From Bricks to Buildings
Adapting the Medical Research Council Framework to Develop Programs of Research in Simulation Education and Training for the Health Professions

Faizal A. Haji, MD;
Celina Da Silva, MN;
Delton T. Daigle, PhD;
Adam Dubrowski, PhD

Summary Statement: Presently, health care simulation research is largely conducted on a study-by-study basis. Although such “project-based” research generates a plethora of evidence, it can be chaotic and contradictory. A move toward sustained, thematic, theory-based programs of research is necessary to advance knowledge in the field. Recognizing that simulation is a complex intervention, we present a framework for developing research programs in simulation-based education adapted from the Medical Research Council (MRC) guidance. This framework calls for an iterative approach to developing, refining, evaluating, and implementing simulation interventions. The adapted framework guidance emphasizes: [1] identification of theory and existing evidence; [2] modeling and piloting interventions to clarify active ingredients and identify mechanisms linking the context, intervention, and outcomes; and [3] evaluation of intervention processes and outcomes in both the laboratory and real-world setting. The proposed framework will aid simulation researchers in developing more robust interventions that optimize simulation-based education and advance our understanding of simulation pedagogy.

Key Words: Framework, Program of research, Health care simulation, Medical research council.

Research in health care simulation, while still young, is maturing rapidly. As our experience with simulation-based education (SBE) grows, questions arise about what the focus of inquiry should be and how best to advance knowledge in the field. Such “growing pains” are hardly unique to health care simulation. The broader scientific community has often benefited from similar introspective moments, as in 1963 when Bernard Forscher1 published the now widely cited editorial “Chaos in the Brickyard.” In this article, Forscher uses the analogy of bricks and buildings to demonstrate the perils of supplanting the purpose of scientific inquiry with its methods. He argues that every scientific discipline starts with small-scale projects to explore unstudied phenomena; these form bricks of knowledge that serve as the foundation for future inquiry. However, as the field matures, a tipping point is reached where a programmatic approach to research (ie, the systematic integration of knowledge through sequential studies that build on one another)2 must replace a project-based strategy. This programmatic approach prevents the foundations of true knowledge from being “buried under an avalanche of random bricks.”1

Recent arguments within health care simulation scholarship suggest that the field has reached a similar “tipping point.”3–5 Gaba6 has identified that health care simulation largely depends on innovation from “investigator-initiated” research and such individual projects address “only one small part of a much larger whole.” Moreover, he argues that there is a “need to articulate more fully the key themes and questions and to demonstrate the linkage between them and our projects.”7 Moving forward, the space science concepts of the Decadal Survey and the Science Traceability Structure (STS) that Gaba has introduced8 will likely become important processes that guide research prioritization and link individual programs, projects, and study designs to overarching themes, goals, and objectives in the field. Similar thinking can be found in the reports from 2 recent agenda-setting conferences: the Society for Simulation in Healthcare Research Summit9 and a recent Utstein-style meeting.8 These seminal meetings identified a number of research questions in health care simulation requiring further investigation. Many of the questions “cannot be answered in a single study, but require several studies that build upon one another.”4

As these efforts continue to clarify the research priorities in health care simulation, investigators may benefit from guidance on how to effectively build research programs capable of addressing these areas of inquiry. Among the directions for future research uncovered during the Utstein-style meeting, three major themes emerged: instructional...
design, outcome measures, and translational science research (TSR). Although frameworks have been published to guide investigators in TSR, in the development of simulation for assessment, and in outcome measurement, there is comparatively less guidance on building programs of research to optimize instructional design of SBE interventions.

This is particularly important given recent systematic reviews and meta-analyses, which leave little doubt that SBE yields large positive effects on participants’ satisfaction, knowledge, skills, and practice behaviors. Notwithstanding these results, much of the current scholarship has focused on justifying the effectiveness of simulation in various contexts and has largely been conducted on a “study-by-study” basis. Although this is in part due to political, economic, and bureaucratic factors limiting time, resources, and funding available for programmatic simulation research, the result is a body of evidence that is at times chaotic, contradictory, and limited in advancing the understanding of “what works for whom, and under what circumstances.” For instance, a recent critical review demonstrated mixed results regarding the impact of simulation fidelity on transfer of knowledge and skills. The studies included in this review were individual efforts that did not systematically build on one another, involved a variety of learner populations and clinical contexts, and did not share a common operational definition of “fidelity.” Consequently, substantial uncertainty remains about this important feature of simulation instructional design. This “uncoordinated accrual of information” amounts to little more than adding bricks to an already full brickyard; it neither builds robust understanding of SBE for researchers nor provides useful information to educators on optimizing simulation programs and integrating them into existing curricula.

