found self-training to be as effective as the traditional instructor-led training. In the fourth study, the authors compared a ‘two-step’ training model (i.e., see one, do one), with a more sophisticated ‘four-step’ training model (i.e., instructor demonstration and deconstruction, followed by trainee formulation and performance), and found no statistically significant difference between the two approaches.

These studies add to the body of evidence relating feedback and SRL by suggesting that the source of information (i.e., instructor vs. simulator) has an influence on how trainees utilize and learn from that information. As this is an expanding area of the literature, additional aspects of feedback will need to be considered including feedback timing and frequency (Hatala et al., 2014), as well as the perceived credibility of the feedback source (Eva et al., 2012).

### 3.4 Discussion

#### 3.4.1 Summary of Findings

We identified 42 studies comparing at least two different interventions involving nursing simulation. The majority of those (54%) used nursing students as participants and measured performance using mostly subjective outcomes (84%). While more than half of the outcomes were measures of process skills (69%), many studies relied on self-reported satisfaction as the primary outcome (40%). Upon applying our inclusion criteria and completing the realist synthesis, we excluded 69% of the nursing simulation studies due to missing information. Having to exclude the majority of studies was disappointing because we expected to have a large dataset of included studies to synthesize and derive a list of pragmatic recommendations for nursing simulation educator’s instructional design decisions.

Given the small number of included studies (N=13), our synthesis findings are far from conclusive. Our preliminary interpretations suggest that learners benefit from more rather than less feedback, and that the source of feedback (i.e., instructor vs. simulator) matters and depends on the learning objectives and skill of interest. The synthesis also suggests educational benefits
when pairing multiple instructional design features (i.e., feedback and debriefing) and multiple feedback sources (i.e., simulator and instructor). When allowing trainees to engage in self-regulated learning, it seems important to have an instructor present for some aspect of performance to provide support for trainees.

3.4.2 Limitations

The limitations of this review relate directly to the limitations of the primary studies. Many of the existing nursing simulation studies lack details regarding the context, mechanism and outcomes of their studies. While we excluded 69% of the studies due to these limitations, the remaining studies were heterogenous, which limits our interpretations and results in cautious recommendations. Further, we cannot comment on the longer-term effects of nursing simulation-based interventions because delayed outcome assessments are rare in this literature. That absence is despite many educational literatures having agreed that it is expected practice to include follow-up and delayed testing in order to investigate the longer-term effects of interventions (e.g., see Schmidt & Bjork, 1992).

Additionally, we used a modified version of the realist review process outlined in the RAMESES standards. We made subjective decisions about what data constituted context – mechanism - outcome (C-M-O) and which studies to include based on those criteria. Our subjectivity extended to our decisions for how to synthesize the findings. While we believe we applied the realist review standards well, the nature of realist reviews is that other reviewers likely would have followed a different process and made different decisions.

3.4.3 Integration with Prior Work

Hayden et al. (2014) concluded that there is “substantial evidence” that comparable outcomes can be achieved when up to 50% of traditional clinical hours are replaced by high-quality simulation. Through our realist review, we are unable to confirm that conclusion, specifically because we did not synthesize studies comparing simulation and clinical training interventions. Our results do, however, cause us to pause and question whether nursing simulation-based training interventions have been refined sufficiently to ensure the highest quality of education. Indeed, we suggest much work is to be done to optimize simulation-based training for nurses and nursing trainees.
With regards to the previous nursing simulation reviews, our realist review complements some of their conclusions. Harder (2010) and Kardong-Edgren et al. (2010) indicate that interprofessional simulation research based on scenarios that replicate the true clinical environment are required, while our findings confirm that the majority of studies used nursing students as participants (54%). Additionally, a number of researchers (Kaakinen, & Arwood, 2009; Rourke, Schmidt & Garga, 2010; Waldner & Olson, 2007; Franklin et al., 2015) recommend using theories from teaching, learning, and nursing to design and interpret findings in research studies, and we found a pattern whereby those studies that did not use a theoretical framework were often excluded for not outlining the study’s context – mechanism - outcome (C-M-O) data.

