Knowing How and Knowing Why: Integrating Conceptual Knowledge in Simulation-Based Procedural Skills Training to Support Learning Transfer

by

Jeffrey J.H. Cheung

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Institute of Medical Science
University of Toronto

© Copyright by Jeffrey J.H. Cheung 2019

Knowing How and Knowing Why: Integrating Conceptual Knowledge in Simulation-Based Procedural Skills Training to Support Learning Transfer

Jeffrey J.H. Cheung

Doctor of Philosophy

Institute of Medical Science
University of Toronto

2019

Abstract

In health professions education (HPE), simulation technologies are used to imitate aspects of real-life clinical skills and scenarios with goal of preparing clinical professionals for real-world practice. Though much effort has gone into incorporating simulation into HPE curricula, questions remain about simulation-based instruction can be designed to support trainees' ability to transfer skills learned in simulation to novel contexts. Instructional designs that promote trainees' conceptual understanding have shown to enhance their ability to transfer across a variety of domains but has yet to be operationalized the context of healthcare simulation.

In this dissertation, I investigate the relationship between instructional design, conceptual understanding, and skill retention and transfer in the context of simulation-based procedural skills training with novice medical trainees. Building on research in education, cognitive psychology, and clinical reasoning, this work assesses the impact of instructional designs that seek to integrate two types of knowledge that underlie procedural flexibility and expertise: conceptual knowledge (i.e., knowing why) and procedural knowledge (i.e., knowing how). In three randomized, controlled experiments, I test the cognitive mechanisms of effective integration of conceptual and procedural knowledge by manipulating the availability and presentation of these knowledges in instructional material.

Results show the integrated instruction increases in trainees' conceptual knowledge, which in turn mediates improvements in trainees' skill retention and transfer. Study 1 establishes the benefits of video-based instruction that integrates conceptual knowledge (in addition to procedural knowledge) for trainee's skill retention and transfer. Study 2 demonstrates that video-based integration is most effective when it encourages trainees to create causal linkages between procedural and conceptual knowledges at the level of cognition, a process called *cognitive integration*. In Study 3, we replicate previous findings and attempted to bolster cognitive integration by designing simulators that make the causal relationships between procedural and conceptual knowledge visible and interactive.

Taken together, the dissertation operationalizes cognitive integration in simulation-based procedural skills training and provides evidence that integrated instruction can enhance simulation-based skill retention and transfer. Hence, lower fidelity training that emphasizes developing an integrated understanding of procedural and conceptual knowledge may be superior to repeated practice with more realistic simulations.

Acknowledgments

Though it is perhaps cliché to say, the process of completing a PhD has been transformative. It has been a life changing journey of highs and lows; but each crest and trough has been accompanied by support, mentorship, and guidance from those who have come before and those on academic journeys of their own. As I close this chapter of my life and begin the next, I would like to extend my deepest gratitude to all who have been part of this journey.

To my committee members, thank you for your generosity in sharing your time and wisdom. It has been an honour and privilege working with you all. I look forward to continuing our collaborations in the years to come! To Ryan Brydges, my supervisor – it has been an exciting process learning from and with you. Thank you for your openness and support throughout the PhD. Your willingness to share your vulnerabilities has helped me get through many a difficult time. Working with you has helped me become a better academic and human being. To Mahan Kulasegaram, you have been an example of the kind of researcher and mentor I hope to one day become. Thanks for always having your door open for impromptu questions and conversations, and for connecting everyone with academic opportunities. Someday I will figure out what Science is and tell you all about it. To Nikki Woods, thank you for your fearlessness in pushing my thinking and understanding. Your candor and support (despite what your facial expresses may convey) have been a great source of strength through the years. To Carol-anne Moulton, thanks for being there when it counts and sharing your tremendous wisdom and intellect.

To the Wilson Centre community – it was here that I was first titillated by the endless ways of viewing and studying the world of healthcare and education. In the past 6 years, what I've learned through my interactions with all the scientists, fellows, visitors, and staff make me truly feel like a 'child' of the centre. Thank you all for fostering and supporting my boundless curiosity and most of all for your kindness and warmth. Though life may lead me to other places, I will always consider The Wilson Centre my first intellectual home. To Mariana, I will never forget your hugs and welcoming attitude to all who walk through the door. To Doug, thank you for tolerating my presence as you taught me how to better use pictures and words to tell a story. To the scientists and researchers: Maria, Tina, Vicki, Walter, Mathieu, Stella, Vicki, Geoff, Laura, Rodrigo, Elise, Shiphra, Ayelet, Nancy, Cynthia, Brian, Paula, Laura, Sarah, Catharine, & Niall – thank you for sharing your wisdom over the years in both formal and informal settings; I

have become the person I am today because of your teachings and support. To my Wilson Centre contemporaries: David, Julian, Jean-Marie, Kinga, Rob, Cristian, Faizal, Cheryl, Aaron, Gianni, Naomi, Jamie, Priyanka, Natasha, Kristina, Ahmed, Brandon, Dorotea, Mike, Jon, & Chris - I will miss our days of morning coffee runs, intellectual debates, constant learning, and overall mischievousness! To the Danes: AnnSofia, Flemming, Mikael, Kirsten, Christine, Charlotte, Ebbe, and Andreas, – it was so nice to spend time with you all during your times at the Wilson Centre.

To my loving wife, Victoria – thank you for your support and understanding. You've always been there to calmly listen to my intellectual rants and have tolerated me changing my mind about 50 times before finally deciding on something. I don't know how you do it sometimes! You've brought balance into my life and given me a future I look forward to building together with you in years to come. Love always.

Contributions

Jeffrey Cheung (author) prepared this dissertation. JC led each of the projects from inception, execution, analysis, and writing of original research articles. The contents of this dissertation include three original manuscripts, with co-author contributions described below each.

1. Cheung, J. J. H., Kulasegaram, K. M., Woods, N. N., Moulton, C., Ringsted, C. V., & Brydges, R. (2018). Knowing How and Knowing Why: testing the effect of instruction designed for cognitive integration on procedural skills transfer. *Advances in Health Sciences Education*, 23(1), 61–74. doi:10.1007/s10459-017-9774-1

JC designed the experiment in the study, completed all data collection and analysis, and led manuscript writing and journal submission process. Guidance for each project was provided by PhD co-supervisors Dr. Ryan Brydges, Dr. Charlotte V. Ringsted, and Dr. Carol-anne Moulton, as well as PhD thesis committee members, Dr. Kulamakan Mahan Kulasegaram, and Dr. Nicole Woods.

2. Cheung, J. J. H., Kulasegaram, K. M., Woods, N. N., & Brydges, R. (2019). Why Content and Cognition Matter: Integrating Conceptual Knowledge to Support Simulation-Based Procedural Skills Transfer. *Journal of General Internal Medicine*. doi:10.1007/s11606-019-04959-y

JC designed the experiment in the study, completed all data collection and analysis, and led manuscript writing and journal submission process. Guidance for each project was provided by PhD co-supervisors Dr. Ryan Brydges, Dr. Carol-anne Moulton, as well as PhD thesis committee members, Dr. Kulamakan Mahan Kulasegaram, and Dr. Nicole Woods.

3. Cheung, J. J. H., Kulasegaram, K. M., Woods, N. N., & Brydges, R. Material concepts: exploring simulation as a medium to enhance cognitive integration and skill transfer. Unpublished manuscript.

JC designed the experiment in the study. Snezhana Besonova served as the project research assistant completed all data collection. JC led the manuscript writing and journal submission process. Guidance for each project was provided by PhD co-supervisors Dr. Ryan Brydges, and

Dr. Carol-anne Moulton, as well as PhD thesis committee members, Dr. Kulamakan Mahan Kulasegaram, and Dr. Nicole Woods.

Table of Contents

List of Abbreviations

ANOVA – Analysis of Variance

ANCOVA – Analysis of Covariance

GRS – Global Rating Scale

HPE – Health Professions Education

LP – Lumbar Puncture

M – Mean

PFL – Preparation for Future Learning

SD – Standard Deviation

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction and Literature Review

1.1 General Introduction and Overview

In its broadest interpretation, simulation can be defined as a technique that imitates some aspect of reality (Gaba, 2004; Rosen, 2008). Simulation may conjure images of futuristic ‘Matrix’-like technologies that create alternate or augmented realities for humans to experience, but on a more fundamental level, simulation can be viewed as an educational strategy with the ultimate goal of preparing trainees for reality. Beyond replicating reality, simulation educators aim to influence what trainees learn, and how they perceive and behave in future situations.

Though simulation has experienced incredible uptake in the last 30 years, there is much left to be understood (Gaba, 2011). As with other areas of research in health professions education (HPE), research practices have lacked theories or conceptual frameworks to guide programs of inquiry (Cook, Bordage, & Schmidt, 2008; Haji, 2015; Hodges & Kuper, 2012; Norman, 1996, 2004). In setting a research agenda for the field, leaders in simulation have called for research on simulation instructional design that aims to clarify what about simulation works, for whom, and under which circumstances (Cook, 2010; Cook et al., 2011, 2012; Cook, Hamstra, et al., 2013). In these calls, a priority area for research is the instructional design of simulation; specifically, the impact of different instructional design features on simulation-based learning outcomes (Dieckmann et al., 2011; Issenberg, Ringsted, Østergaard, & Dieckmann, 2011).

To address these gaps in the simulation instructional design literature and to advance the science of HPE, this dissertation draws on theories from cognitive psychology and education to inform the study of simulation instructional design. The studies contained in this dissertation focus on clarifying the relationship between the types of knowledge that trainees acquire from the instructional content of simulation (e.g., anatomy, equipment, steps of a skill, patient safety), how this content is organized (e.g., anatomy first, followed by steps of the skill), and trainees’ learning outcomes from simulation.

The main thesis of this dissertation is that simulation-based training will be most effective when instruction imparts upon trainees knowledge pertaining to *how* to perform a skill along with knowledge pertaining to *why* the skill is performed in the prescribed manner and sequence. Put another way, clinical expertise for skills taught in simulation will develop from integrated instruction, which aims to combine procedural knowledge (knowing *how*) and conceptual knowledge (knowing *why*). The primary phenomenon of interest is the influence of integrated instruction on trainees' conceptual knowledge and further, the relationship between trainees' conceptual knowledge and their ability to apply their simulation-based procedural skills to novel contexts, an ability we define in this dissertation as *transfer* (Goldstone & Son, 2005).

This dissertation presents three independent studies in self-contained chapters (Chapters 3 – 5). The studies were conducted iteratively and share common research goals, questions, and methodology (Chapter 2). These three studies are preceded by the current introduction (Chapter 1) and followed by a general discussion of the findings (Chapter 6), conclusions (Chapter 7) and the future directions and limitations of this program of work (Chapter 8). This first chapter provides an overview and simulation in healthcare, synthesizes literature on how procedural and conceptual knowledge are defined and how they relate trainees' expertise development, describes strategies for integrating knowledge that have been applied in clinical reasoning, and discusses the learning outcomes of retention and transfer as operationalized in this thesis.

1.2 Simulation in Health Professions Education

1.2.1 The Proposition of Simulation for Health Professions Education

The impetus for the modern adoption of simulation in HPE arose due to perceived shortcomings in the traditional apprenticeship model of training. First introduced in 1889 at John Hopkins Hospital, the traditional Halstedian model of training is a time-based model of training, which requires intensive apprenticeship and exposure to a large quantity and wide range of clinical experiences (Carter, 1952). However, given the increasing cost of healthcare, restrictions to resident duty hours, increasing specialty sub-specialization, and public concerns about medical errors - the efficiency and safety of immediate health care delivery have become paramount - time and resources for health care education have become increasingly limited (Aggarwal & Darzi, 2006; Moorthy, Munz, Sarker, & Darzi, 2003; Reznick & MacRae, 2006; Ziv, Wolpe, Small, & Glick, 2003). These pressures have brought the educational effectiveness of time-based training models into question. Can we continue to rely on apprenticeship to provide trainees with the necessary exposure to a suitable representation of what they will encounter in independent professional practice? Are there enough training opportunities for trainees to gain adequate expertise? How can we be certain that all trainees receive comparable training experiences? Is it ethical to have novice trainees 'learn' on patients?

These concerns have bolstered efforts to establish competency-based models for training health professionals (Frank et al., 2010; Hodges, 2010; Hodges & Lingard, 2012; Leung, 2002; Long, 2000) and, in tandem, propositions for simulation to meet unmet training and assessment needs (Issenberg, 2006; Issenberg & Scalese, 2008; McGaghie, Issenberg, Barsuk, & Wayne, 2014; McGaghie, Issenberg, Petrusa, & Scalese, 2016; Motola, Devine, Chung, Sullivan, & Issenberg, 2013; Scalese, Obeso, & Issenberg, 2008). Unlike the traditional apprenticeship model of training, where clinical expertise is inferred from a trainees' time-spent in a program, competency-based models of training use predefined outcomes of knowledge, skills, and attitudes to determine when to deem trainees ready for independent practice (Frank et al., 2010; Holmboe, 2014). In contrast to the opportunistic nature of learning inherent in time-based education (Reznick & MacRae, 2006), simulation offers a means to standardize the training and assessment experiences that are necessary in competency-based models of education. Rather than learning being dependent on factors such as demographics of the local patient population, trainees' and educators' physical location within a hospital (e.g., operating room versus emergency room), time of day, a patient's comorbidities, and other clinical work pressures or concerns; educators can leverage simulation to provide structured (simulated) clinical experiences that align with a learners' needs and requisite competencies (Gaba, 2004; Issenberg et al., 1999; Ziv et al., 2003). Simulation is seen as a vital component in the shift toward competency-based education for the health professions.