To build the existing knowledge base, the focus of research should move away from asking if SBE is effective and toward understanding why simulation interventions are or are not effective and clarifying how to use SBE to optimize outcomes. This requires a clearer conceptualization of the instructional design features of SBE programs that bring about desired effects—the so-called “active ingredients” in simulation-based learning and research environments. As identified in the aforementioned agenda-setting conferences, the latter line of inquiry should consider questions at several levels: (1) What is the theoretical basis of these instructional design features? (2) What are the resource requirements and system challenges impacting the application of these theories to SBE interventions? (3) How can programs address SBE intervention implementation given local resource challenges and within the context of the broader health care system?

To advance understanding along this continuum, investigators may benefit from a systematic approach to the design, refinement, evaluation, and implementation of SBE interventions. In developing this approach, one can draw upon the experience found in fields with similar challenges. As others have acknowledged, simulation (whether delivered as a stand-alone experience or as an augmentation of other educational modalities) is a complex sociotechnical intervention. In this article, we present one approach to building programs of research adapted from the Medical Research Council (MRC). Although its application to health professions education is novel, the MRC framework has been successfully applied to health services and public health research.

In the sections to follow, we expand on the concept of simulation as a complex intervention. Next, we present our adaptation of the MRC framework as it pertains to programmatic health care simulation scholarship. To supplement this discussion, we describe numerous studies from the health services and health care simulation literature that may serve as successful examples in each phase of the proposed framework. Finally, we provide general considerations and discuss the strengths and limitations of the MRC approach, given the current climate of health care simulation research and the need for adequate resources, funding, and expertise for investigators to perform this work.

SIMULATION AS A COMPLEX INTERVENTION

Complex interventions are “built up from a number of components, which may act both independently and interdependently.” Such interventions are characterized by a multitude of active ingredients that make standardization of intervention delivery difficult, long causal chains linking them with intended outcomes, and features that mutate upon adaptation to local contexts. Thus, complex interventions represent open systems that feed back onto themselves, potentially changing the conditions that made them work in the first place. Consequently, it is difficult to define, develop, document, and most importantly, replicate complex interventions.

In the health care simulation context, these challenges may be exacerbated by (1) a lack of clarity in the theoretical foundation or pedagogical features underpinning an intervention, (2) limited resources or time to refine and pilot the intervention and to ensure the intervention’s feasibility and acceptability, (3) a small pool of research subjects leading to underpowered designs, (4) the lack of validity evidence supporting assessment measures, and (5) the difficulty with controlling for confounding variables, typically owing to participants not being isolated from other educational experiences or local contexts. Likewise, the processes linking simulation interventions and outcomes are often nonlinear and may be mediated by the context of the health care system that the intervention is embedded within (especially when SBE interacts with other educational methods), making transitions from the laboratory to the “real world” problematic. In the next section, we describe a framework for designing, refining, evaluating, and implementing complex interventions in health care simulation that may help address some of the aforementioned challenges.

ADAPTING THE MRC FRAMEWORK FOR HEALTH CARE SIMULATION

The MRC framework was published in 2000 and subsequently revised in 2008, to reflect the experience that had accumulated since the initial publication. Although the
2000 version proposed a stepwise, linear approach to designing and evaluating complex interventions (following a structure analogous to the development of pharmaceutical therapies),37 the 2008 framework proposed a nonlinear approach that emphasizes both piloting and implementation studies; integrates process and outcome evaluation; acknowledges the importance of alternative, nonexperimental research methods; and highlights the importance of understanding the intervention’s context.29

In adapting this framework, although we have retained much of the original MRC language, we have modified the overall structure and the elements within each phase to reflect the unique needs of health care simulation research. The adapted framework calls for research grounded in established theories of instruction and learning and that investigates the mechanisms underpinning simulation interventions in advance of evaluating their effectiveness. Ultimately, we promote research that seeks to understand the relationship between an intervention, its context, and its outcomes, both in controlled research settings and in the real world. In our conceptualization, each of the phases described in the original MRC framework are embedded within larger “cycles” of the research process (Fig. 1), which are described in detail later. The specific research questions, methodological considerations, and outcomes for each phase are delineated in Table 1, whereas the example research studies corresponding to each phase are presented in Table 2.

Cycle A: Theory and Modeling

The first cycle of the research program encompasses the “development” phase, which encompasses 2 subphases: theory and/or evidence identification and intervention modeling. In the first subphase, the purpose of the research program and primary research question(s) is identified. Commonly, this is achieved by identifying a gap in the literature or through the identification of a clinical or educational “problem.” To ground the work, investigators should identify theories or conceptual frameworks that help to refine research question(s), generate hypotheses, identify suitable outcomes, and inform the design of subsequent simulation intervention(s) to be studied (by delineating mechanisms or “active ingredients”). *This is important in evolving health care simulation scholarship, as articulation of theoretical assumptions is important in informing each research study and linking the results of individual studies together in a meaningful way.4,8