3.4.4 Implications for Practice and Future Directions

As the nursing simulation literature grows, it is essential that studies be replicable, both for researchers wishing to expand upon evidence and educators wishing to incorporate evidence into practice. We used the realist review process as a guide and, unfortunately, this led to our excluding 69% of relevant studies. Authors missed reporting a broad range of information about the context – mechanism - outcome (C-M-O) connections that is required in this type of review. Knowing that research requires a significant investment of financial, human and other resources, we suggest that the amount of missing information in the original studies limits the generalizability of findings and therefore calls those investments into question. As the use of the realist review methodology increases, we suggest that this has implications for the reporting practices of authors, and the expectations for reporting of peer reviewers, and journal editors.

Educators must be selective and purposeful when using simulation because instructional design is more important than the mere presence or absence of simulation. The instructional design features that have garnered early interest in the nursing simulation literature are feedback, debriefing, self-regulated learning, and the intersection between feedback and SRL. We recommend that nursing education researchers consider the broad range of design features for which robust evidence has accumulated (e.g., see Cook et al., 2013).

The lack of consistency in the literature suggests that nursing simulation researchers need a common definition or understanding of simulation as well as the concept of fidelity. Researchers and educators must separate the simulator equipment from the instructional design and
understand the unique contributions of all kinds of simulation and consider using the terms physical resemblance, and functional task alignment (Hamstra et al., 2014).

To move simulation research forward, there is a need to compare one simulation intervention to another simulation intervention, with manipulations of instructional design features being the key focus. Of critical importance, however, is the full reporting of the contexts, mechanisms and outcomes because the realist review method is only as good as the primary research on which it is based.

Finally, we recommend that future studies be underpinned by a theoretical framework so that they are methodologically rigorous and grounded in sound instructional design principles, with each context – mechanism - outcome (C-M-O) component being deliberately contemplated and reported. Only then will the value of the research be realized with pragmatic recommendations for simulation educators, along with theory-building or theory-refuting data for researchers.
4 General Discussion

4.1 Summary of Findings

We conducted this review recognizing that a more focused and systematic synthesis of the literature was necessary after we identified repeated limitations and associated opportunities in the existing nursing simulation reviews. Driven by a desire for our findings to be useable in the daily practice of nursing simulation educators, we used an innovative, novel approach to synthesis – the realist review methodology - that produced a different type of appraisal of the nursing simulation-based training literature. We believe a novel approach is needed because reviews of the broader healthcare simulation literature have shown the value of simulation, yet those reviews do not specifically guide nursing educators on how and when to use simulation optimally. Therefore, we endeavoured to answer the question “What simulation design features are most effective in enhancing nursing trainees' learning and clinical outcomes?” and to synthesize the literature to produce specific recommendations on the use of simulation as an instructional tool for nursing educators.

4.1.1 Primary Findings – Poor Study Reporting and High Exclusion Rate

Our initial search detected 42 studies comparing at least two different interventions involving nursing simulation. Of those, the majority used nursing students as participants (54%) and measured performance using mostly subjective outcomes (84%). While more than half of the outcomes were measures of process skills, many studies relied on self-reported satisfaction as the main outcome.

Upon applying our inclusion criteria and completing the realist synthesis, our primary findings were surprising and disheartening for two reasons. First, given the volume of studies published, we expected to have a large dataset of included studies to synthesize. Second, we expected to be able to make pragmatic recommendations to nursing simulation educators to aid in making instructional design decisions. However, as demonstrated in the results above, 69% of the nursing simulation studies had to be excluded from this synthesis due to missing information. Therefore the volume of data to analyze was limited and the resulting recommendations are
similarly limited. The nature of the missing information ranges broadly within the context – mechanism - outcome (C-M-O) categories. Data missing from the context description included: the type of nurse trainee, the simulation setting, and the content or topic being taught. For mechanism, missing information included: descriptions of the educational process utilized, which educational mechanism guided the intervention, and a hypothesis about how the intervention would affect the outcome. Finally, common gaps in outcomes included: lack of alignment between the study question and the assessment tool used, issues with confounding variables, and inconsistent assessment methods between control and intervention groups.

Knowing that research requires a significant investment of financial, human and other resources, we suggest that the amount of missing information in the original studies limits the generalizability of findings and therefore calls those investments into question.