Despite the potential for simulation to enhance HPE as it transitions to competency-based education models, the educational value of simulation is unclear. In many cases, efforts to implement simulation into curricula have resulted in purchases of costly simulators that are eventually left to collect dust after the initial enthusiasm of using a new educational tool wanes (Issenberg, 2006). In addition to the cost of purchasing and maintaining simulation equipment, adopting simulation into a training program requires space, support staff, and time commitments from both learners and educators. The potentially high costs associated with adopting simulation necessitates an evaluation of the "value proposition" of simulation, both economically and otherwise (Cook, Andersen, Combes, Feldman, & Sachdeva, 2018; Zendejas, Wang, Brydges, Hamstra, & Cook, 2013). Though there is utility in considering concepts such as opportunity cost - what we miss out on when we invest resources in simulation that could otherwise have been put towards other resources – establishing a consensus on what constitutes value or how to measure it has proven difficult (Lin, Cheng, Hecker, Grant, & Currie, 2018; Nestel, Brazil, & Hay, 2018). Evaluating the value of simulation presents some difficult questions. For example, given a particular clinical skill and clinical context, is simulation the most effective approach to teach it? What are the most appropriate measures to capture this? Money? Time? Educational outcomes? Which educational outcomes? What alternative educational interventions should we be comparing simulation-based education against? Further, any effort to establish a value proposition for simulation is muddled by the currently poor understanding of what elements of healthcare simulation enable it to be effective as an educational technique, and thus, how it may be applied to effectively achieve educational goals, even once these are defined.

In the age of competency-based education, simulation holds much promise, yet to fulfill its potential, much work remains in clarifying what elements of simulation result in improved educational outcomes. To-date, rather than from evidence of its educational benefits, the prevalence of simulation in HPE stems more so from an ethical imperative to ensure HPE does not directly harm patients, which is rooted in calls for greater public awareness and professional accountability (Ziv et al., 2003). Though simulation has become an integral component of North American post-graduate training programs, how it can be most effectively incorporated to enhance HPE outcomes is unclear and challenging to evaluate. However, before the proposition of simulation to improve HPE can be adequately evaluated, educators and researchers must first understand how different uses of healthcare simulation relate to important educational outcomes.

1.2.2 How Does Simulation Work? Evidence Supporting Simulation

Before addressing how simulation can improve HPE or its value proposition as compared to other modes of education, an obvious first question to address is whether simulation "works" at all. Namely, does it improve trainees' clinical knowledge, skill, or behavior?

Multiple reviews of the healthcare simulation literature have identified the potential of simulation to improve learning (Cook, Hamstra, et al., 2013; Cook et al., 2011; Issenberg, Mcgaghie, Petrusa, Lee Gordon, & Scalese, 2005; McGaghie, Issenberg, Petrusa, & Scalese, 2006, 2010). These reviews make a strong case that simulation can improve learning outcomes and have outlined some instructional design features of simulation that may help optimize its use for enhancing learning. Some of the top recommendations include features such as the availability of feedback (either built into the simulator or provided externally via instructors), opportunities for repetitive practice and deliberate practice (practice with the intent of consistently improving performance) (Ericsson, 2004), the integration of simulation within curricula, higher fidelity simulation, and clear learning outcomes (Issenberg et al., 2005; McGaghie et al., 2010).

One review of significance to the question of whether simulation "works" is a quantitative systematic review and meta-analysis from Cook et al., (2011). Importantly, the article quantitatively summarizes a large number of simulation research articles, conclusively demonstrating that simulation does indeed "work". From an initial pool of 10,903 articles, the authors analyzed 609 studies and determined effect sizes for simulation-based interventions compared to no-intervention controls on various learning outcomes. Using Hedges g to report standardized effect sizes (Hedges, 1981), the review found large effect sizes of over a standard deviation ($g > 1$) for knowledge outcomes, time skills (time to complete a simulated task), process skills (measures overall effectiveness of efficiency, such as global ratings), and product skills (measures of procedural success or quality of final products). Further, the review found medium effect sizes ($g > 0.5$) for behaviors in practice with patients (e.g., time to complete procedures and instructor ratings of competence).

While findings from these reviews of the literature support the use of simulation, the same authors have also identified significant shortcomings in the methodological and theoretical rigour of simulation research. In the 2005 review from Issenberg et al., when assessing the quality of the reported outcomes of research articles, only 20% were deemed to be "clear and probably true". Further, in the review from Cook et al., (2011), only 4% of studies failed to show a benefit to using simulation; however, two-thirds of those studies reporting a benefit conducted a single-group pretest-post-test comparison. Though simulation appears to positively impact educational outcomes, the majority of simulation research studies lack comparison groups (single-group studies); comparing simulation with no-intervention controls. That is to say, simulation "works", at least when compared to no intervention. As discussed by these authors, it is no longer informative for research to demonstrate that any form of educational intervention (simulation or otherwise) will improve learning (if we teach them, they will learn!) (Cook, 2012; Cook, Brydges, Zendejas, Hamstra, & Hatala, 2013).

The strongest evidence demonstrating simulation's effectiveness comes from studies of simulation interventions that have adopted the principles of deliberate practice (McGaghie et al., 2014, 2016). Unlike traditional approaches to skill acquisition that assume adequate time spent practicing in a domain will result in expert performance, the deliberate practice approach argues that expert performance requires, as its name implies – *deliberate* practice (Ericsson, 2006a); which can be defined by: i) focused attention on a well-defined task that is targeted for improvement, ii) immediate feedback on performance, and iii) repeated practice with the intent of deliberately improving performance beyond current skill levels (Ericsson, 2004; Ericsson, Krampe, & Tesch-romer, 1993). Decades of research of expert performance have shown that differences in levels of expert performance can largely be explained by differences in the amounts of deliberate practice individual experts have engaged in (Ericsson, 2006b); a finding shown across a variety of domains of performance, from competitive chess, music, athletics, and medicine.

A series of simulation studies have paired deliberate practice with the concept of mastery learning and have shown impressive educational outcomes. Mastery learning is an outcomes-based approach to training, which uses repetitive practice until trainees meet a predefined level of competency or mastery (McGaghie et al., 2014; McGaghie, Issenberg, Cohen, Barsuk, & Wayne, 2011). Interventions that have adopted simulation-based mastery learning with deliberate practice have resulted in trainees with higher quality performances of bedside procedures (Barsuk, McGaghie, Cohen, Balachandran, & Wayne, 2009; Barsuk, McGaghie, et al., 2009) and trainees who are also better able to retain their skills over prolonged periods of time (Ahya et al., 2012; Barsuk, Cohen, McGaghie, & Wayne, 2010). Further, the benefits of simulation-based mastery learning with deliberate practice have shown to translate to outcomes at the broader institutional level, such as reductions in complications rates that arise from clinical procedures that have been taught using the approach, and reductions the associated costs that these complications would have incurred (Barsuk, Cohen, Feinglass, et al., 2014; Barsuk, Cohen, Feinglass, McGaghie, & Wayne, 2009; Barsuk, Cohen, et al., 2009; Barsuk, Cohen, Potts, et al., 2014; Cohen et al., 2010).

Despite the impressive benefits cited by studies of simulation-based mastery learning with deliberate practice, as with much of the majority of simulation literature, current research remains at the level of description or justification research, and does not help us understand how or why simulation has these positive effect (Cook et al., 2008). That is, without comparing deliberate practice approaches with other active educational interventions, this body of work fails to clarify the mechanisms that mediate simulation's benefits. For example, a systematic review of mastery learning found that of the only 3 studies that compared simulation-based mastery learning approaches to non-mastery simulation-based training, though there was improved learning, mastery learning was also associated with a greater amount of time spent in simulation training (Cook, Brydges, et al., 2013). Hence, from the current studies available, it is not clear whether the benefits of mastery learning with deliberate practice result from these instructional features themselves, or simply greater time spent practicing. And thus, it might be that these benefits would no longer be apparent after controlling for the additional time and resources such interventions appear to require.

To advance the field of simulation in healthcare, studies must pit different theory-informed simulation-based instruction in head-to-head comparisons, allowing researchers to elucidate the impact of various instructional design features of simulation on educational outcomes. (Cook, 2010; Cook, Hamstra, et al., 2013; Cook et al., 2011). Though reviews of the literature have identified some evidence for various instructional design features, much is left to be clarified given the weak evidence base in support of them. For example, although the literature suggests feedback is an important instructional design feature for effective simulation, the current literature provides little guidance for best educational practices regarding the type, timing, content, or scheduling of feedback, and how these factors might impact educational outcomes (Cheng et al., 2014; Cook, Hamstra, et al., 2013; McGaghie et al., 2010). The primary finding from the research to date appears to be that greater time and effort will enhance educational outcomes, with relatively less information about the effectiveness of the different strategies to achieve them or the factors that drive them to be effective. The more time spent engaged in training activities in the pursuit of expertise, via simulation or other modalities, the more expertise will develop (Cook, 2012; Ericsson, 2006b). Though promising, further research is needed to clarify the mechanisms of learning that instructional design features such as mastery learning with deliberate practice support so that it can be compared to other pedagogical approaches in HPE.

Akin to calls for better research quality in HPE more broadly, simulation research must focus on producing evidence clarifying *why* simulation works, for whom, and under what circumstances (Bordage, 2009; Cook, 2012; Cook & Beckman, 2008; Cook et al., 2008; Hodges & Kuper, 2012; Norman, 2004); to quote McGaghie et al., (2016):

> *"There is no longer any doubt that simulation-based medical education (SBME) can be a powerful educational intervention when it is used under the right conditions. The challenge for the medical education research community is to figure out how to use SBME efficiently and cost-effectively to educate and evaluate individual doctors and health care teams."*

1.2.3 Technology versus Technique - Fidelity, Realism, and Transfer

Though the reproduction of some aspect of reality is a defining feature of simulation (Gaba, 2004; Rosen, 2008), the ultimate goal of simulation is trainee's ability to transfer learning acquired from simulation into reality (Hamstra, Brydges, Hatala, Zendejas, & Cook, 2014). This notion of transfer is poorly understood and understudied in the context of simulation instructional design, where the majority of comparative research studies have examined educational outcomes using pretest-post-test designs or delayed retention tests (Cook, Hamstra, et al., 2013) Unlike transfer, which assess trainees' ability to adapt previously learned knowledge and skills (Goldstone & Son, 2005; Salomon & Perkins, 2015), pretest post-test and delayed retention test designs assess knowledge recall or skill reproduction either immediately after training or at a delay (e.g., one-week later). These constructs of skill reproduction and application represent fundamentally different educational purposes (Broudy, 1977), as well as cognitive processes, which suggests that different instructional designs and assessments may be better suited in developing and measuring them. For example, increasing the challenge of an intervention by mixing practice of various to-be-learned concepts can depress immediate test performance (e.g., a post-test delivered immediately after a learning session), and yet support delayed test performance (e.g., a retention test delivered a week after the same learning session) (Bjork, 1994; Dubrowski, 2005; Schmidt & Bjork, 1992a). In the pursuit of improved skill transfer, there is similar evidence that instructional design features can differentially affect trainees' performances on tests of skill reproduction and transfer (Holyoak & Thagard, 1997; Kapur, 2016; Salomon & Perkins, 1989). Hence, by appreciating the theoretical distinctions between transfer and retention, we also appreciate that instructional design features that support these two constructs, and the assessments that capture them, may differ from what the current reviews of simulation suggest will be effective for learning.

Unfortunately for simulation science, rather than clarify how to optimize whether simulation training produces skill transfer, much work has focused on developing and advancing simulation technology, in order to better reproduce reality. The past twenty years has seen an exponential growth in the number of publications on high-fidelity simulation (Issenberg et al., 2005; McGaghie et al., 2016), which can be loosely defined as simulations that reproduce reality with a high-degree of authenticity or realism (Norman, Dore, & Grierson, 2012). These efforts have resulted in numerous simulators developed for a wide range of clinical skills, utilizing various technologies and modalities for educational delivery. These simulators can vary in their sophistication, ranging from inanimate benchtop task trainers, designed to allow individual trainees to practice technical skills (e.g., suturing or venipuncture), to fully immersive scenario-based simulations using either virtual reality or full-body patient mannequins that simulate breathing, bodily fluids, or even blinking, designed to allow for practice of team-based communication and other cognitive skills across a myriad of clinical scenarios (e.g., deteriorating patient / crisis resource management scenarios) (Issenberg & Scalese, 2008; McGaghie et al., 2010). Despite these advances in the technology of simulation and the expanding list of features that simulators wield, this focus on technology and innovation though fruitful, has created a barrier to understanding and optimizing simulation instructional design. For example, what is meant when authors denote a simulator of 'high-fidelity' as opposed to 'low-fidelity' lacks established consensus. Advances in simulation technology have outpaced advances our ability to discuss the features of simulation with precision and our understanding of how these features work.