Theory identification may be coupled with a thorough review of the literature to identify the evidence-base related to the research question and/or the proposed intervention. The choice of what type of review to conduct depends largely on the goals of the literature synthesis. For instance, where the goal is to summarize an existing body of evidence related to the research question, a systematic review can be considered. If there is limited research on a given topic, investigators may expand the search beyond the health care simulation domain, to related areas (eg, human factors, cognitive psychology, or motor learning) through a scoping review. Alternatively, where the goal is to identify a “gap” in the literature, a “critical” review may be more appropriate. Finally, where a relevant theoretical framework cannot be identified, investigators may consider a realist review to identify essential components and mechanisms at play within a given intervention.38,61

In the second subphase, investigators use the selected theory and knowledge of the evidence base to model the intervention. This involves mapping out the content, structure, and delivery of the intervention and its components and

---

*In choosing relevant theories, “grand theories” may be too abstract or broad in scope to identify active ingredients or generate specific, testable hypotheses. “Mid-range” theories—derived from empirical evidence that have shown logical adequacy over time and across multiple contexts (eg, cognitive load theory66 or Fitz and Posner’s model on motor skills learning67)—may be more suitable. In addition, investigators should also identify competing theories that address the research question(s), which may lead to alternate instructional designs to be tested in the next cycle.*
TABLE 1. Detailed Description of Adapted MRC Framework Phases

<table>
<thead>
<tr>
<th>Phase</th>
<th>Questions</th>
<th>Methods</th>
<th>Outcome/Evidence Generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development: Theory and</td>
<td>■ What is the “problem,” “gap,” or “question” the research will address?</td>
<td>■ Needs assessment of stakeholders</td>
<td>■ Clear purpose, research question(s) and hypotheses</td>
</tr>
<tr>
<td>evidence identification</td>
<td>■ Which theories/conceptual frameworks can be used to refine the question</td>
<td>■ Literature review (critical, narrative, systematic, scoping or realist)</td>
<td>■ Theories to potentially explain “how” and “why” the intervention will generate the expected outcomes and identify the “active ingredients” that allow it to do so</td>
</tr>
<tr>
<td></td>
<td>and generate hypotheses, and design the simulation intervention?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>■ What are the theoretical mechanisms behind this intervention and what outcomes will it generate? What competing theoretical perspectives can inform alternate interventions?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>■ What evidence exists in the health care simulation literature or related domains (eg, education, human factors, motor learning) on the design/ effect of similar interventions?</td>
<td>■ Review common theories or conceptual frameworks (eg, learning, behavior, and organizational theories)</td>
<td>■ Empirical evidence to support proposed intervention design and anticipated outcomes</td>
</tr>
<tr>
<td>Development: Intervention</td>
<td>■ What will the content, delivery, and structure of the program be?</td>
<td>■ Conceptual maps, logic models, causal model, flowcharts, etc., linking context, intervention, and outcomes</td>
<td>■ Detailed map of the content, structure, and delivery of the intervention and its components based on theoretical and empirical understanding of the research question</td>
</tr>
<tr>
<td>modeling</td>
<td>■ What are the outcomes of interest (eg, skills, behavior) and how will they be measured?</td>
<td>■ Focus groups, descriptive studies, key informant interviews, case studies, and surveys (eg, Delphi)</td>
<td>■ Outline of the relationship between intervention components and the program context</td>
</tr>
<tr>
<td></td>
<td>■ What are the active ingredients of the program (mode of simulation, duration/location of training, instructional design, learner characteristics, etc), and how do they interact with program context?</td>
<td>■ ■ What are the mechanisms linking components with desired outcomes?</td>
<td>■ Articulation of the proposed mechanisms, ie, refinement processes leading from the intervention to the anticipated outcomes</td>
</tr>
<tr>
<td></td>
<td>■ What is the context that the intervention will be implemented in (eg, identity/culture of stakeholders, individual/system level barriers and facilitators to uptake)? What are the relationships between the intervention components and the context?</td>
<td>■ Data from primary research regarding program content and outcome measures</td>
<td>■ Data from primary research regarding program content and outcome measures</td>
</tr>
<tr>
<td>Piloting</td>
<td>■ Is the proposed intervention acceptable and feasible to target learners, other stakeholders, and in this context? What are the costs? How will barriers to implementation be addressed?</td>
<td>■ Surveys, focus groups, interviews with stakeholders, cost-benefit analysis</td>
<td>■ Feasibility, acceptability, and cost-effectiveness of the intervention established</td>
</tr>
<tr>
<td></td>
<td>■ How do varying “active ingredients” (eg, sequence of tasks, amount of feedback, frequency of training) affect outcomes, and what is the “optimal” design of the intervention?</td>
<td>■ Small-scale experimental and quasi-experimental comparative studies (pilot/exploratory trials)</td>
<td>■ Clarification of intervention design questions and understanding of how components of the intervention affect outcomes</td>
</tr>
<tr>
<td></td>
<td>■ What is the “optimal design” of training of the comparison group?</td>
<td>■ Observational studies</td>
<td>■ Identification of an appropriate comparison group (ideally based on a competing theory)</td>
</tr>
<tr>
<td></td>
<td>■ What validity evidence supports the proposed outcome measures?</td>
<td>■ Ethnography</td>
<td>■ Resolution of potential methodological issues before the definitive evaluation</td>
</tr>
<tr>
<td></td>
<td>■ What is the estimated effect size of the intervention? How will biases and confounders be mitigated? How will participants be randomized?</td>
<td></td>
<td>■ Establish validity evidence for assessments</td>
</tr>
</tbody>
</table>