We used rigorous inclusion criteria that led us to exclude the majority of published studies, which demonstrates a quality issue in the current literature and limits the recommendations that can be made for best practices in simulation-based training. Indeed, our inclusion of only 31% of studies limited our capacity to pool the evidence in a way that reduced bias and confounding (Higgins, & Green, 2011). Thus, as reviewers, we could not fully answer our question and are challenged to make specific recommendations that move the literature and practice forward. As the realist synthesis methodology continues to increase in use, the design, implementation, and reporting of research needs vast improvements in the nursing simulation literature if the findings are to be included in realist reviews.

In alignment with previous reviews of the nursing simulation literature, we found that many research studies were not grounded in a theoretical framework. While the absence of a theoretical framework did not directly result in our excluding a study, we argue that there is a connection. We observed that the lack of research, educational or nursing theories led to methodological limitations which were often associated with the C-M-O gaps identified above; the researchers either did not report all the relevant data, or did not consider them.

We also found that authors do not use a consistent, operational definition of simulation or of the term fidelity (high, medium and low). What educators and researchers consider to be “simulation” varies broadly, and there is an implied assumption amongst authors that the terms “high, medium, and low-fidelity” are self-evident. We argue that “fidelity” has become a “God
term” (Lingard, 2009). That is, nursing researchers appear to be presuming a shared meaning and understanding of what high, medium and low-fidelity actually mean, which leads to a superficial treatment of the terminology that precludes deeper investigation. We suggest that this is problematic because such flippant use of terminology appears to be associated with an absence of important methodological details being reported, specifically, how the investigators implement an educational intervention and their rationale for doing so. Hence we are particularly concerned by nursing researchers’ near singular focus on “high-fidelity simulation” (Leigh, 2008; Harder, 2010; Hyland, & Hawkins, 2009; Foronda et al., 2013; Dufrene & Young, 2014; Skrable & Fitzsimons, 2014; Jansson et al., 2013; Norman, 2012; Rourke et al., 2010; Shearer, 2013; Weaver, 2011; Yuan et al., 2011; Yuan et al., 2012). This focus on high-fidelity appears to have become ubiquitous without consideration, study or analysis of the value and limitations of all types of simulation. There may be an appropriate context (time, place, learner and lesson) for all types of simulators that is likely being missed.

4.1.2 Secondary Findings – Few Studies to Synthesize using Realist Methods

We identified 13 studies that included sufficient information regarding the study contexts, mechanisms and outcomes. As we analyzed these studies, we found that we could synthesize them across instructional design themes - feedback, debriefing, SRL, and the intersection between feedback and SRL – and thus, could show some preliminary patterns of how nursing researchers are studying them. These patterns are explained below for nursing simulation educators’ consideration. In comparison to the 11 instructional design features identified by Cook et al. (2013), the 13 nursing simulation studies that could be included addressed two (feedback and individualized learning).

As a whole, our findings of the context – mechanism - outcome (C-M-O) data were noteworthy. Regarding context, because of our inclusion criteria, there was homogeneity of the participants (primarily practicing nurses and nursing students) in 13 studies. However, there was heterogeneity of the other reported components (setting, equipment used, content being taught). More consistency is required in reporting and researching the components of context for realist reviewers to make recommendations to nursing simulation educators. While we were able to identify the themes of feedback, debriefing, SRL, and the intersection between feedback and
SRL, the mechanisms of the 13 studies were heterogeneous and were sometimes not made explicit. For example, it was unclear how conceptual frameworks influenced instructional design because the theory was absent or mentioned at a cursory level or required extrapolation based on the information provided. Similarly, the authors did not explicitly describe the interactions occurring between the components (participants, researchers, equipment). These ambiguous and implied connections make the realist review process challenging and also prohibit replication of the studies.

In terms of outcomes regarding feedback and debriefing, the small number of included studies (N=13) limits what we are able to conclude from our synthesis findings. Our preliminary analysis suggests that learners benefit from more rather than less feedback, and that what educators choose to be the source of feedback (i.e., instructor vs. simulator) should be aligned with the learning objectives and skill of interest. Our synthesis also indicates educational benefits when combining multiple instructional design features (i.e., feedback and debriefing) and multiple feedback sources (i.e., simulator and instructor). When allowing trainees to participate in self-regulated learning, studies suggest the need to have an instructor present to provide support and guidance during practice and skill performance, though when that should occur has not been adequately investigated. Ultimately, the low number of included studies demonstrates that there are opportunities to optimize C-M-O linkages when designing interventions for the various learner populations and content areas in nursing. The heterogeneous and superficial descriptions of and connections between C-M-O, limit our ability to directly answer the inquiries associated with realist reviews. Indeed, many of the included studies replicated findings from other professions in healthcare simulation. At present, nursing simulation educators are left with only vague recommendations because research has not uncovered which instructional design features are optimal for nursing simulation, as well as how best to employ these instructional design features.