One highly-cited framework for simulation realism suggests different modes of thinking about reality can help educators design simulation scenarios to achieve educational outcomes (Dieckmann, Gaba, & Rall, 2007). This model proposes three modes of thinking required to fully describe and specify a given situation: physical, semantical, and phenomenal. The physical mode considers the functional material features of a simulation (e.g., shape, weight, size, texture, movement, sound). The semantical mode considers the conceptual meaning that underlies the sensory information a simulation presents (e.g., recognizing that the water within a syringe represents an anesthetic and will have the intended effect on the simulated patient). The phenomenal mode considers the emotional experience that accompany simulation (e.g., feeling the simulation scenario has relevance to real clinical situations). The authors of this framework suggest that these three modes of thinking about reality can be used to describe and design simulation scenarios that align with the goal increasing participants' willingness to "suspend disbelief".

According to this framework, buy-in into the simulation as facilitated by physical, semantical, and phenomenal realism, determines whether a simulation "works". Greater alignment of simulation realism with educational purposes of simulation can encourage trainees to more willingly treating the simulation "as-if" it was a clinical situation. If the goal of simulation is to teach procedural steps of a skill, it would be important to have physical realism. If the goal of the simulation scenario is to teach diagnosis and treatment planning, semantical realism of the pupillary light reflex through a verbal announcement may be more important than having simulator that exhibits a realistic pupillary light reflex. Similarly, if the goal of the simulation scenario is emotional engagement with a patient, phenomenal realism of the simulated patient's voice when describing their own symptoms may be paramount. By titrating each of these modes of realism, it is suggested that trainees will have greater engagement in simulation and thus, simulations will ‘work’ (Rudolph, Simon, & Raemer, 2007).

Though this framework of realism proposes hypotheses about how realism can promote trainee engagement in simulation when viewed as a complex social practice, many empirical questions remain about whether, and under which circumstances, increased engagement and immersion will result in improved learning outcomes. For example, does increasing physical realism increase engagement the same way increasing semantical or phenomenal realism? Can these types of realism be independently manipulated? How can their impacts be measured? Other than engagement or motivation, are there other potential learning mechanisms that increasing realism acts upon?

Further, despite the purported importance of aligning simulation realism to its educational purposes, the framework also suggests that in some cases “it might be beneficial or even necessary to purposefully depart from realism to provide the most effective training” (Dieckmann et al., 2007). The authors provide examples of speeding up or slowing down physiological responses, providing hints, feedback, or even pausing the entire scenario to allow repeated practice. Though the authors refer to such manipulations as creating a ‘hyper-reality’, it more likely alludes to learning mechanisms beyond trainee engagement enabled through simulation realism. For example, repetitive practice (Cook, Brydges, et al., 2013), feedback (Hattie & Yates, 2013), reflection and elaboration (Reigeluth & Carr-Chellman, 2009). Though an important step, description alone is insufficient in advance of more effective healthcare education – empirical research must also clarify how changes these aspects of realism impact education outcomes, and further justify why realism and engagement should take precedence over learning outcomes such as retention and transfer.

A challenge for simulation research in clarifying the active instructional design features that support trainees skill transfer, is the allure of simulation technology, and along with it, the assumption that greater fidelity to reality will yield better transfer to reality (Norman, Dore, & Grierson, 2012). Though on the surface, just as simulation “works” for improving educational outcomes described in the previous section of this chapter, current evidence suggests simulation-based training (of the high-fidelity variety) improves transfer outcomes. Simulation has shown to facilitate transfer of learning to real patients (Brydges, Hatala, Zendejas, Erwin, & Cook, 2015;

Teteris, Fraser, Wright, & McLaughlin, 2012) and other "translational" outcomes to other aspects of health care systems (e.g., reduced infection rates) (Barsuk, Cohen, et al., 2009; McGaghie, Draycott, Dunn, Lopez, & Stefanidis, 2011; McGaghie et al., 2014). However, few research studies have compared the effectiveness of different instructional design features of simulation on transfer outcomes (Cook, Hamstra, et al., 2013), and where comparisons of high and low fidelity modalities of simulations have been made, the relationship between fidelity and skill transfer is not apparent (Norman, Dore, & Grierson, 2012). This finding has led some to question whether simulation, when viewed only as a modality for delivering realistic educational content, will ever improve learning (i.e., skill retention and transfer), particularly if educators do not also consider the learning mechanisms simulation affords trainees (Artino & Durning, 2012; Hamstra et al., 2014).

Put more generally, a key challenge to advancing simulation research may be the tendency to view simulation as a modality for educational delivery, without also considering the content being delivered. This focus on technology and technique, obfuscates the underlying knowledge and skills that are packaged within simulation-based training. Thus, rather than using robust research methodology to clarify how simulation can complement and build upon the knowledge and skills trainees acquire from the classroom and clinical settings (i.e., educational content), comparative research in simulation has predominantly explored ways of more efficiently delivering content through simulation. For example, research studies have considered instructional features of simulation such as practice scheduling (distributed practice) (Moulton et al., 2006), repetitive practice (Cook, Brydges, et al., 2013; McGaghie et al., 2006), variability of practice (Brydges, Carnahan, Rose, Rose, & Dubrowski, 2010; Dubrowski, Park, Moulton, Larmer, & MacRae, 2007), complexity of practice (Haji, Cheung, Woods, Regehr, deRibaupierre, et al., 2016; Van Merriënboer & Sweller, 2010), and feedback on practice (Boyle et al., 2011; Porte, Xeroulis, Reznick, & Dubrowski, 2007). To use an analogy, simulation is operationalized in education as if it were a ship that carries precious educational content into the minds of learners, with little consideration for the nature of the cargo itself. Unfortunately for ship builders and sailors – educators in this analogy – a ship that can effectively deliver one type of cargo may not be appropriate for delivering another type of cargo.

Though technological developments in the delivery of simulation-based education represent significant contributions to HPE, for simulation to realize its full potential, simulation science must clarify how simulation instructional design relates to the knowledge and skills that enable effective skill and transfer. The path forward requires theoretically informed and methodologically robust evidence that moves beyond comparing something versus nothing, and also accounts for both *what* is being learned in simulation and *how* we structure this learning through the modality of simulation.

1.3 Procedural Knowledge and Conceptual Knowledge: The Ingredients of Adaptive Expertise

1.3.1 The pursuit of expert performance

Though greater expertise can be associated with more effective skill transfer, this positive relationship may not always be found. In the health professions, expertise is more often based on peer nomination and time spent within the domain than evaluations of an expert's ability to demonstrate skill transfer. Thus, most graduating trainees will be considered "experts" at some point in time after their formal training and accreditation as a health professional. As such, the question of whether health professions trainees will become experts or not is moot. The more important question for education researchers and society is just what kind of experts they will become. As put by Bereiter & Scardamalia (1993):

> *"Eventually they will quit being novices, without our having to do anything about it. The important question is what they will become. Will they become experts in their lines of work or will they swell the ranks of incompetent or mediocre functionaries?"*

Expertise development is a central goal of competency-based education frameworks, which aim to ensure not only a basic level of competence in graduate health professions trainees, but excellence and adaptability. In addition to the promise of excellence is the promise of improved training efficiency; adopting a competency-based approach to education brings with it a potential hope for reduce training times, by focusing trainees' educational experiences on lists of to-be-achieved competencies, rather than allowing time to be the sole determinant of competence (Hodges, 2010). These aspirations of producing better experts in less time have been met with skepticism from the researcher community, who challenge the supporting evidence base that informs these projections (Boyd et al., 2018; Norman, Norcini, & Bordage, 2014; Whitehead, Kuper, & Webster, 2012; Whitehead, Austin, & Hodges, 2012; Whitehead & Kuper, 2017).

Instead, decades of research on expertise show that it arises as the result of years of vigorous study and effortful practice (Ericsson, 2004; Norman, Eva, Brooks, & Hamstra, 2006). However, owing to the constraints of time and other resources, voluminous practice on its own (deliberate or otherwise) is an inadequate strategy to prepare trainees with all the experience they will need in their future professional careers. Put another way, unlike the steps to a dance that can be known prior to a performance and repeatedly rehearsed until mastery, health care delivery is more complex, variable, and its steps are not entirely knowable. Addressing how clinical expertise can be developed more effectively and efficiently requires research about how instructional designs enhance learning, beyond prescribing more practice and effort. To this end, some have argued that HPE be designed to teach trainees clinical knowledge and skills in a manner that enables their *application* and *re-interpretation* when they inevitably encounter novel clinical problems or situations (Broudy, 1977; Mylopoulos, Brydges, Woods, Manzone, & Schwartz, 2016; Schwartz, Chase, & Bransford, 2012).

1.3.2 Types of Expertise and Requisite Knowledge

Not all experts are created equal. Though most experts can efficiently manage routine problems encountered in the everyday practice in their domain of expertise, not all are able to innovate upon previously learned strategies to address challenging novel problems. These two courses of expertise can be defined as routine expertise and adaptive expertise, respectively (Hatano & Inagaki, 1986). Individuals on both courses of expertise differ in their ability to adapt because of differences in their processes of cognitive development that arise from differences in educational experiences. Namely, the development of routine expertise requires practice of the procedural knowledge required to efficiently execute a given procedure; whereas the development of adaptive expertise requires a comprehension of the nature of the procedure alongside procedural knowledge, which can be referred to as conceptual knowledge (Baroody, Feil, & Johnson, 2007; de Jong & Ferguson-Hessler, 1996; Star, 2005). Procedural knowledge pertains to knowing *how* to perform procedural steps of a skill to effectively and efficiently perform it, whereas conceptual knowledge pertains to knowing *why* these steps are performed in the prescribed manner and order. Educational environments that equip trainees with procedural and conceptual knowledge prepare trainees to demonstrate both routine and adaptive expertise.

The procedural and conceptual knowledge required for routine and adaptive expertise are obtained through varying degrees of effort and instructional designs. The body of procedural knowledge consistent with routine expertise is readily acquired with little difficulty through repeated practice in reoccurring environments (Hatano & Inagaki, 1986). Participation in these environments and the common problems experienced within them are often adequate educational experiences for individuals to become proficient in performing relevant procedural tasks, all without a need for conceptual understanding. In contrast, the body of conceptual knowledge consistent with adaptive expertise depends on experiences that highlight the inter-relatedness of procedural skills (Hatano & Inagaki, 1986; Karpov & Bransford, 1995). Individuals must be able to create meaning to each step of a skill and have a rationale for the selection of each over possible alternatives. This necessitates variability in experiences of a skill, but also an understanding of the relationship between variants of a procedure – which in turn fosters procedural flexibility in trainees, that is, the ability to adapt procedural knowledge to novel situations and invent new problem-solving procedures (based on old ones) when confronted with never before seen problems (Holyoak, 1991).

1.3.3 Relating Procedural and Conceptual Knowledge

"If meaningful knowledge is the basis of adaptive expertise (the flexible application of knowledge) and can be characterized as well-connected knowledge, then the construction of well-connected knowledge should be the basis for fostering adaptive expertise or flexibility, as measured by transfer. Analogously, the acquisition of unconnected knowledge should be the basis for promoting routine expertise or inflexibility."

- Baroody, 2003

Though it is generally recognized that both procedural and conceptual knowledge are necessary components for adaptive expertise, how these two types of knowledge should be related to one another is less conclusive. Is it better to start with concrete examples of problems and teach the problem-solving procedures to address them? Or is it better for learners to start from instruction that explains the theoretical concepts that underlie the procedures? Or should they be taught together in an integrated approach?

Tension between procedural versus conceptual instruction originate from differing philosophies of educational goals, outcomes, and thus instruction. This includes debates about the primacy of empirical learning versus theoretical learning (Karpov & Bransford, 1995), skills practice versus knowledge of concepts (Anderson, Fincham, & Douglass, 1997; Baroody, 2003; Holyoak & Thagard, 1997; Novick & Holyoak, 1991), and discovery learning approaches versus direct instruction (Bransford & Schwartz, 1999; Kapur, 2016; Kirschner, Sweller, & Clark, 2006; Mayer, 2004; Schwartz, Lindgren, & Lewis, 2009). What each disagreement holds in common is whether understanding will be more effective when trainees gain experience by working through a class of problems (empirical learning, skills/rote learning, discovery learning) or through explicit teaching of the general methods or principles of the essential features of a class of problems (theoretical learning, concepts-first, and direct instruction). Further complexity is added when considering what measure constitutes effective understanding. For example, when gauging the effectiveness of a particular approach to teaching trainees' mathematical multiplication, is it effectiveness in solving similar problems on a mid-term examination a few weeks later? Or perhaps we are interested in the process and efficiency with which learning takes place for the related concept of division?