(continued on next page)
Then linking these with anticipated outcomes. This can be achieved using a variety of methods, ranging from simple paper-and-pen exercises (e.g., logical models and flowcharts) to full computer-based models. Primary research (using focus groups, informant interviews, case studies, ethnographic and survey-based methods) may also be required to clarify elements of the intervention (e.g., program content). To be effective, investigators must model “active ingredients” and identify how these relate to each other as well as how they relate to the program’s context. Clear articulation of these mechanisms and interrelationships is critical to “standardizing” the process and function of the intervention. This ensures the integrity of the SBE program during the evaluation and implementation phases, even if nonessential features vary with local conditions.

Completion of the first and second subphases is likely to be both closely related and proceed in a cyclical, iterative manner. As evidence is reviewed and research questions are refined, the selected theories may be revisited. Similarly, as the intervention is modeled, uncertainties may arise (e.g., regarding mechanisms of action, choice of research outcomes), which may prompt a search for additional evidence or theory to support the design-related decisions made. If uncertainties remain, these can be articulated as specific research questions (e.g., regarding competing instructional design features) that will be addressed during the next cycle of the research program.

**Cycle B: Piloting**

The second cycle begins with piloting, in which investigators populate the intervention model with primary research data. Piloting is divided into four subphases that are independent and can be completed in any order (and in some cases, possibly in tandem). However, all subphases should be completed before evaluating the intervention’s effectiveness.

The first subphase establishes the feasibility and acceptability of the intervention for program stakeholders. In addition, potential barriers to program implementation should be considered and mitigated to the extent possible. Addressing these issues at this early stage ensures that the program will be appropriate to implement, given the target population and context.

The second subphase clarifies uncertainties in the design of the intervention (i.e., a “formative” evaluation of the intervention) and outcome assessment. This is critical because the first research cycle will likely result in several questions, which may prompt a search for additional evidence or theory to support the design-related decisions made. If uncertainties remain, these can be articulated as specific research questions (e.g., regarding competing instructional design features) that will be addressed during the next cycle of the research program.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Questions</th>
<th>Methods</th>
<th>Outcome/Evidence Generated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation</td>
<td>■ Are the participants randomly assigned? Is the study sufficiently powered? Are outcome assessments and allocation blinded? Is loss-to-follow-up addressed?</td>
<td>■ Process evaluation (e.g., focus groups with learners, observational studies during intervention delivery)</td>
<td>■ Definitive/summative evidence about the effect of the intervention on desired outcomes in a controlled setting, as well as unintended outcomes</td>
</tr>
<tr>
<td></td>
<td>■ Was the intervention delivered as planned, and were the anticipated outcomes realized? Were any unexpected outcomes observed? What additional processes emerged, and how do they relate to the observed outcomes? What can be learned about how and why the intervention works from these data?</td>
<td>■ Outcome evaluation using randomized, controlled experiments (preferred) or quasi-experimental/observational studies</td>
<td>■ Confirmation that the intervention was delivered in accordance with the model from phase 1</td>
</tr>
<tr>
<td></td>
<td>■ How generalizable are the results?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation</td>
<td>■ What are the unintended effects of the intervention? What are the long-term costs?</td>
<td>■ Observational studies</td>
<td>■ Insight into unanticipated processes, which allowed or prevented the intervention from producing the desired effect</td>
</tr>
<tr>
<td></td>
<td>■ How can the intervention be incorporated into existing training curricula? How can other pedagogical approaches be used alongside the simulation intervention?</td>
<td>■ Qualitative methods (ethnography, follow-up focus groups, interviews, and surveys of stakeholders and participants upon return to “normal duties”)</td>
<td>■ See how the program works in the real world.</td>
</tr>
<tr>
<td></td>
<td>■ How does the context change the intervention, and vice versa? How do both change over time?</td>
<td>■ Economic analyses</td>
<td>■ Identify long-term and unintended outcome</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>^Identify how intervention processes/outcomes change in new locations and impact local context</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>■ Synergies, conflicts, and optimal strategy of integrating the program into the curriculum or with other pedagogical approaches</td>
</tr>
</tbody>
</table>
TABLE 2. Mapping of MRC Phases to Publications in Health Services and Health Care Simulation Research