4.2 Study Strengths and Limitations

Strengths - This review analyzed and synthesized a subset of the studies identified through a systematic literature search. While the quantity of syntheses and integrative reviews of nursing simulation studies are increasing, we followed the PRISMA and RAMESES standards when conducting this review, which resulted in a methodologically rigorous process. Our process of
using four raters to code, extract and analyze the data independently and in duplicate also added methodological strength to our method and the results. The realist review approach offered a systematic process for analyzing qualitative data that resulted in pragmatic, contextual, specific guidance for simulation educators. Furthermore, this new methodology provided a fresh lens through which to synthesize the evidence and led to important insights about the state of the literature in nursing simulation-based education.

**Limitations** - The limitations of this review directly relate to the limitations of the primary studies. As explained above, in a realist review, because the context, mechanism and outcome are all connected, gaps in any part of the context – mechanism - outcome (C-M-O), limit the usability of the study data. Many of the existing nursing simulation publications do not provide the required details regarding the context, the mechanism and the outcomes of their studies. When a large majority of the studies cannot be included, the results and recommendations we make must be tentative at best. Other limitations with the primary studies included a lack of alignment between the study question and the evaluation assessment used, confounding, or inconsistent evaluation methods between the control and intervention groups.

Due to the iterative nature of conducting a realist review, and because it is a relatively new way of conducting a systematic review, the process was onerous and challenging. At times it felt as if we were regressing rather than progressing. For example, we developed a data extraction tool, however the desired C-M-O information was not consistently reported, so we were forced to revise the tool multiple times and modify what we looked for based on what was reported. Additionally, while coding each study in duplicate increases the rigour of the findings, having four coders understand and agree with such a complex process was challenging and time-consuming. Furthermore, achieving a common understanding of context or mechanism was ambiguous for some variables, which arguably could have been either. For example, we categorized the type of simulator being used as an educational mechanism in one round of coding, and as a contextual factor in the next. This need to be iterative and repeat the process multiple times using revised forms, was time-consuming and led to a drop in team motivation that required group discussion to mediate.

With regards to our process, as noted above, while Pawson et al. (2004) indicate that the initial stage of a realist review is to use the literature to identify existing conceptual theories, we
allowed the theories to emerge during stage two when we analyzed the extracted data instead. Thus we used a modified version of the realist review process outlined in the RAMESES standards. Moreover, we made judgement calls about what data constituted C-M-O and which studies to include based on those criteria, and made subjective decisions regarding how to synthesize the findings. While we believe we were true to the process and applied the realist review standards well, other reviewers may well have made different, but similarly justifiable decisions.

The rigor of the process is also what made it interesting, and led to our recommendations about study reporting, use of educational theory, and methods for conducting knowledge syntheses in the nursing education literature. Rather than adding another synthesis to the body of evidence that indicates that simulation is effective, or that simulation is sometimes effective, we now present evidence, albeit preliminary, for promising instructional design themes and findings. More broadly, we identified some surprising findings about the state of nursing simulation literature, which should give pause to educators and researchers and then hopefully propel the industry forward.

4.3 Comparison with existing literature

As indicated above, the body of evidence regarding the use of simulation in nursing and other health professions education is expanding. While there is evidence that clearly indicates that simulation is more effective than no intervention, beyond that, the evidence is mixed and does not explain how or why simulation works as an educational modality. Thus, simulation educators lack clear direction on what types of simulation to use with specific trainee groups in particular contexts. However, academic organizations and practice settings alike have “bought-in” literally and figuratively to the idea of simulation and thus will continue to use simulation as a teaching and learning tool. By undertaking this realist review, we tried to begin to provide this specific, clear, contextual direction that was absent in previous reviews.