In general, research findings in education show that trainees' skill transfer can be supported through a variety of instructional strategies that aim to support trainees in developing meaningful understanding of problems, be that through teaching procedural or conceptual knowledge first, or emphasizing one or the other (Baroody & Dowker, 2003; Chi & VanLehn, 2012; Rittle-Johnson & Schneider, 2015). Studies in mathematics education suggest that improvements in both knowledges are strongly correlated, where developments in one also yield improvements in the other; that is, improving trainees' procedural knowledge will also improve trainees' conceptual knowledge, and vice versa (Rittle-Johnson, Schneider, & Star, 2015; Rittle-Johnson, Star, & Durkin, 2009, 2012; Schneider, Rittle-Johnson, & Star, 2011). While encouraging, some argue that beyond determining the instructional organization of procedural and conceptual instruction, researchers must understand the impact such ordering and emphasis has on trainees' cognitive knowledge structures that enable understanding and transfer (Bransford, Brown, & Cocking, 1999).

Findings from cognitive psychology also suggest that transfer is supported through an inter-related understanding of procedural and conceptual knowledge (Chase & Simon, 1973; Chi, Glaser, & Farr, 1988). Research on analogical transfer has repeatedly demonstrated its dependence on trainees' deep structural understanding of problems (Goldstone & Son, 2005; Holyoak & Koh, 1987; Holyoak & Thagard, 1997; Needham & Begg, 1991). Whereas novices can identify the surface features of familiar problems and subsequently solve them (i.e., procedural knowledge), experts can also look beyond the surface to see the deep structural relationships a novel problem has with familiar problems and situations (i.e., conceptual knowledge); thus solving these novel problems by adapting previously acquired knowledge and problem-solving skills (Bédard & Chi, 1992; Chi, Feltovich, & Glaser, 1981; Chi, Glaser, & Farr, 1988; Chi & VanLehn, 2012; Salomon & Perkins, 1989). In this sense, adaptive experts have been prepared to see the “old in the new” in ways that allow them to effectively transfer learning (Bransford & Schwartz, 1999; Mylopoulos et al., 2016).

1.4 Knowledge Integration in Health Professions Education

1.4.1 Integrated Competencies in Health Professions Education

In HPE, the need to foster adaptive expertise and integrate knowledge is reflected in calls to integrate clinical competencies. For example, accreditation bodies consider competent medical trainees as individuals capable of integrating the full spectrum of competencies outlined in outcomes-based training frameworks (Frank et al., 2010). However, trainees' ability to transfer may be limited by curricula that teach and assess competencies as isolated attributes. To overcome this perceived limitation, leaders in the field have called for a focus on how to teach clinicians to integrate multiple competencies (Sklar, 2013). Moreover, renowned clinicians describe deliberately integrating competencies as central to their development of expertise (Mylopoulos, Lohfeld, Norman, Dhaliwal, & Eva, 2012). As an example, take the case of a resident who knows how to perform a lumbar puncture (Medical Expert), yet has always worked with a supervisor who has set up the room and team for her to complete the procedure. In a new situation where she is now unsupervised, she may find herself lost in coordinating the necessary resources (Leader) and involving key team members (Communication, Collaborator) to perform the procedure. She has not yet integrated these key competencies and instead demonstrates each individually with varying degrees of success. Thus, teaching trainees to integrate competencies would presumably improve transfer and combat the compartmentalization of clinical competencies.

The integrated nature of competencies has not gone unnoticed by proponents of competency-based education. Indeed, competency frameworks such as CanMEDS provide structured curricula-level programming to promote integration across the years of medical training (vertical integration) and across courses within curricula (horizontal integration) (Frank et al., 2010). What is missing from these frameworks, however, are recommendations for how to teach trainees to integrate competencies during individual educational sessions.

1.4.2 Integrating Basic Science and Clinical learning

Predating todays calls for the integration of competencies, are calls for the integration of basic science and clinical learning. Since the Flexner report in 1910, basic science has been recognized as a fundamental component of clinical expertise and competence (Finnerty et al., 2010).

Following suit, the 20th century saw medical school curricula strive to achieve stronger integration of basic science and clinical learning. The assumption behind these efforts being that by grounding clinical practice in the basic sciences, there would be higher quality health care borne out from higher quality medical education and thus expertise.

Strategies for integrating basic science and clinical learning have predominantly focused on creating "integrated" curricula (Goldman & Schroth, 2012; Harden, 2000). Two broad strategies for curricular integration are vertical and horizontal integration (Bandiera, Boucher, Neville, Kuper, & Hodges, 2013; Hays, 2013). Vertical integration seeks to connect curricular content across time, whereas horizontal integration seeks to connect curricular content at a given time. For example, vertical integration occurs when educators design a 2nd year biology course to reference content learned in the preceding 1st year biology course; and horizontal integration occurs when content in a 1st year biology course is built with an understanding of the content being taught in the coinciding 1st year chemistry course. Vertical and horizontal integration can occur across various content areas within a curriculum, whereby linkages are created at either the level of programs (e.g., across institutional goals), courses (e.g., scheduling and organization of classes), or sessions (e.g., content selection and sequencing within a teaching session) (Goldman & Schroth, 2012).

1.4.3 The Case for *Cognitive* Integration – Linking Cause and Effect

Rather than a product, curricular integration may more appropriately be thought of as an activity, with the ultimate goal of helping trainees appreciate the connections and relationships between different content areas contained within a curriculum (Case, 1991). Evidence from studies that teach trainees to integrate basic sciences and clinical sciences (two types of foundational knowledge) suggests that integration efforts at the level of curriculum planning must be reinforced through careful design of instructional sessions (i.e., the way content is presented to trainees) (Kulasegaram, Martimianakis, Mylopoulos, Whitehead, & Woods, 2013). Such integrated instruction has been shown to improve trainees' diagnostic reasoning skills by prompting them to create conceptual linkages between the signs and symptoms of a disease (clinical science) and its underlying pathophysiology (basic science), a process termed 'cognitive integration' (Kulasegaram et al., 2013). Specifically, instruction that encourages cognitive integration has been linked to improvements in novice trainees' retention of clinical reasoning skill (Baghdady, Carnahan, Lam, & Woods, 2014a, 2014b, 2013; Baghdady, Pharoah, Regehr, Lam, & Woods, 2009; Woods, Brooks, & Norman, 2005, 2007a; Woods, Neville, et al., 2006), their transfer of learning to novel problems (Woods, Brooks, & Norman, 2007b), and their ability to learn new diagnostic categories (Mylopoulos & Woods, 2014).

Unlike curricular integration, cognitive integration emphasizes the interaction between instructional content (in this case basic and clinical science), and trainees' cognition. The authors of this work argue that the benefits of cognitive integration arise from greater conceptual coherence (Murphy & Medin, 1985) among features of clinical signs and symptoms of disease. By presenting basic science as a causal explanation to *why* a constellation of clinical signs and symptoms are connected, trainees develop a more efficient organization of this content beyond merely improved memory for these features (Baghdady et al., 2009). Conceptual coherence provides an organization framework for trainees' to make sense of cause and effect relationships between basic science and clinical features, creating more robust mental representations of disease and greater understanding (Kulasegaram et al., 2015).

The benefits of cognitive integrations may apply more broadly to domains of knowledge beyond the basic and clinical sciences required in clinical diagnosis. Though scholars have made calls for the expansion of what constitutes basic science in HPE (Bandiera et al., 2013; Bandiera Glen et al., 2017; Lucey, 2013), the majority of research on integration has been confined to teaching clinical reasoning and to traditional basic sciences such as the biology and pathophysiology of disease that are examined at the level of program as opposed to level of session, or cognition (Kulasegaram et al., 2013). Research is required to clarify the nature of what constitutes basic and clinical sciences, and their respective impacts on trainees' skill retention and transfer across the spectrum of clinical competencies trainees will be required to become expert in.

1.5 Synthesis

Literature on expertise development and cognitive integration suggest that when applying simulation to HPE, we must be cognizant of how simulation instructional design impacts the types of expertise trainees will develop. Though instructional design features such as deliberate practice may prove effective in enhancing educational outcomes from simulation-based training, clearly there is an under-studied link between the active ingredients of instructional designs and improved expertise and skill transfer. Any education that does not consider the building blocks of knowledge that undergird expertise may merely prepare trainees to *reproduce* knowledge and skills – not transfer. Using simulation as a means to replicate real clinical scenarios and skills may only be teaching trainees the procedural knowledge required for routine expertise and may be missing out on the added benefits of conceptual knowledge.

In contrast to prescriptions of more effort, more practice, and more realism, answers for instructional design may lie in research findings from psychology and education, which have clarified the structure of expert knowledge that enables adaptive expert behaviours (Bédard & Chi, 1992; Chase & Simon, 1973; Chi et al., 1988; Starkes & Allard, 1993). These findings have been applied to research exploring the instructional strategies that assist trainees in building expert-like knowledge structures in their own cognition (Chi, 2006; Chi & VanLehn, 2012; Chi & Wylie, 2014). Through a focus on imparting the relevant conceptual and procedural knowledge to trainees, simulation instructional design may be designed in ways that promote their integration and facilitate adaptive expertise development. However, whether the model of cognitive integration of basic and clinical sciences may apply beyond diagnostic reasoning to other clinical competencies taught using simulation remains to be evaluated. It remains to be tested, for instance, whether basic science and clinical science are specific instances of the more general typologies of conceptual and procedural knowledge. Further, it is unclear how integrated instruction should be designed in the context of learning novel psychomotor skills. Unlike clinical reasoning, psychomotor skill acquisition operates on sensory-motor, haptic, spatial, auditory information that are not present in the materials used to teach clinical reasoning (e.g., visual and textual). Thus, how to design simulation instruction that imparts conceptual and procedural knowledge to facilitate skill transfer is an empirical question that remains to be addressed.

Prescriptions for greater simulation realism to improve HPE have yet to demonstrate that the greater trainee engagement they are purported to instill (through physical, semantical, and phenomenal realism), translate to improved learning outcomes. Though the "suspense of disbelief" is a plausible explanation (i.e., mechanism) for why we might hypothesize trainees learn more effectively from simulation, without empirical data of its impact on learning outcomes, it is an inadequate criterion for determining whether a simulation or simulator "worked". Theoretical descriptions of realism then, offer hypotheses about how greater realism and fidelity *could* impact trainee learning through greater trainee engagement; however, these hypotheses remain to be tested. This dissertation defines realism as a measure of the degree of similarity between training and clinical scenarios – and uses the terms realism, fidelity, and authenticity interchangeably. In exploring these concepts, it offers an alternative perspective to the role of realism and engagement. Rather than aligning realism to suit perceived educational needs and functions (e.g., physical, semantical, phenomenal), I propose that, simulation instructional design should be aligned with the transfer of learning; and thus, the procedural and conceptual knowledges shown to support transfer. Simulation realism should be secondary to the integration of procedural and conceptual knowledge.

In this dissertation, "transfer of learning" is operationalized primarily as an outcome measure – a transfer test of participants' ability to respond to clinically relevant changes between training and testing contexts; changes made based on practicing clinicians' judgements. Defined this way, the transfer test used in this dissertation aligns with what Broudy (1977) describes as *applicative knowing*, whereby trainees are assessed on their ability to apply their previously acquired knowledge and skills to novel contexts. Further, based on the manipulations between training and transfer contexts described below, the transfer test can also be considered a form of *far transfer*, where the changes extend beyond the superficial features of a problem or skill (i.e., *near transfer*) and also affect its deep structural features (Perkins & Salomon, 1992). Such changes are conceptual in nature, affecting the way a problem looks or feels in ways that manipulate the relationship between the previously acquired problem-solving procedure (or skill) and the desired outcome or solution. In other words, the transfer test requires more than trainees simply directly applying what they learned from their training to "solve" the test, and instead requires them to adjust their approach to the problem based on structural changes to the problem itself. Thus, the solution is not immediately evident, even when a trainee recognizes that the transfer test is related to what they previously learned during their training.

Taken together, it appears that practice on its own, even with high-fidelity simulators or voluminous deliberate practice, will be insufficient if simulation instructional designs do not also address the knowledge trainees acquire and how these knowledges relate to the types of experts we are seeking to develop in health care. If our goal is the development of adaptive expertise, our instructional designs should facilitate mechanisms known to support trainees' procedural flexibility, and primary outcomes of research studies should utilize transfer tests as opposed to post-tests or retention tests. Are we setting trainees along a path toward routine or adaptive expertise?

Chapter 2

Aims, Hypotheses, & Methodology

2.1 Research Aims and Purposes

This dissertation consists of three clarification studies that reveal novel possibilities for simulation instructional design to enhance trainees' skill transfer and contributes to our understanding of how simulation writ large can be optimally employed to advance HPE. Importantly, each of the findings rests upon the foundation of theory and arises from a program of studies informed by the same theoretical framework and designed to build upon one another. The research program is an example of clarification research that aims to address *why* an educational intervention works (Cook et al., 2008). It accomplishes this by comparing different active interventions against one another (Cook, 2010, 2012) and by building and testing theory (Norman, 2004). Theory serves as the backbone of this dissertation, iteratively woven into each study design to allow for the building and synthesis of old and new knowledge.

The overarching goal of this dissertation is to explore and test the role of conceptual understanding in how trainees develop improve skill transfer in the simulation-based training setting. Building on theory from education, psychology, and clinical reasoning, this program of research operationalizes concepts and practices shown to be effective in building adaptive expertise in domains outside those typically used in simulation-based training. Specifically, the study designs outlined below have been informed by understandings of the procedural and conceptual knowledge that underpin adaptive expertise in education, and the framework of cognitive integration that has been applied to link basic and clinical sciences in the study of clinical reasoning.