<table>
<thead>
<tr>
<th>MRC Phases</th>
<th>Health Services Research Example Publications</th>
<th>Health Care Simulation Example Publications*</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Development A—theory</td>
<td>Byrne et al (2006)32 describe the development and piloting of a program for secondary prevention of coronary disease. Their study used a literature review of key components in existing interventions and patient interviews to model the intervention, which was pilot tested in 4 practices and evaluated using follow-up interviews with practitioners.</td>
<td>A. Theory and evidence identification66 conducted a critical review of debriefing, including existing principles, techniques (self vs. facilitated, video-playback, etc.) and gaps in the literature. The authors identify the need to build models or theories of debriefing specific for simulation-based learning.</td>
</tr>
<tr>
<td>and evidence identification</td>
<td>MRC Phases Health Services Research Example Publications Health Care Simulation Example Publications*</td>
<td></td>
</tr>
<tr>
<td>B. Intervention modeling</td>
<td>Thompson and Weiss (2006)42 conducted a pilot study to identify &quot;active ingredients&quot; in a homeopathy intervention, interviewing patients before and after delivery of a 5-visit &quot;package of a care&quot; homeopathy program.</td>
<td>Kneebone et al (2004)68 presented their model for SBE that combines psychomotor and communication skills training by incorporating standardized patients with inanimate models within simulation environments. The authors suggest that such &quot;patient-focused simulation&quot; facilitates training in authentic simulated contexts that reflect the realities of clinical practice.</td>
</tr>
<tr>
<td>II. Piloting</td>
<td>Murchie et al (2007)33 modeled and pilot tested an integrated follow-up program for cutaneous melanoma. The intervention was designed from an iterative process that included a literature review, consultation with an expert steering group, and semistructured interviews with patients.</td>
<td>Kneebone et al (2006)64 explored the feasibility of their patient-focused simulation concept for assessment of technical and communication skills through a pilot study of the Integrated Procedural Performance Instrument (IPPI), whose psychometric properties were subsequently described in a later study (LeBlanc et al, 2009).</td>
</tr>
<tr>
<td>III. Evaluation</td>
<td>Murchie et al (2010)35 conducted an RCT of the efficacy of their cutaneous melanoma program on guideline compliance, health status, and patient satisfaction. To evaluate the process of the intervention and identify &quot;how it worked,&quot; they also conducted a qualitative interview-based study of participating GPs to determine their experience with running the program (Murchie et al, 2009)34</td>
<td>Moulton et al (2009)47 applied the IPPI as a communication skills teaching tool for medical students and residents during simulation-based training in wound closure and urinary catheterization. The RCT compared feedback from a standardized patient using video-review of IPPI scenarios with no feedback on subsequent technical performance and patient communication.</td>
</tr>
<tr>
<td>IV. Implementation</td>
<td>To follow up on the implementation of their cutaneous melanoma program, Murchie et al (2010)30 conducted a qualitative study to determine the experiences and feelings of participating patients with regard to comfort, continuity of care, and any concerns with the intervention.</td>
<td>Stefanidis et al (2012)46 subsequently conducted an RCT demonstrating training to automaticity leads to superior transfer of skills to the operating room for novices compared with training to &quot;proficiency.&quot;</td>
</tr>
</tbody>
</table>

Barsuk et al (2009)49 conducted a cohort study evaluating central venous catheter (CVC) insertion skills of residents receiving clinical training vs. simulation training featuring mastery learning and deliberate practice.

Moulton et al (2009)47 applied the IPPI as a communication skills teaching tool for medical students and residents during simulation-based training in wound closure and urinary catheterization. The RCT compared feedback from a standardized patient using video-review of IPPI scenarios with no feedback on subsequent technical performance and patient communication.

In a preobservational/postobservational study, Barsuk et al (2009)51 examined the incidence of catheter-related bloodstream infections over a 32-mo period on a medical intensive care unit staffed by simulation-trained residents. Barsuk et al (2010)52 demonstrated CVC insertion skills acquired in the simulation laboratory under mastery training and deliberate practice conditions are retained at 12 mo and revealed that simulation training resulted in medical care cost savings (Cohen et al, 2010).53 Recently, Barsuk et al (2011)54 demonstrated unintended positive effects of the intervention, with improved CVC insertion skills of junior trainees taught by simulation-trained seniors.

*GPs = general practitioners. Only a few examples of studies from each phase are provided here for illustrative purposes; however, the health care simulation literature contains many more examples. A more comprehensive review of existing piloting and evaluation studies can be found in recent systematic reviews of studies comparing simulation with no intervention (Cook et al, 2011) and those comparing various simulation instructional design features (Cook et al, 2013).
outcomes? and (3) what is the best way to measure the outcomes of interest? Successful completion of this subphase also requires investigators to identify or gather validity evidence concerning the proposed outcome measures and conduct small-scale experimental and quasi-experimental studies in which active ingredients are varied, to study resulting changes in outcomes of interest. The chosen variations in these instructional design features should be informed by theory, to help clarify the mechanisms linking the intervention components under investigation with the outcome observed.