In line with our primary finding, which is that many studies did not report data regarding context – mechanism - outcome (C-M-O), we found such information missing in the study by Hayden et al. (2014). Those authors did not provide specific details about the simulations, including how they were conducted by each school, in each program, and within each course. Without that
information, other educators cannot replicate the training and evaluate the curricula at their own institutions. With regards to the conclusion by Hayden et al. (2014) that there is “substantial evidence” that comparable outcomes can be achieved when up to 50% of traditional clinical hours are replaced by high-quality simulation, we are unable to confirm that claim because they compared simulation to clinical training and we compared different forms of simulation-based training. We argue that our findings suggest that as simulation replaces clinical learning experiences for students, there is an opportunity to use it more effectively to achieve the desired outcomes.

While we are unable to answer the realist review questions of “What works? For whom? Under what circumstances? And why?” our findings offer insight into the gaps and opportunities in the nursing simulation literature. Cook et al. (2013) presented eleven instructional design features that are consistently associated with positive outcomes when they are available. Our secondary findings build slightly on the features of feedback (trainees need more, consistent and regular) and individualized learning (novice learners need appropriate educational supports).

With regards to the previous nursing simulation reviews, while the realist review enabled us to examine and analyze the data from a completely different lens, our findings complement some of their conclusions. Harder (2010) and Kardong-Edgren et al. (2010) indicate that interprofessional simulation research based on scenarios that replicate the true clinical environment are required, while our findings confirm that the majority of studies used nursing students as participants (54%). Additionally, a number of researchers (Kaakinen, & Arwood, 2009; Rourke, Schmidt & Garga, 2010; Waldner & Olson, 2007; Franklin et al., 2015) recommend using teaching and learning, and nursing theory to design and interpret findings in research studies, and we found that without a theoretical framework, C-M-O data not was not considered and reported in the studies.

Our findings also add evidentiary support to the concerns and strategies recently identified by Franklin et al. (2015), including to ground the study in a theoretical framework and to use rigorous research methods with evaluation instruments that match the outcomes. Our review supports Franklin et al.’s (2015) recommendations to learn from simulation research knowledge within and beyond the nursing discipline, to adopt and use consistent terminology regarding simulation, and to collaborate rather than publishing single author reviews.
4.4 Implications for Educators

If the results of Hayden et al. (2014) are any indication, the use of simulation will continue to increase in nursing education in academic as well as practice settings. When using simulation, educators need specific direction on which equipment to use, where to provide the training, what to do regarding instructional design and how to use simulation to achieve the relevant course and program learning outcomes. We undertook this review with the aim of providing this pragmatic information, at least in a preliminary fashion. However, as demonstrated, due to the gaps in context – mechanism - outcome (C-M-O) data provided, 69% of the studies had to be eliminated.

What we can reiterate is that there is a role for simulation in nursing education because simulation consistently leads to better outcomes than no intervention. Furthermore, simulation is sometimes better than other instructional methods, but not always. More importantly, educators must be selective and purposeful when using simulation because instructional design is more important than the mere presence or absence of simulation. Specifically, educators’ thoughtfulness and intentionality in facilitating task alignment between the skills that trainees practice and demonstrate during simulation and the desired learning and clinical outcomes expected of graduates, may help educators achieve the full potential of simulated learning. We have shown that the areas of early interest in the nursing simulation literature are feedback, debriefing, self-regulated learning, and the intersection between feedback and SRL. Beyond these design features, we recommend that nursing educators consider using many others for which robust evidence has accumulated (e.g., mastery learning, curricular integration, range of difficulty and more – see Cook et al., 2013).