To examine the role of conceptual knowledge, instructional strategies for integration, and skill transfer, this dissertation includes a series of three experimental studies to address the following research aims:

1. To characterize the relationship between conceptual knowledge and skill transfer.
2. To examine the impact of integrated instruction (i.e., integrating conceptual knowledge with procedural knowledge) on skill retention and transfer.

3. To test whether the process of cognitive integration mediates the effectiveness of skill transfer in simulation-based procedural skills training. That is, integrated instruction that supports cognitive integration is more effective than integrated instruction that does not.

4. To design and test novel simulation-based instruction to enhance cognitive integration and transfer.

2.2 Hypotheses

From these aims I had the following hypotheses:

1. Greater conceptual knowledge (but not procedural knowledge) would be associated with improved skill retention and transfer.

2. Integrated instruction would improve skill retention and transfer through its positive effect on trainees' conceptual knowledge.

3. Compared to integrated instruction that merely presented both procedural and conceptual knowledge, integrated instruction that supports cognitive integration of these two knowledges would better enhance conceptual understanding, and thus, skill retention and transfer.

4. Procedural and conceptual knowledge serve the same cognitive function as clinical and basic sciences in clinical reasoning; that is, basic science represents a subset of conceptual knowledge and clinical science represents a subset of procedural knowledge.

5. Simulation itself (e.g., simulators) can be designed to accentuate the causal linkages between procedural and conceptual knowledge and thereby support skill retention and transfer.

2.3 Overview of Methods

All three studies in this dissertation examine novice medical students learning how to perform a Lumbar Puncture (LP) using instructional videos and simulation-based training. I selected LP because of the rich procedural and conceptual knowledge that could be readily manipulated in our instruction and the availability of a simulator the research team felt could be modified to assess skill transfer and the potential benefits of trainees' integrated knowledge. Each of the three studies use experimental methods to compare the impact of different instructional interventions for LP (integrated and unintegrated instruction) on participants LP performance on immediate post-tests, retention tests, and transfer tests, as well as written tests of procedural and conceptual knowledge.

The following subsections below provide a general overview of each study's protocol, instructional materials, simulated scenarios, assessments, and statistical analyses. Further details of study comparisons and methods are provided in the individual chapters for each study.

2.3.1 Protocol

Each of the three studies was conducted in two sessions separated by about one-week: a training session, and follow-up session. During the training session, participants viewed video-based instruction about LP, and were given one-hour of self-regulated LP in a simulated scenario, which was followed by an immediate post-test of the same training scenario. During the follow-up session, participants returned to the laboratory to complete both a retention test on the same LP scenario experienced the week prior, and a transfer test.

All studies utilize a delayed retention and transfer period of about one-week to allow for washout of any transient effects of instruction that can artificially inflate immediate post-test performance (Schmidt & Bjork, 1992a). This one-week timeframe was chosen based on previous studies of psychomotor learning and simulation-based procedural skills training also conducted with LP (Brydges, Nair, Ma, Shanks, & Hatala, 2012; Cheung et al., 2016; Haji, Cheung, Woods, Regehr, de Ribaupierre, et al., 2016).

2.3.2 Instructional Materials and Simulation Scenarios

LP instructional materials and simulation scenarios were adopted and modified from a previous study of simulation-based LP in the same student population (Haji, Cheung, Woods, Regehr, de Ribaupierre, et al., 2016).

For the instructional materials (instructional videos), instructional videos were modified to reflect differences in the type of knowledge they contained, as well as how these knowledges were related through sequencing and verbal instructions. These adjustments allowed us to operationalize the presentation of procedural and conceptual knowledge and manipulate the extent to which cognitive integration would be supported.

For the training/post-test/retention scenario, participants had to complete the LP procedure on a healthy patient laying on their side with normal anatomy. For the transfer scenario, I adjusted the scenario such that the LP procedure required that the procedure be done with the patient (and thus simulator) sitting upright as opposed to laying on their side as participants would have encountered during their training and post-test scenarios. These scenarios were consistent across all three studies, except for study 3, where I counter-balanced the order of retention and transfer tests.

2.3.3 Outcome Measures

Our primary outcome was LP performance in the retention and transfer scenarios. Expert raters reviewed and assessed video-recorded LP performances on a global rating scale (GRS). Post-test performances were deemed unreliable assessments of learning given the transient effects of practice on immediate performance, and thus I sought to focus on the more robust measures of retention and transfer (Dubrowski, 2005; Schmidt & Bjork, 1992a; Schmidt & Lee, 2005). Though raters were also given task-specific checklists to complete for each performance, these were intended only as an orienting tool. The GRS was used as a primary outcome of interest because it permits assessing how participants apply knowledge and skills in a gestalt manner, as opposed to how they perform all the necessary steps listed on a procedural checklist; as a less reductive tool, the GRS was judged to be likely more sensitive for detecting the subtleties of expertise we expected to observe in our transfer test (Ilgen, Ma, Hatala, & Cook, 2015; Ma et al., 2012; Regehr, Macrae, Reznick, & Szalay, 1998).

To complete our assessments of LP performance, I recruited expert raters with experience and training in performing LPs (or spinal anesthesia techniques) to rate video-recordings of participants' simulated LP performances. Raters consisted of senior residents or fellows in either anesthesia or neurology who completed rater orientation sessions prior to completing ratings. Specifics details of rater recruitment and orientation are described for each study in their respective chapters.

The written procedural and conceptual knowledge tests were developed through expert consultation with staff and residents who were involved in clinical teaching of LP and/or related procedural skills. The procedural knowledge test consisted of a sorting exercise, whereby participants were tasked with placing a list of 13 steps for LP in the appropriate order. The conceptual knowledge test consisted of a list of short answer questions, whereby participants were tasked with providing answers to questions about why particular steps of the LP skill were performed in a particular manner, or what might happen if these instructions were not followed. I made modifications to the conceptual knowledge test between the studies in attempts to improve item statistics (e.g., difficulty). Any adjustments in test content or delivery are described in the individual study chapters.

To address questions about the role of procedural and conceptual knowledge in simulation-based procedural skills transfer, the term procedural is used in two meanings: procedural knowledge, and procedural skill. Procedural knowledge is consistent with previous definitions in the education literature as a type of knowledge related to "knowing how" to perform a specific (problem-solving) procedure in a given situation or problem. It is qualitatively different from conceptual knowledge (knowing why). Procedural skill is used to describe proficiency in clinical procedures, which may or may not involve both procedural and conceptual knowledges, and thus, adaptive expertise and the ability to demonstrate skill transfer.

2.3.4 Simple Mediation Analyses

Our analyses included simple mediation analyses in all three studies (Hayes, 2013). Simple mediation clarifies the simultaneous relationship between three variables: instructional group (control versus intervention), a mediator variable, and an outcome variable (Leppink, 2015). Unlike traditional approaches to determining mediation that are criterion-based (e.g., full or partial mediation), the simple mediation approach allows for inferential statistical testing of mediation using boot-strapping methods (Hayes, 2009). This statistical method allowed us to capture and test the effect of our intervention *mediated* through its impact on conceptual knowledge (indirect effect), in addition to the overall effect of our interventions on retention and transfer (total effect) as would be captured using either an Analysis of Variance (ANOVA) comparison between groups, or a regression model. Further, unlike an Analysis of Covariance (ANCOVA), which controls for the effect of a confounding variable, the mediation approach allows us to consider the positive (or negative) impact of a mediator (or suppressor) variable on our outcome of interest; in the case of this dissertation, the positive impact of conceptual knowledge on retention and transfer outcomes.

An overview of this statistical method is provided in each of the individual study chapters; however, to illustrate the use of simple mediation analysis, consider for example, addressing whether a new method of teaching was superior to an old method of teaching. Let us assume our *a priori* hypothesis was that the new method would improve student motivation and result in students spending more time studying (which we were able to capture), and thus improve student knowledge. In our analyses, we could compare the knowledge tests scores of trainees' who received the new method versus the old method, but how might we capture the influence of studying time? Using ANOVA or linear regression, we could use three separate analyses to test the relationship between group and time on task, group and test scores, and time on task and test scores. We find that the new method of teaching increases participants' knowledge test scores as well as their time on task (studying) and that greater time on task is associated with higher knowledge test scores. However, these analyses would not be able to capture the relationship between all three variables. Using ANCOVA, we could clarify this relationship by accounting for time on task within a single model. However, time on task is inputted as a covariate, we now find there are no group differences after controlling for the effects of time on task. This may lead us to prematurely conclude that our new teaching method had no effect on the trainees' knowledge. This however would be misguided, because we know there was a positive effect of the intervention on time on task (students who received the new method of teaching studied more), and also that students who studied more had higher test scores. In essence, the ANCOVA obfuscates the impact of the intervention through our hypothesized mechanism, time on task. On the other hand, simple mediation analyses capture all three paths (as assessed through the three ANOVAs) and avoid the interpretive challenges of an ANCOVA – which do not allow us to test the hypothesized learning mechanism through which the new method of teaching ‘works’. An in-depth review of the statistical procedure can be observed in Hayes (2013).

2.4 Significance

Simulation provides a controlled setting to test the impact of cognitive integration on the learning and transfer of core procedural competencies (e.g., lumbar puncture, central line insertion) (American Board of Internal Medicine, 2016; The Royal College of Physicians and Surgeons of Canada, 2012). Additionally, recent systematic reviews have revealed simulation to be an ideal setting to test the effectiveness of instructional design features (Cook et al., 2012, 2011; Cook, Hamstra, et al., 2013), and have identified simulation-based assessments of procedural skills competency that have robust validity evidence (Ilgen et al., 2015). Therefore, simulation-based procedural skills present a unique opportunity to design and study the impact of integrated instruction on transfer of learning.

In testing these theories and concepts in the context of simulation, this dissertation extends theories on the knowledge underlying adaptive expertise and cognitive integration into the realm of clinical psychomotor skills taught in simulation. By building theory and conducting comparative experimental studies of different instructional interventions, this dissertation also addresses the lack of clarification studies available in simulation and the overall lack of theory. Finally, by designing and comparing the impact of novel instructional designs rooted in theory, findings from these studies can serve educators in informing them of instructional design features and strategies that may support the effective use of use simulation as an educational tool to enhance health profession education.

Chapter 3

Knowing How and Knowing Why: Testing the impact of instruction designed for cognitive integration on procedural skills transfer

Adapted with permission from Springer Nature Customer Service Centre GmbH: Springer Nature. Cheung JJH, Kulasegaram KM, Woods NN, Moulton CA, Ringsted CV, Brydges R. Knowing How and Knowing Why: testing the effect of instruction designed for cognitive integration on procedural skills transfer. *Advances in Health Sciences Education: Theory and Practice*. 2018. DOI 10.1007/s10459-017-9774-1. License #: 4542110213573.

3.1 Preamble

One of the principle gaps in simulation instructional design is the linkage between how specific instructional design features influence learning transfer. This chapter represents a first attempt to operationalize and test the impact of integrated instruction (designed to support cognitive integration) on simulation-based procedural skills transfer. The primary goal of this study was to clarify the relationship between conceptual knowledge and skill transfer and explore the efficacy of an integrated instructional video (containing both procedural and conceptual knowledge) to support both. A second goal of the study was to delineate the effects of integrated instruction, and thus adaptive expertise development, on different learning outcomes – those that required adaptation (i.e., application or skill transfer) and those that theoretically would not (i.e., reproduction of a previously acquired skill). Hence, we opted to also compare the impact of integrated instruction on trainees' post-test and retention test performances.

In this study, we developed the experimental protocol, instructional materials, assessments, and statistical approaches that the subsequent two studies of this dissertation would build upon. This includes the experimental administration (research assistant scripts) of the protocol, integrated and procedural only instructional videos, written procedural and conceptual knowledge tests, and mediation analyses.

This paper was accepted for publication in Advances for Health Sciences Education: Theory and Practice, a journal whose primary focus is linking education theory and practice through publishing theoretically and methodologically rigorous studies and commentaries. The discussion from this chapter contributes to literature on both simulation-instructional design and clinical reasoning. The study hypothesizes and tests a relationship between cognitive integration, as studied in clinical reasoning, and skill transfer in psychomotor, procedural skills taught through simulation. The study identifies a potential role for instructional interventions that emphasize conceptual knowledge, such as the integrated instruction developed and tested in the study. However, without a comparison with another integrated instructional intervention, the study on its own does not confirm whether cognitive integration is the optimal means of delivering conceptual knowledge to enhance skill retention and transfer for simulation-based procedural skills.