The third subphase involves identifying and designing the training protocol for a comparison group, to which the primary intervention may be compared in the evaluation phase. The comparison group should be identified, modeled, and piloted with a competing instructional or learning theory in mind and with the same rigor as the intervention under examination.

The final subphase addresses methodological issues before undertaking the “summative” evaluation. Where this evaluation will involve determining the effectiveness of the intervention, it is important to generate estimates of the sample size needed to power a large-scale prospective randomized controlled trial (RCT). This may necessitate a pilot RCT comparing the intervention and comparison group to determine an expected effect size.

In completing each of these subphases, important data will be generated about the appropriateness of the selected theories or conceptual framework(s), outcome measures, model of the intervention and comparison groups, and understanding of the context in which the program operates. Thus, investigators should adopt a reflexive approach, where the assumptions and decisions made during the development phase are revisited, to reflect on “if and how” these decisions need to be amended based on the data accumulated. It is likely that some elements of the intervention will need to be remodeled in light of this new knowledge and underlying theories may be challenged. In turn, new or additional conceptual frameworks may need to be considered to help understand what findings have been uncovered. Investigators can then initiate further studies that build on these findings, demonstrating a programmatic approach. In addition, the findings from pilot studies should be disseminated, especially if established theories are challenged and novel theoretical positions are adopted. This type of research advances our fundamental understanding of SBE.

**Cycle C: Evaluation**

The third cycle of the research program involves a summative evaluation of the intervention’s effectiveness, to justify its implementation in the real world. To be most informative, the evaluation should consider not only whether the intervention achieved its intended outcomes but also any unintended outcomes as well as intended and unintended processes. Evaluation of outcomes is the predominant practice in simulation research. Well-structured outcome evaluation ensures that the study is sufficiently powered (potentially necessitating multicenter trials), carefully considers and controls for potential confounders, and randomizes participants or uses other quasi-experimental approaches (ideally with a comparison group) when RCTs are not possible. Specific biases, threats to validity, and implications for generalizability of results may need to be addressed in the design of the evaluation, including the appropriateness and method of randomization, whether concealment of allocation from participants and faculty (eg, expert raters) is possible, management of participant preferences for the intervention or comparison group, whether inclusion and exclusion criteria are representative of the target population, the implications of these criteria on recruitment and retention, and the context in which the trial should be conducted to balance control of confounding variables with generalizability of results.

Although program processes are infrequently studied in simulation research, it is also important to understand whether the intervention was delivered as designed. When coupled with outcome evaluation, process evaluation can help clarify causal mechanisms and identify contextual factors associated with observed variations in intervention delivery or outcome. In doing so, the evaluation generates essential knowledge regarding the next steps of the research program. When intended outcomes are realized (ie, a “positive” result), the process evaluation may yield valuable insight regarding what intervention components must be in place for effective implementation of the intervention in the real-world setting. Conversely, when a process evaluation demonstrates delivery of the intervention in accordance with the theory and/or logic model but without achieving the intended outcomes, this should prompt investigators to disseminate these findings. The failure of the theory to generate expected outcomes is a finding of interest in and of itself. Researchers should then identify an alternate theory and explore other features of instructional design, returning to the summative evaluation once the intervention has been revised through additional cycles of the development and piloting phases. As the evaluation acts as the final “gatekeeper,” the findings from this cycle must be carefully reviewed before proceeding to implementation, to ensure that programs which are “rolled out” to end users have been thoroughly developed and evaluated.

**Cycle D: Implementation**

In the final cycle of the research program, the SBE program is implemented into a health care education setting and investigators evaluate the intervention’s effects on “real-world” outcomes. The goal of any simulation intervention is to be successfully taken up by “end-users” (ie, educators and trainees) as a usable “product.” Thus, it is important that investigators devote the necessary time and resources to evaluate the impact of the intervention on educational practice. A number of issues must be considered as the intervention is transferred from the controlled setting of the
evaluation study to the inherently messier real-world setting. This may include, but is not limited to, the following: (a) evaluating the short- and long-term, intended and unintended processes and outcomes of the program (which may or may not differ from those observed during the controlled evaluation study); (b) evaluating the rate of uptake and stability of the intervention (including how it changes when adapted in different contexts); (c) developing an understanding of how contexts change after implementing the intervention (eg, any broadening of subject groups, long-term costs, and cultural changes initiated by the intervention’s presence); and (d) characterizing how these factors change across geographical frontiers and over time. Because simulation will likely augment rather than replace existing educational modalities, this phase may also include research aimed at understanding the optimal strategy for integrating the simulation intervention within existing educational curricula and with other pedagogical approaches (eg, problem-based learning, apprenticeship training).