4.5 Implications for Researchers and Future Directions

For researchers studying simulation, what is required first is a common definition or understanding of simulation that is accepted and used by all simulation experts. Currently, researchers’ and educators’ meaning of simulation varies from very narrow, to very broad. For some, it is what happens in the lab, using mannequins. For others, it can happen anywhere and includes any activity that represents another activity. Clarity of definition will facilitate more effective and accurate comparisons of interventions and outcomes. Related, is a need to re-examine the concepts of low, medium and high-fidelity, as these loaded terms need to be
unpacked, studied and analyzed to ensure researchers and educators are able to separate the simulator equipment from the instructional design and to understand the unique contributions of all kinds of simulation. Rather than assuming that “high-fidelity simulation” is always necessary and leads to the best outcomes, further research and analysis defining and comparing the different types of simulation equipment as well as the instructional design strategies that are employed when utilizing the equipment would be valuable. Or, consider abandoning the use of the term fidelity completely and use the terms physical resemblance, and functional task alignment as suggested by Hamstra et al. (2014). Furthermore, authors are attributing “more feedback” to ‘high-fidelity’ mannequins and “less feedback” to ‘low-fidelity’ mannequins, but these concepts and functions are mutually exclusive. Fidelity appears to have been used mostly as a reflection on the appearance and use of the equipment, whereas feedback is an instructional design feature that can be manipulated as an intervention. Conflating the two is problematic and should be a practice we avoid.

Then, it is imperative to move beyond the types of studies that currently exist. Nursing simulation researchers can stop studying whether simulation works or not, or even comparing simulation with other instructional modalities. To move simulation research forward, there is a need to compare one kind of simulation against another kind of simulation, with manipulations of instructional design features being the intervention. Rather than adding additional types of technology, and studying the effects, exploring what works in instructional design (debriefing; simulation design characteristics; feedback; faculty interaction, etc.), for whom (novice nurses; associate degree nurses; critical care nurses, etc.), under what circumstances (fundamental nursing skills courses; continuing education; clinical practice; patient outcomes, etc.) will have pragmatic and theoretical value for nursing simulation educators. Of critical importance however, is the full reporting of the contexts, mechanisms and outcomes because the realist review method is only as good as the primary research on which it is based.

A theoretical framework must also underpin any new studies. By using a theoretical framework, the study will be more methodologically rigorous and grounded in sound instructional design principles, with each context – mechanism - outcome (C-M-O) component being deliberately contemplated and reported. When considering the context, grounding the study in research theory would cause the researchers to consider the replicability of their study, which would
facilitate an explicit description of the participants, the setting, the content being taught and the
equipment being used. Similarly, grounding the study in nursing and educational theory would
facilitate thoughtfulness about the context and mechanism – what is being taught, by whom, to
whom, where, why, and how? As well as what are the assumptions underlying why the authors
expect the intervention to impact education in a certain way? The C-M-O links must align so that
an outcome related to the expected mechanisms is in fact what is being measured. One way to do
this might be for researchers to use the course and program learning outcomes for which
simulation is being employed as the study outcome measures. As a very simple example, if
simulation is being used to train nursing students on how to measure vital signs, the intervention
could be feedback (with everything else about the context and mechanism being the same), and
the students’ ability to measure vital signs would be the study outcome measure. Furthermore,
when the C-M-O questions are considered, theory is often the best way to provide the language
and way of thinking necessary to develop the appropriate linkages. For the example above, that
would involve framing the simulation training around a nursing or educational theory that
supports nursing students learning how to measure vital signs. Only then will the value of the
research be realized with pragmatic recommendations for simulation educators, along with
theory-building or theory-refuting data for researchers.

While limitations in the primary data tempered our findings, the realist review process is a
rigorous, valuable, informative method that can provide very specific, clear direction to nursing
simulation educators, not about whether simulation works, but how to use it. If nursing
simulation researchers are able to improve the reporting of their studies based on our
recommendations, the pragmatic suggestions we hoped to make will be possible for future
knowledge syntheses. The result, we suggest, would be that nursing educators would use
simulation to optimize their learners’ results, and researchers would be able to build evidence for
educational principles in the simulation-based training context.
5 Conclusion

The use of simulation in the training of health professionals is increasing, and there is no indication that the trend will stop. Nor should it. There is clear evidence that simulation has a role to play in health professions education, and that in some cases, it can play a valuable role. Schools and hospitals have invested time, money and other resources in acquiring simulation equipment and in training their educators. Similarly, vendors continue to invest in improving the simulation equipment they are manufacturing.