3.2 Abstract

Transfer is a desired outcome of simulation-based training, yet evidence for how instructional design features promote transfer is lacking. In clinical reasoning, transfer is improved when trainees experience instruction integrating basic science explanations with clinical signs and symptoms. To test whether integrated instruction has similar effects in procedural skills (i.e., psychomotor skills) training, we studied the impact of instruction that integrates conceptual (*why)* and procedural (*how)* knowledge on the retention and transfer of simulation-based lumbar puncture (LP) skill. Medical students (N=30) were randomized into two groups that accessed different instructional videos during a 60-minute self-regulated training session. An unintegrated video provided procedural *How* instruction via step-by-step demonstrations of LP, and an integrated video provided the same *How* instruction with integrated conceptual *Why* explanations (e.g., anatomy) for key steps. Two blinded raters scored post-test, retention, and transfer performances using a global rating scale. Participants also completed written procedural and conceptual knowledge tests. We used simple mediation regression analyses to assess the total and indirect effects (mediated by conceptual knowledge) of integrated instruction on retention and transfer. Integrated instruction was associated with improved conceptual (p<.001) but not procedural knowledge test scores (p=.11). We found total effect of group (p>.05). We did find a positive indirect group effect on skill retention (B_{ab}=.93, p<.05) and transfer (B_{ab}=.59, p<.05), mediated through participants improved conceptual knowledge. Integrated instruction may improve trainees' skill retention and transfer through gains in conceptual knowledge. Such integrated instruction may be an instructional design feature for simulation-based training aimed at improving transfer outcomes.

3.3 Introduction

An assumed benefit of simulation-based training is that trainees will transfer what they learn effectively into subsequent learning experiences, and ultimately, into their clinical practice (Hamstra et al., 2014). Despite this common assumption (Boet et al., 2014; Teteris et al., 2012), research evidence on how best to facilitate transfer through instructional design of simulation-based training is incomplete (Cook, Hamstra, et al., 2013). Lists of evidence-based instructional design features do exist (Cook et al., 2012; Issenberg et al., 2005; McGaghie et al., 2010), however, these lists are mostly based on studies assessing trainee performance on post-tests (immediately after simulation training) or retention tests (after a delay), but not on transfer tests (Cook et al., 2012). Post-tests and retention tests assess a trainee's ability to recall information or reproduce performance, whereas transfer tests assess a trainee's ability to apply previous knowledge and skills to new contexts (Chi & VanLehn, 2012; Needham & Begg, 1991). Hence, most empirical studies and reviews of the literature on simulation-based training have not addressed whether and how educators can design training to improve transfer.

One instructional design manipulation linked to improved transfer outcomes involves directing learners to create meaningful relationships between relevant types of knowledge – a process called cognitive integration (Kulamakan Mahan Kulasegaram et al., 2015, 2013). When teaching learners to make clinical diagnoses, for example, cognitive integration appears to be best supported by instructional materials presenting conceptual explanations about the underlying mechanisms of disease (e.g., basic science knowledge) alongside the clinical signs and symptoms required for diagnosis (e.g., clinical science) (Baghdady et al., 2014a; Kulamakan Mahan Kulasegaram et al., 2015, 2013). Such integrated instruction appears to enhance novice learners' retention of clinical reasoning performance (Baghdady et al., 2014a, 2014b, 2013, 2009; Woods et al., 2005, 2007a; Woods, Neville, et al., 2006), as well as their ability to transfer learning to novel problems, as measured by the ability to diagnose more difficult clinical cases accurately (Woods et al., 2007b).

Like in clinical reasoning, cognitive integration may be a process that helps learners develop the underlying memory structures for psychomotor skills learning and transfer (Schmidt, 1975). Psychomotor skills are highly cognitive, requiring numerous decision-making processes such as planning, coordinating, regulating, and interpreting movement tasks (Starkes & Allard, 1993). Much research shows that encouraging learners to engage in these cognitive problem-solving operations improves skill acquisition and transfer (Guadagnoli & Lee, 2004; Lee, Swinnen, & Serrien, 1994). These benefits have been demonstrated via the superiority of random vs. blocked practice (J. B. Shea & Morgan, 1979; J. B. Shea & Zimny, 1983, 1988) delayed vs. concurrent feedback schedules (Lee, White, & Carnahan, 1990; Winstein & Schmidt, 1990), and modelling performances with augmented feedback vs. without (McCullagh & Caird, 1990). Learners engaged in these beneficial elaborative conceptual processes (J. B. Shea & Morgan, 1979) are thought to create more meaningful and memorable representations of the movement task (Schmidt & Lee, 2005). Those representations are believed to improve subsequent retention and transfer for learning both psychomotor and cognitive skills (Lee et al., 1994; Schmidt & Bjork, 1992a). Integrated instruction, then, may provide learners with the conceptual substrate to more efficiently and effectively elaborate their learning of psychomotor skills.

As a next step in this research program on cognitive integration, simulation-based training offers a controlled setting for testing how integrated instruction impacts the learning and transfer of core invasive bedside procedures, like lumbar puncture and central line insertion (American Board of Internal Medicine, 2016; The Royal College of Physicians and Surgeons of Canada, 2012). Simulation is proven as a useful modality for testing the effectiveness of instructional design features (Cook et al., 2012, 2011; Cook, Hamstra, et al., 2013), and simulation-based assessments of procedural skills have robust validity evidence (i.e., they are likely sensitive enough to detect learning changes in a research study) (Ilgen et al., 2015).

To translate the work on cognitive integration from clinical reasoning skills to simulation-based procedural skills training, we need to identify the types of knowledge that learners must integrate. In clinical reasoning, the focus has been on integrating clinical science and basic science. These two types of knowledge can be more broadly categorized as procedural and conceptual knowledge, respectively (de Jong & Ferguson-Hessler, 1996). Procedural knowledge can be characterized as ‘knowing how’, and is defined by knowledge of the specific steps to achieve a particular goal, including the ability to execute these steps (Baroody et al., 2007). Conceptual knowledge can be characterized as ‘knowing why’, and is defined as knowledge of generalizations and principles that are not necessarily tied to particular problems or procedures (Baroody et al., 2007). This knowledge can include what is thought of as basic sciences, including anatomy and physiology. Similar to findings in clinical reasoning, when teaching mathematics problem-solving, instructional approaches integrating procedural and conceptual knowledge are associated with improved retention and transfer outcomes (Baroody, 2003; Rittle-Johnson & Schneider, 2015; Rittle-Johnson et al., 2015, 2012). Hence, in procedural skills, and likely for clinical skills more generally, we propose integrating procedural and conceptual knowledge, which parallel clinical knowledge and basic science respectively.

To investigate the impact of instruction designed to prompt cognitive integration during simulation-based procedural skills training, we compared how integrated instruction versus procedural instruction alone impacts novice medical students' immediate post-test, retention and transfer test performances of lumbar puncture (LP). We propose post-tests and retention tests require learners to replicate knowledge, whereas transfer tests require them to apply knowledge (Broudy, 1977). Previous conceptual work on retention tests suggests that participants who better understand a skill may be able to better maintain that skill over time (Dubrowski, 2005). As we expect integrated instruction will improve trainees' understanding, we hypothesize that performance will be improved on retention and transfer tests, but not on immediate post-test. Further, we hypothesize that improved transfer will be associated with improved conceptual knowledge, and do not expect the same relationship for procedural knowledge.

3.4 Methods

3.4.1 Participants

Upon receipt of University of Toronto Research Ethics Board approval, we recruited 30 undergraduate pre-clerkship medical students (Year 1 and 2; ~250 students per year). Without previous studies on the effect size of integrated instruction on procedural skills retention and transfer, we used the principle of selecting sample size based on previous studies in a related domain (Norman, Monteiro, & Salama, 2012). We chose to recruit 30 participants, which represents the median sample size of studies that have detected significant large effects when comparing instructional design features in simulation (Cook, Hamstra, et al., 2013). Notably, 30 participants is within the range (on the lower end) of studies in social psychology that employ simple mediation analyses (Rucker, Preacher, Tormala, & Petty, 2011).

3.4.2 Learning Materials

We modified pre-existing videos (Haji, Cheung, Woods, Regehr, deRibaupierre, et al., 2016) to develop the instructional videos for the How and How + Why groups. Both videos included the same procedural *how* verbal instructions and expert demonstrations of the steps of LP on a simulated patient in the lateral decubitus position. We used the Lumbar Puncture Simulator II (Kyoto Kagaku Co., Ltd, Kyoto, Japan), a part-task simulator of the lower torso. For the How + Why group's video, we integrated conceptual *why* verbal and visual instructions (i.e., causal explanations underlying key procedural steps) into the How group's video. Conceptual knowledge included principles related to anatomy, tool function, and patient safety. An example comparing the two videos is provided in Figure 3-1. Three experienced physicians (PGY5 Neurosurgery resident, PGY5 Anesthesia Resident, and an Internal Medicine Staff) provided input and feedback on all content in the instructional videos, especially related to the interpretability and accuracy of the how and why explanations. The How group video was 13:36 minutes and the How + Why group video was 17:39 minutes.

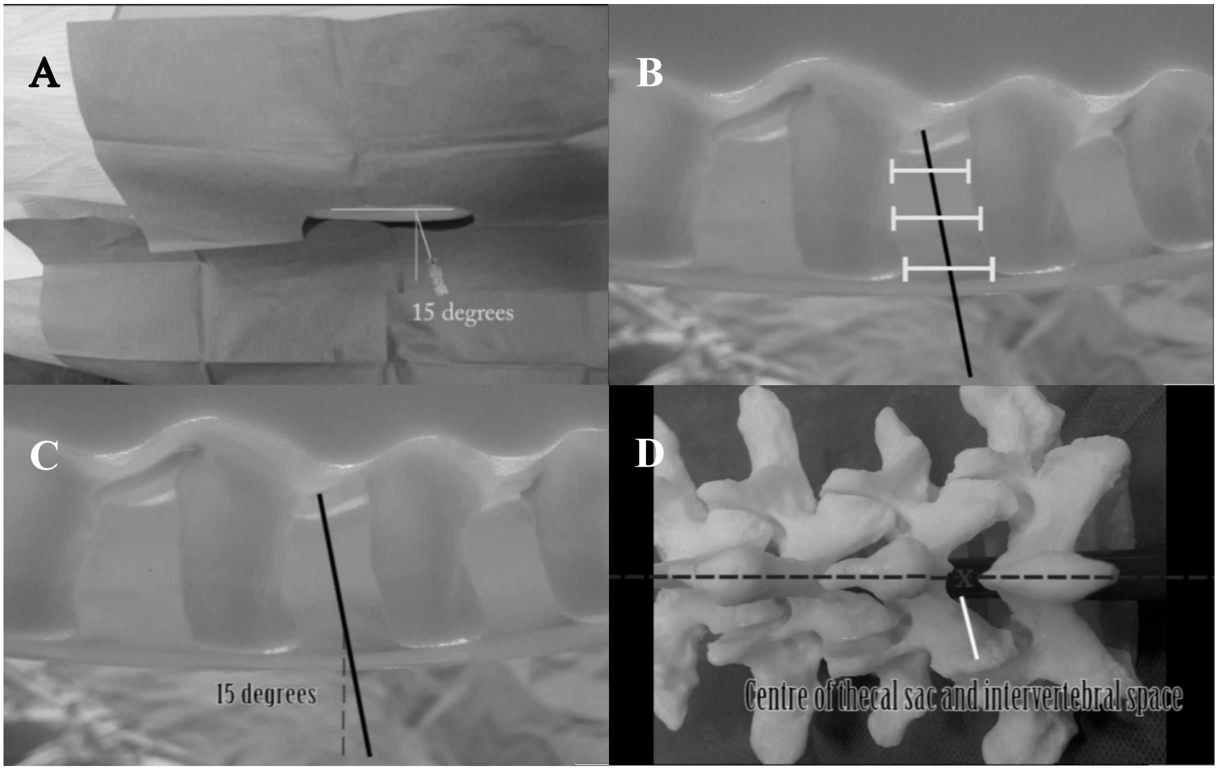


Figure 3-1. Screen captures of instructional video demonstrating procedural how instruction and conceptual why explanations. **Procedural *how* Instruction:** when inserting the spinal needle, both groups were provided these instructions in the video: “Confirm the site with your non-dominant hand and position the needle at the centre of the intended puncture site. You want this to be precisely at the midline of the patient’s body and angled at a slight 15 degrees towards the patient’s head. This corresponds with a trajectory aimed at the patient’s belly button. [diagram/animation (**A**)].” **Conceptual *why* explanation:** only the Why group received the additional causal explanation of the procedural step: “Being at the centre of the interspace ensures that you will have maximal clearance between the two spinal processes and the intervertebral space to eventually access the thecal sac. [diagram/animation (**B**)]. As well, the slight 15 degree angle allows the needle to slide comfortably between the slightly angled spinal processes of the vertebrae [diagram/animation (**C**)]. Being precisely at the midline of the patient’s body ensures that the needle is targeting the centre of the thecal sac and the intervertebral opening. [diagram/animation (**D**)].”

3.4.3 Procedure

Participants completed this study individually. We randomized participants into two groups that received access to different instructional materials during a self-regulated simulation-based LP training session. One group received the procedural *how* instructional video (How group) and the other received a video with the same procedural instructions, along with conceptual *why* explanations (How + Why group). We obtained written informed consent from each participant.

Upon arrival at their simulation-based training session, participants received free access to their assigned instructional video for 20 minutes, which we designed a form of observational practice/skill modelling. Immediately after, participants completed a procedural and a conceptual knowledge (pre-training) tests without access to their instructional video. Next, participants practiced LP repeatedly for 60 minutes, receiving no external feedback on their performance. We chose not to provide external feedback to control for potential differences in the type and quantity of knowledge provided to trainees. While the lack of external feedback may seem problematic for learning, our study design ensured participants had self-regulated access to the instructional video before, during, or after each LP attempt. Indeed, self-regulated observational practice has been repeatedly shown to enhance learning of motor skills compared to instructor-controlled practice (Cheung et al., 2016; Domuracki, Wong, Olivieri, & Grierson, 2015; Wulf, 2007). During training, the first author stopped and reset the scenario only if the participants exceeded 12 needle passes on that LP attempt.