As in previous cycles, the data gathered from this phase can be used to generate additional research questions, which may further refine future versions of the intervention, specify learner populations and contexts that will have the greatest benefit from the intervention, and provide valuable lessons on how contextual factors influence intervention uptake and delivery. In turn, investigators may choose to return to the development and piloting phases to address unanswered questions. In situations where unintended processes or outcomes emerge or where questions are raised regarding how the intervention can synergize with other pedagogical approaches, investigators may be prompted to evolve the line of inquiry in a new direction and initiate a new program of research.

**GENERAL CONSIDERATIONS**

Although our framework is intended to serve as a guide rather than a prescription for developing research programs in health care simulation, there are some specific considerations in moving through the phases. First, although we have suggested that relevant theories should be articulated at the outset of the research program, investigators may need to draw on multiple theories in various phases of the research program. For instance, learning or instructional design theories are appropriate to ground development and modeling of a proposed simulation intervention; however, theories from cognitive and health psychology may need to be considered if participant behavior change or transfer of knowledge and skills is of primary interest. Where team-based training is the focus (as opposed to individual learning), sociocognitive theories (eg, distributed cognition) may be appropriate. Similarly, theories from implementation and complexity science are helpful in engaging issues of curriculum integration and how interventions and contexts change with time.

Second, although the model is iterative within each cycle of the research program, we do suggest an ordering between research cycles. For example, the cycle of modeling the intervention and pilot testing active ingredients should precede the evaluation cycle, to ensure the intervention has been developed to a point where it is expected to have a worthwhile effect and where the mechanisms by which this effect is generated are reasonably understood. For a variety of reasons, current simulation scholarship has not readily engaged in this process. However, by taking the time to carefully develop and pilot an intervention, including comparing various instructional design features “head to head,” major strides can be made in clarifying how these design features operate (both individually and in concert) and can further our understanding of how and why interventions generate desired outcomes.

The final consideration relates to the choice of research methods for the phases outlined earlier. In general, both quantitative and qualitative methods will provide important and often complementary knowledge. Thus, researchers should seek to generate evidence using a variety of methods that do not share common weaknesses. Although we have provided some methodological considerations for each phase of our framework in Table 1, in general, we concur with McGaghie et al that decisions about research methods and measurement procedures must be fitted to specific research questions, which will likely change as investigators move from one phase to another.

**STRENGTHS, LIMITATIONS, AND IMPLICATIONS OF THE MRC FRAMEWORK**

It is our hope that the proposed framework extends current conceptualizations of simulation as an “intervention” and compliments existing frameworks that also call for cumulative, sustained, thematic, and theoretically based programs of research in SBE. In our view, the strength of the MRC approach lies in its emphasis on modeling simulation interventions based on theory, pilot testing active ingredients, and the iterative nature of the research process. This may account for the success of the framework in health services research, where research programs following each of the phases can be found.

Although applicable to a range of research areas, the MRC framework is particularly suited to guide the research programs in simulation instructional design that have been called for. Because it provides detailed guidance on the process of developing and refining simulation interventions, this framework can also complement approaches such as the TSR model, which focuses on the outcomes measured in simulation research (Fig. 2). Finally, the adapted MRC framework may also support broader attempts to organize research in health care simulation. For instance, the framework can identify and prioritize studies within individual research programs, which can subsequently be linked to overarching themes in the field through the STS.

Conversely, the MRC framework falls short in its guidance regarding the methodology and selection of translational outcomes during the implementation phase. Investigators should look to the TSR and similar models for guidance in these latter phases, to ensure that simulation will have appropriate links to improved health service delivery and patient care. In addition, recent criticisms...
aptly point out that the MRC framework provides limited guidance on evaluating interventions to change complex systems, including interventions at the population, policy, and health economics levels.\textsuperscript{27,69} Thus, analogous issues in simulation (eg, research designed to improve system-level efficiencies, such as simulation for mass casualty events, distributed education models, or blended learning approaches that incorporate simulation with other pedagogical approaches) may need to consider guidance and methods from the science of complex systems to adequately address their needs.\textsuperscript{27}

Achieving success using the proposed approach will likely require a team of investigators investing sustained effort over multiple years (as can be seen by the dates of sequential research studies cited in Table 2). Some investigators may feel the need to focus on “fundable” and “publishable” projects and thus prioritize evaluation studies over both implementation studies and the “early work.” However, in our view, each cycle of the proposed framework may generate information that advances knowledge in health care simulation and should be considered scholarly in its own right. Thus, we advocate for the dissemination of results from each phase, with journals possibly even preferring work clearly rooted in a programmatic framework. For instance, in addition to rigorous literature syntheses, a publication detailing the application of an established theoretical framework to model the instructional design of a potential simulation intervention (which could be considered a “protocol” for the intervention) or a pilot study that clarifies aspects of the intervention’s instructional design will provide valuable data that the authors and other researchers in the field can use to test the proposed intervention components and mechanisms. We recognize that this departs from traditional views of “publishable” work; however, we hope that well-constructed, articulate reports from the aforementioned phases will be viewed as worthwhile contributions by journal editors and other leaders in the field. To be useful, however, authors need to articulate clearly the theoretical basis and context in which their intervention was designed and piloted, provide a “thick description” of the intervention, and situate the results within a larger program of research and the broader research agenda set by the simulation research community.\textsuperscript{6} Without these elements, other researchers and educators cannot determine the relevance of the results to their own educational settings or research programs and will have difficulty replicating or synthesizing the evidence generated.\textsuperscript{29,70}