This review offers those involved in nursing simulation (educators, researchers, manufacturers) an opportunity to take a step back and regroup. The existing evidence conclusively demonstrates that using simulation as an instructional strategy is better than doing nothing. But the evidence is inconclusive when comparing simulation to other types of instructional design because the construct of the design (the how) is more important than the specific strategy (the what). The evidence comparing two different types of simulation is similarly inconclusive. This is why the study of simulation is ripe for the realist review methodology. The question is no longer a simple one asking “does it work or not?”. Instead, simulation researchers must explore what kind of simulation works (instructional design), for which type of trainee, and under what circumstances. This will require the study design to be informed by a theoretical framework, with explicit context – mechanism - outcome (C-M-O) links that are clearly considered during study conception and fully reported during study dissemination. Researchers are encouraged to consider designing studies that vary the context (trainees, learning space, simulator) or the mechanism (theoretical framework, instructional design method) and to evaluate relevant outcomes for which favourable validity evidence has been collected. While realist research studies are more complex and time consuming to design and conduct, the results should contribute significantly more pragmatic knowledge to the body of simulation evidence.

Using our team-designed process, we excluded 69% of the published primary nursing studies comparing two types of simulation because one or more components of the C-M-O data were not provided. The remaining 31% of the studies demonstrated that feedback, debriefing, self-regulated learning (SRL), and the intersection of feedback and SRL are the only four themes that have received adequate attention to date. Nursing education researchers and the editors of the
journals in which they publish must improve the standards of reporting so that future knowledge syntheses are better able to make recommendations related to when simulation should be used, for which trainees, under which circumstances, and why.
References


audiovisual feedback and debriefing. Critical Care Medicine, 36(10), 2817 – 2822. doi: 10.1097/CCM.0b013e318186fe37


Internal Medicine, 14, 625 – 628. doi: 10.1007/s11606-012-2198-y


to nursing students’ confidence and competence: A systematic review. *International Nursing Review, 32*(1), 26 – 33.


## Appendix 1
Data Extraction Form 1

<table>
<thead>
<tr>
<th>Context</th>
<th>Mechanism</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who? (learner / instructor)</td>
<td>How?</td>
<td>How well?</td>
</tr>
<tr>
<td>What? (task / simulator)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Where? (setting)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>When? (program / training)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Is this meaningful?

Which theory? (explicit or implicit)
# Appendix 2
## Data Extraction Form 2

<table>
<thead>
<tr>
<th>Context</th>
<th>Mechanism</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who is there? (RNs, RPNs, undergrads, post-grads, etc.)</td>
<td>(If any) What learning principles are noted as having influenced the instructional design?</td>
<td>What are the expected outcomes?</td>
</tr>
<tr>
<td>What is there? (equipment, supports, instructors, etc.)</td>
<td>What is everyone actually doing? (i.e., how is the activity experienced by everyone?)</td>
<td>What are the explicit (authors described) outcomes?</td>
</tr>
<tr>
<td>How is the learning setting organized?</td>
<td>What interactions are occurring (i.e., between people, between equipment, between all components)?</td>
<td>What are the implicit or unintended consequences of the activity?</td>
</tr>
<tr>
<td>Where is the learning setting?</td>
<td>College / University / hospital? Lab / classroom?</td>
<td>How does the setting impact the outcomes? What, if any, are the differences in outcomes based on setting?</td>
</tr>
</tbody>
</table>
# Appendix 3
Data Extraction Form 3

<table>
<thead>
<tr>
<th>Context</th>
<th>Mechanism</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is there (equipment, supports, instructors, etc.)?</td>
<td>(If any) What conceptual framework / theory is noted as having influenced the instructional design?</td>
<td>How well does the outcome match the conceptual framework? That is, are the authors testing the theory well or not?</td>
</tr>
<tr>
<td>How is the learning setting organized?</td>
<td>How well is the theory / conceptual framework utilized?</td>
<td>How are the outcomes measured? (e.g., by whom? when?)</td>
</tr>
<tr>
<td>Where is the learning setting (College / University / Hospital / other)?</td>
<td>What interactions are occurring: between participants, between investigators/research assistants, between participants &amp; equipment, between all components?</td>
<td>How does the setting impact the outcomes? What, if any, are the differences in outcomes based on setting?</td>
</tr>
<tr>
<td>What simulator is being used (e.g., mannequin, task trainer, etc.)?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4
List of Excluded Studies (missing C-M-O data)


McCormick, M. J., Burton, T., & Werts, N. (2010). Case scenarios and simulations as techniques to facilitate asthma education and team building among health care students. *Journal of*