Immediately after practice, participants completed a post-test consisting of one additional trial of the same scenario, without access to the instructional video. Participants returned one-week later to complete the retention and transfer test scenarios. As a final step, participants completed the same procedural and conceptual knowledge tests. Figure 3-2 presents a summary of the study design.

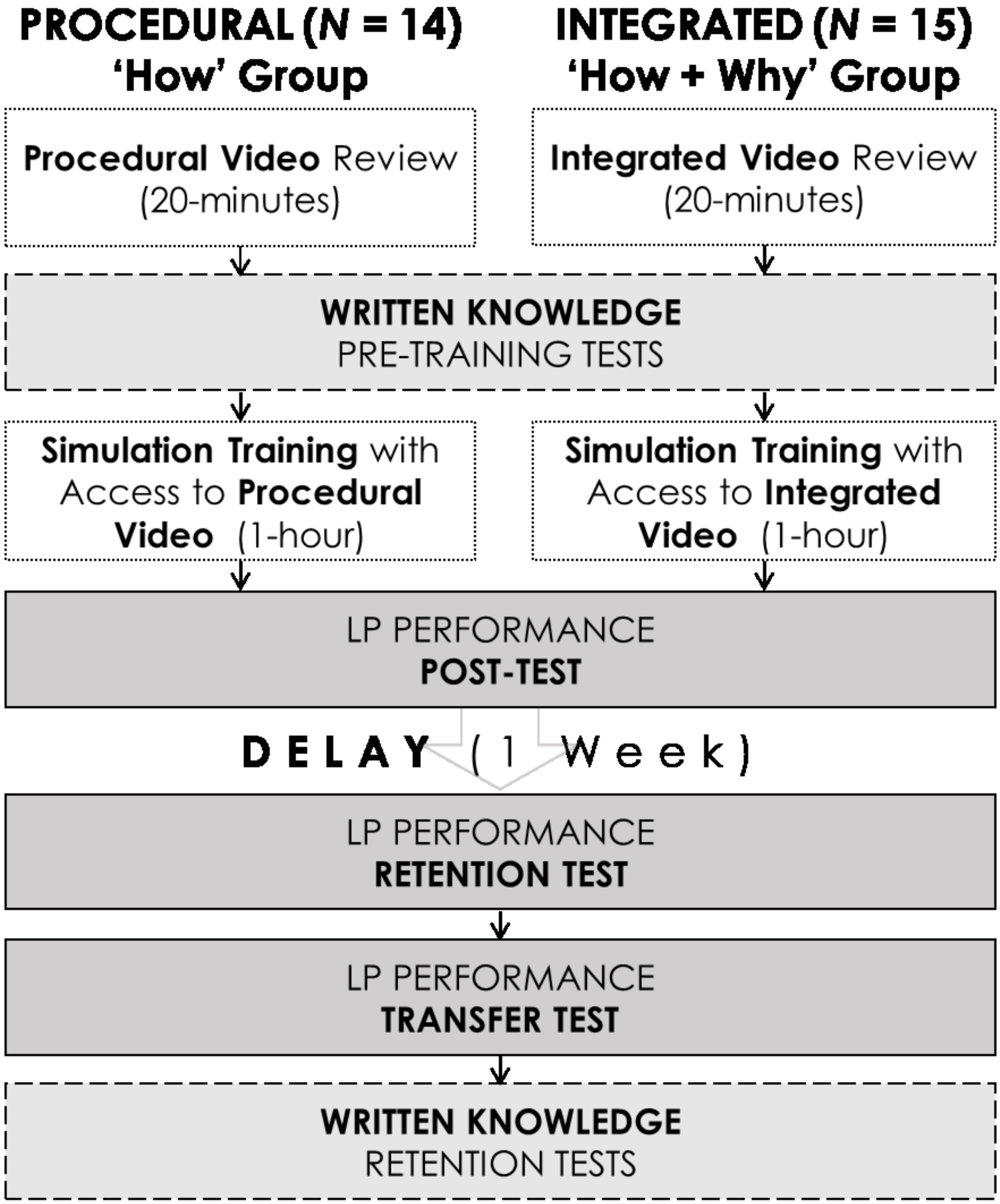


Figure 3-2. Summary of study design and procedure with group sample sizes from final analysis

3.4.4 Procedural and Conceptual Knowledge Testing

To test how well participants understood the procedural and conceptual knowledge provided in the instructional videos, we developed two written tests. The procedural knowledge test required participants to sort 13 steps of LP into the appropriate order. The conceptual knowledge test required participants to answer five short answer questions. The same three experienced physicians developed the questions and scoring rubric for this test (out 10 points):

i. Why do you angle your needle at 15 degrees? (1 points)

ii. Why do you have the bevel of the needle facing the patient's side? (3 points)

iii. Why do we you enter at the L3-L4 or L4-L5 area? (1 points)

iv. Why do we you clean in an outward fashion from the centre of the insertion site? (2 points)

v. Why do we you insert and remove the needle with the stylet in place? (3 points)

3.4.5 Lumbar Puncture Performance Testing Scenarios

Participants experienced training, post-test, and retention scenarios consisting of a healthy patient requiring a routine LP to rule out multiple sclerosis (Haji, Cheung, Woods, Regehr, deRibaupierre, et al., 2016). For training, post-test, and retention test, we positioned the simulator in the lateral decubitus position, and used a spine insert representing normal soft tissue and normal spinal anatomy. For the transfer test scenario, we positioned the simulator in the upright sitting position, and used a spine insert representing an obese patient with normal spinal anatomy. The clinical stem noted that the patient was feverish, required LP to rule out meningitis, and could not tolerate lying on his side. Notably, the 'obese' insert has a thicker layer of soft tissue that increases the difficulty of landmarking for the LP procedure.

3.4.6 Study Outcomes

For the procedural knowledge tests, the first author produced a score out of a possible 13 points (i.e., single rater, no inter-rater reliability metrics). For the conceptual knowledge tests, two blinded raters (PGY4 anesthesiology residents) independently produced a score out of 10 points.

We video-recorded the three test scenarios (post-test, retention test, and transfer test) using two cameras, one wide-angled and the other close-up view of the participants' hands and equipment (e.g., orientation of needle). The same two blinded raters independently scored each performance using a 45-item task-specific checklist (CL) developed for LP (Lammers, Temple, Wagner, & Ray, 2005) and a global rating scale (GRS) with 6 dimensions scored on a 5-point Likert scale (Martin et al., 1997). Previous studies have demonstrated that both of these assessment tools have favourable validity evidence for use in a research study (Brydges et al., 2012). Based on consistent findings of better expert-novice discrimination for GRS versus CL data (Brydges et al., 2012; Ilgen et al., 2015), we consider GRS scores on retention and transfer test as the primary outcomes, because we expect the GRS will likely be most sensitive for detecting nuanced differences in skill observed on those two tests. During a 1-hour rater training session, the raters agreed they would interpret '*competence'* (i.e., score of 3/5 on the GRS) as indicating the medical student could perform LP in the clinical context under direct supervision. We averaged GRS scores across all 6-dimensions (total GRS score divided by 6), and calculated the average score of the two raters. To assess inter-rater reliability, we calculated the intraclass-correlation coefficient (ICC) between the two raters' scores for conceptual knowledge tests, and GRS scores.

3.4.7 Statistical Analyses

To compare conceptual and procedural knowledge test scores between groups and across pre-training and retention tests we used a 2x2 (group x test) repeated measures mixed analysis of variance (ANOVA).

To conduct our simple mediation regression analyses (i.e., a form of path analysis) (Hayes, 2013; Leppink, 2015; Rucker et al., 2011), we used the PROCESS macro for SPSS Version 22 (provided in Hayes, 2013) to assess relationships between participants' assigned group, their conceptual knowledge scores, and their retention and transfer test scores. These analyses allowed us to assess the total effect of our group treatment on LP performance (Figure 3-3A), as well as the direct and indirect effect of our group treatment when including conceptual knowledge scores as a mediator in the model (Figure 3-3B). We used the conceptual knowledge scores at retention as the mediator variable based on theory (i.e., improved conceptual knowledge would mediate our group effect), and based on positive correlations between outcome variables: conceptual and procedural knowledge test scores at retention, and GRS scores at retention and transfer (where detected, results reported below). Post-test LP GRS scores were not significantly correlated with any outcome variables and thus mediation analyses were not conducted. On a technical note, these path analyses require the use of bootstrapping methods to statistically test the product of *a* and *b* paths (i.e., *a* x *b*).

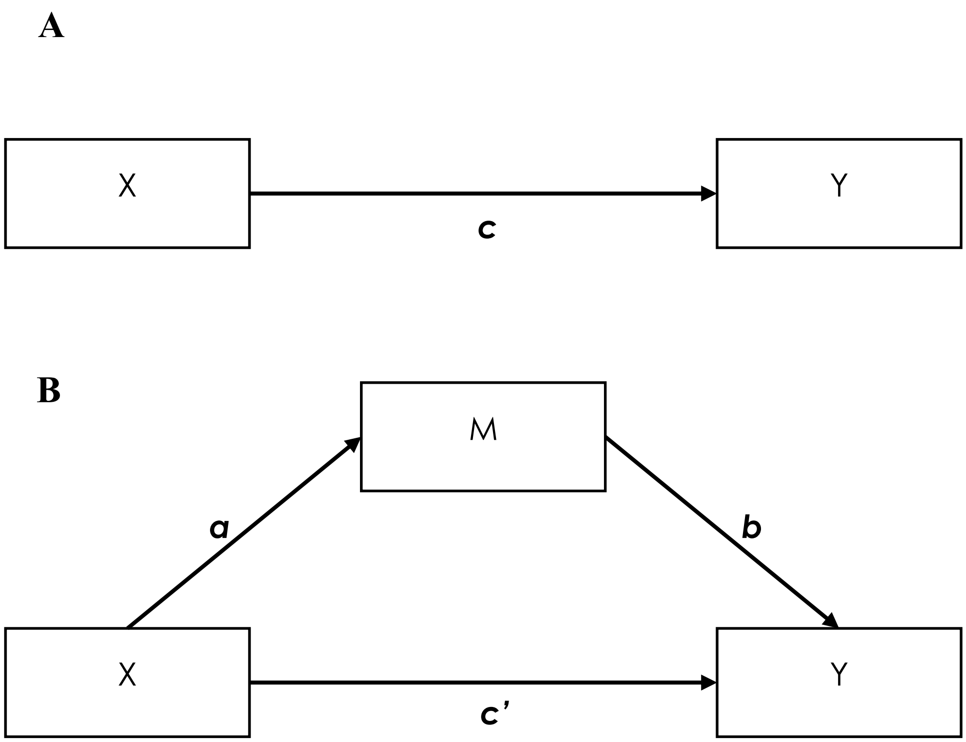


Figure 3-3. Regression analyses for total effect of intervention X on outcome variable Y (*c*) (**A**) and regression analysis for indirect effect of intervention X via mediator variable M (*a* x *b*) (**B**), i.e., simple mediation analysis. Respective path coefficients are represented by *c, a, b, and c’* variables with the path coefficient for an indirect effect represented by *a* x *b*.

3.5 Results

One participant in the How group was unable to attend the retention session and was excluded from our analyses. Inter-rater reliability was excellent for the conceptual knowledge test (ICC = .85), and was fair for the GRS (ICC = .64). All data are reported as mean ± standard error. All outcome measures are summarized in Table 3-1.

Table 3-1. Mean group scores ± standard error for written procedural and conceptual knowledge tests, and Global Rating Scale scores for lumbar puncture performance on post-test, retention test, and transfer test.

	Group	
	How	**Why**
Procedural Knowledge Test		
Pre-training	12.43 ± 0.27	12.00 ± 0.29
Retention	12.79 ± 0.15	12.33 ± 0.25
Conceptual Knowledge Test		
Pre-training	3.32 ± 0.30	5.05 ± 0.29
Retention	2.96 ± 0.26	4.80 ± 0.33
Lumbar Puncture GRS Score		
Post-test	3.68 ± 0.27	3.28 ± 0.21
Retention Test	3.24 ± 0.20	3.11 ± 0.25
Transfer Test	2.65 ± 0.17	2.91 ± 0.26

3.5.1 Procedural and Conceptual Knowledge Test Performance

For the procedural knowledge test, ANOVA revealed no significant effects of time ($F(1,27) = .19$, $p = .15$), group ($F(1,27) = 2.73$, $p = .11$), and their interaction term ($F(1,27) = .003$, $p = .96$). For the conceptual knowledge test, ANOVA revealed no significant effect of time ($F(1,27) = 3.14$, $p = .09$) and a significant effect of group ($F(1,27) = 21.33$, p < .0001, $\eta_p^2 = 0.44$) with the How + Why group (4.93 ± .28) scoring significantly higher than the How group (3.14 ± .28). The interaction term was not significant ($F(1,27) = .09$, $p = .77$).