Successful engagement in the MRC approach also requires a significant investment of resources, which is more problematic in simulation research than in other fields where the MRC framework has been successfully used (eg, cancer care health services). Underfunding\textsuperscript{5,20} has limited programmatic research in health care simulation to date. As a result, investigators are often forced to use short, infrequent training sessions attached to existing educational programs targeting small cadres of learners.\textsuperscript{20} In the present research climate, it may be difficult to conduct sequential pilot studies that manipulate various features in simulation interventions, secure captive learner populations for outcome and process evaluation, or capture long-term effects once SBE interventions have been implemented. Such studies are invariably longer, require larger pools of subjects, and often require multidisciplinary teams with expertise in a range of research methods; without adequate funding, this research simply may not be possible. However, as Gaba\textsuperscript{6} states, “we must take the long view and believe that rational thinking about the key themes and questions about simulation in health care will find a constructive policy and funding climate somewhere, now or in the future.” Strides are already being made to advance the health care simulation agenda at a policy level\textsuperscript{20}; thus, although these constraints may limit short-term progress, as a community, we should position ourselves to engage in the kind of research that advances our field, so that when change occurs, we can capitalize on it.

One strategy that may assist simulation researchers, until a more consistent programmatic funding source becomes available, is to target specific funding sources for different phases of the research program. Preliminary research focused on modeling interventions, testing feasibility and acceptability with stakeholders, generating validity evidence of assessments, and small-scale pilot work may be suitable for short-term seed funding (eg, hospital educational research or innovation funds), grants from professional governing bodies and societies (eg, CSERT Grant from the Association for Surgical Education), and philanthropic sources (eg, Laerdal

![FIGURE 2. This figure demonstrates the relationship between the phases of the MRC framework and the TSR model. As can be seen, the MRC provides detailed guidance in the early phases of developing, refining, and evaluating a complex intervention in simulation (MRC phases 1–3), which would commonly be classified as T1 research in the TSR model. Conversely, phase 4 of the MRC model (evaluation) encompasses T2 and T3 level research, where the TSR model provides more specific guidance. When viewed in this way, both of these models provide overlapping but complimentary guidance for health care simulation researchers.](Image)
Foundation Awards). With the demonstration of these early successes, iterative growth of research questions, and formation of research networks, investigators may be more persuasive with granting bodies in securing sustainable long-term funding. For example, the Canadian Institute of Health Research has a history of funding well-articulated, educationally focused, multiyear research programs with demonstrated early successes and established assessment metrics. Tying larger-scale evaluation and implementation studies directly to translational research outcomes may also help attract funding from national bodies such as the Agency for Healthcare Research and Quality.

Returning to the analogy of bricks and buildings, there are different costs associated with making bricks and constructing buildings. Intuitively, those in the business of funding building construction will be hesitant to invest if there is limited evidence of purposefully made bricks and other building materials necessary for such an undertaking. Similarly, those in the brickmaking business would be hesitant to invest in making bricks without appropriate specifications for their use. Although both investments would fund unique aspects of the ultimate construction, an overarching plan (using methods like the Decadal Survey and STS) is needed to justify both investments. The proposed adaptation of the MRC framework may support such an overarching plan, which in turn may aid researchers in securing both seed and large-scale programmatic funding to advance health care simulation scholarship.

CONCLUSION

The framework proposed in this article calls for simulation researchers to make a change in their approach to scholarship. By taking the time to engage in a theory-based, iterative, programmatic approach such as the one described in this framework, investigators will invariably move away from individual research projects and toward programs of research composed of a number of individual research studies that challenge, refine, and build on existing knowledge over time. Following on the analogy of Forscher, the proposed paradigm shift and its acceptance by simulation researchers, educators, and publishers would ensure that as a simulation research community, we do not produce a plethora of ill-fitting bricks but bricks with the specific shape and color necessary to construct buildings that stand for generations to come. As health care simulation research matures, should we aspire for anything less?

ACKNOWLEDGMENTS

The authors thank Hide Miyagawa for his assistance in creating Figure 1 and Rabia Khan for her suggestions on improving the presentation of the framework. The authors also wish to thank the peer reviewers and journal editorial board for their constructive feedback, which led to critical improvements to the manuscript.

REFERENCES