Our correlation analyses revealed conceptual knowledge test scores related positively and significantly with GRS scores at both retention ($r = .47$, $p = .01$) and transfer ($r = .43$, $p = .02$). Conversely, the procedural knowledge test scores did not correlate significantly with GRS scores at either retention ($r = -.03$, $p = .89$) or transfer ($r = .19$, $p = .32$). Additionally, retention and transfer GRS scores did not correlate significantly ($r = .28$, $p = .14$), and post-test GRS scores did not correlate significantly with any outcome measure.

3.5.2 LP Performance Tests

Path *c* in Figure 4A shows there was no significant total effect of group on retention ($c = -.13$) or transfer ($c = -.26$). Figure 4B depicts the simple mediation analysis of GRS retention and transfer scores when accounting for conceptual knowledge scores in the model as a mediator variable. Path *a* is identical for both retention and transfer and shows participants in the How + Why group had significantly higher written conceptual knowledge test scores ($a = 1.84$). Path *b* reveals that participants (i.e., regardless of group) with higher conceptual knowledge scores exhibited significantly better retention ($b = .51$) and transfer ($b = .32$). The indirect effect of group on retention ($ab = .93$) and transfer ($ab = .59$), mediated by conceptual knowledge scores, was tested by calculating a bias-corrected bootstrap 95% confidence intervals using 15,000 bootstrap samples, which was entirely above zero for retention (95% CI[.44, 1.69]) and transfer (95% CI[.13, 1.16]), and thus significant. After controlling for participants' conceptual knowledge scores, a significant direct effect of group was also detected for retention ($c' = -1.06$), but not transfer ($c' = -.33$).

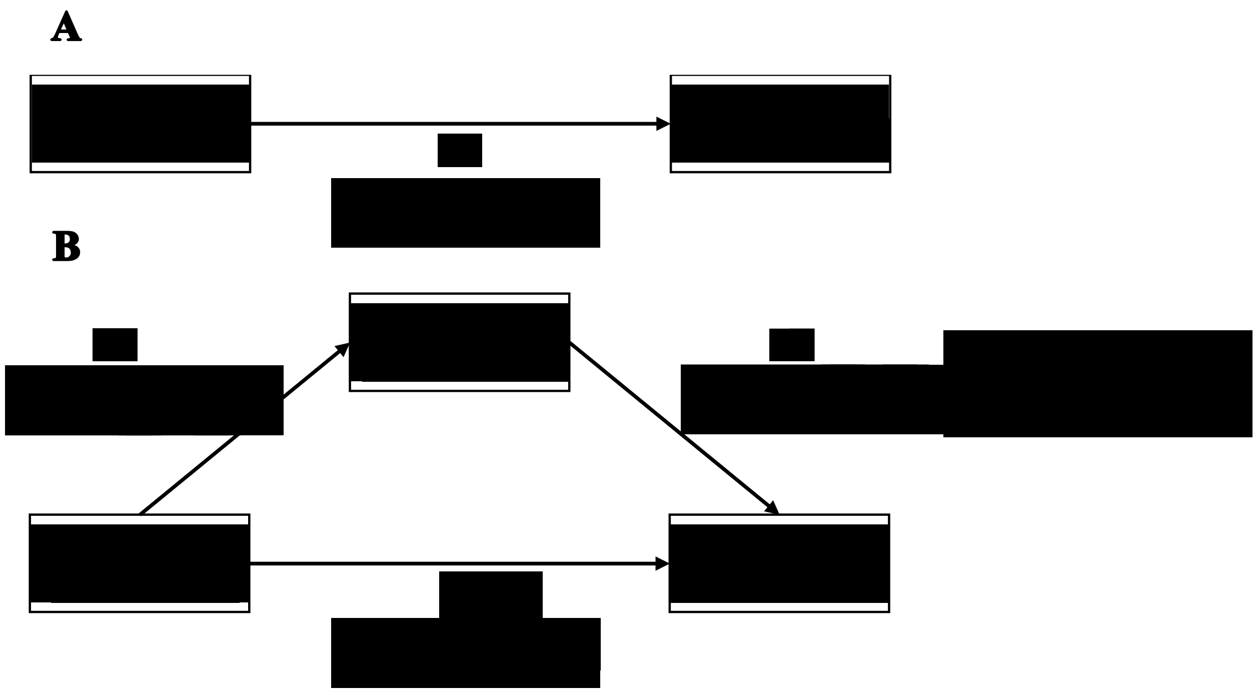


Figure 3-4. Unstandardized coefficients of paths in the mediation analysis are represent by *a*, *b*, *c'* for the outcome variable of GRS of retention (**A**) and transfer performances (**B**). Group Treatment effect on Conceptual Knowledge by *a*, Conceptual Knowledge effect on GRS Retention and Transfer Performance by *b*, and direct effect of Group Treatment is represented by *c'*. The indirect effect of Group Treatment mediated by Conceptual Knowledge is represented by *ab*; the total effect of group treatment without accounting for conceptual knowledge is represented by *c*. *$p < .05$

3.6 Discussion

We examined the effects of integrating conceptual *why* explanations with procedural *how* instructions on novice learners' skill retention and transfer of simulation-based LP skills. Our results demonstrated an indirect effect of instruction designed to enhance cognitive integration on the retention and transfer of LP skills, an effect that was mediated by participants' improved conceptual knowledge. There was however, no significant total effect of integrated instruction on either retention, or transfer. Interpreting our results in terms of educational effect, the data show that for two participants with the same conceptual knowledge test score, a participant in the How + Why group would score an average of .93 more points on the GRS at retention and an average of .59 more at transfer. Scored out of 5, this translates to a difference of 19% and 12%, respectively. Related to our hypotheses, the integrated instruction was associated *directly* with improved conceptual knowledge scores, and it was associated *indirectly* with improved transfer and retention outcomes. As hypothesized, we found no statistically significant relationship between immediate post-test scores and any other variable.

Our results extend previous findings showing benefits of instruction designed to enhance cognitive integration for clinical reasoning skills to the training of simulation-based procedural skills (i.e., skilled performance of a psychomotor skill). When teaching clinical reasoning, cognitive integration of clinical and basic science knowledge is thought to benefit diagnostic ability through the resulting *conceptual coherence* developed by learners (Kulamakan Mahan Kulasegaram et al., 2013). Conceptual coherence, in diagnostic reasoning, is present when learners can organize their clinical knowledge into coherent mental representations, structured by basic science concepts (Woods, 2007; Woods et al., 2007a). When teaching procedural skills, integrated instruction may also help learners to create conceptual coherence that organizes their physical actions. That is, the indirect positive effect of integrated instruction on transfer and retention outcomes implies that improved conceptual knowledge may have better enabled participants to organize and perform the procedural steps of LP. The exact mechanisms of cognitive integration in skills requiring both cognitive and psychomotor skills, like invasive procedures, ultrasonography, and physical examination maneuvers, will need to be studied in future research.

3.6.1 Implications of Integrated Instruction for Instructional Design and Transfer

Our findings suggest that transfer of simulation-based procedural skills can be improved by instructional designs that support cognitive integration of procedural and conceptual knowledge. Current lists of instructional design features recommended for simulation-based training omit integrated instruction (Cook et al., 2012; Issenberg et al., 2005; McGaghie et al., 2010). When integration is described, it is discussed at the curriculum level (e.g., including four simulation half-days in a program) (Cheung et al. 2016), rather than at the session level where cognitive integration is best supported (e.g., scheduling exactly how simulation will be used during each 4-hour half-day) (Kulamakan Mahan Kulasegaram et al., 2013). Findings from education (Bransford & Schwartz, 1999; DeCaro & Rittle-Johnson, 2012; Kapur, 2014; Rittle-Johnson et al., 2012; Schwartz & Bransford, 1998) psychology (Needham & Begg, 1991), and in clinical reasoning (Mylopoulos et al., 2016; Mylopoulos & Woods, 2014) suggest the cognitive mechanisms, and thus the instructional designs, supporting retention may not be the same as those supporting transfer. Thus, instruction that integrates procedural and conceptual knowledge may represent a novel instructional design feature for simulation training, one with a strong theoretical basis for its benefits for transfer of learning.

3.6.2 Limitations and Future Directions

Though our integrated instruction had a significant indirect effect on retention and transfer, the total effect of the intervention, not accounting for conceptual knowledge, was not significant. This may be caused by our intervention being underpowered, and by other limitations of our study design. First, we did not control for participants' previous knowledge, and those in both groups may have engaged in cognitive integration using prior conceptual knowledge not provided in our instructional videos (e.g., anatomy knowledge from their formal curriculum). Second, the written conceptual knowledge test delivered prior to simulation-based training may have encouraged cognitive integration through the effects of self-explanation (Chamberland et al., 2013) or test-enhanced learning (Larsen, Butler, & Roediger III, 2008). Third, participants in both groups scored poorly on the conceptual knowledge test. Though the How + Why group scored significantly higher, this was a difference of roughly 1 point (approximately 10%), and overall participants in both groups scored less than 50%. Fourth, we developed the conceptual knowledge test for this study and further refinement and validation is required to ensure robust assessment in future work.

Given our main finding involves an indirect effect, our study design prevents us from disentangling the effect of integrated instruction versus the effect of improved conceptual knowledge on our learning outcomes. Hence, one area of further inquiry is to test whether providing conceptual knowledge in isolation (i.e., not integrated with procedural knowledge) is sufficient for trainees to improve their retention and transfer outcomes. Such a study would determine if educators must spend the time and effort creating materials for integrated instruction, or whether learners can use conceptual knowledge on their own (i.e., spontaneously integrate) to develop the conceptual coherence necessary for improved future performance.

3.6.3 Conclusions

By extending findings on cognitive integration from clinical reasoning to simulation-based procedural skills training, our study adds an instructional design feature that has largely been over-looked in this domain: integrated instruction that helps learners form relationships between their procedural knowledge of *how* and their conceptual knowledge of *why* when learning a procedure. Crucially, our findings suggest the resulting cognitive integration is associated with improved transfer outcomes, addressing an additional gap in the healthcare simulation literature.

Chapter 4

Why Content and Cognition Matter: Integrating Conceptual Knowledge to Support Simulation-Based Procedural Skills Transfer

Adapted with permission from Springer Nature Customer Service Centre GmbH: Springer Nature. Cheung JJH, Kulasegaram KM, Woods NN, Brydges R. *Why* Content and Cognition Matter: Integrating Conceptual Knowledge to Support Simulation-Based Procedural Skills Transfer. *Journal of General Internal Medicine*. 2019. DOI 10.1007/s11606-019-04959-y. License #: 4542110213573.

4.1 Preamble

Building upon the previous chapter, this experimental study aimed to clarify the relationship between cognitive integration and simulation-based procedural skills transfer. The first study of this dissertation (Chapter 3) established a positive relationship between conceptual knowledge and LP skill retention and transfer. Further, the mediation analyses demonstrated that integrated instruction, designed for cognitive integration, had a positive impact on both skill retention and transfer through improvements in conceptual knowledge. That is, the integrated instruction led to greater conceptual knowledge in trainees, which in turn led to improved performance on LP at retention and transfer. Left unclear, however, was whether cognitive integration was the mechanism at play, or whether the mere inclusion of conceptual knowledge in the integrated instructional video yielded these improvements. This study aimed to address the mechanism through which integrated integration supports skill retention and transfer. Does instruction designed to support cognitive integration via explicit cause-and-effect explanations between procedural and conceptual knowledge represent the optimal means of enhancing simulation-based learning? Or perhaps, participants can engage in spontaneous integration of these knowledges as long as they are both presented?

To test cognitive integration as a mechanism supporting LP skill retention and transfer, the study added an additional integrated instruction group to the previous study design (three groups) and increased the group sample size (n = 15 per group to n = 22 per group). The new integrated instruction intervention was designed to integrate conceptual and procedural knowledge "in sequence", presenting conceptual knowledge first, followed by a step-by-step procedural knowledge. In contrast, the other integrated instruction, was "integrated for causation", which theoretically should better support cognitive integration, and thus skill retention and transfer. Both integrated instruction interventions were video-based and designed to contain the same content, that is, the procedural knowledge and conceptual knowledge instructions were equivalent. The procedural knowledge portion of the integrated in sequence video was the exact same instructional video that was available to the procedural only control group. In addition to addressing whether instruction integrated in sequence would be as effective as instruction integrated for causation, the same overall protocol was maintained to allow for a replication of previous findings of Chapter 3.

There are several differences to note between the study in the current chapter and study 1 (Chapter 3). With the help of expert clinicians and HPE graduate student reviewers (naïve to the LP procedure), the original video-based integrated instruction was modified to bolster its effectiveness and renamed to the "integrated for causation" video. This name more accurately reflected the underlying design principle used for integration, which was to facilitate cognitive integration. Following the same rationale, the new integrated group was named the "integrated in sequence" video, again to reflect the underlying design principle of teaching conceptual knowledge first, followed by procedural knowledge. Another change to the study design was in the conceptual knowledge test. To enhance the sensitivity of the test in detecting group differences, expert clinician educators were consulted to provide additional relevant conceptual knowledge questions to the test. As well, the written procedural knowledge and conceptual knowledge tests were administered just once during participants' follow-up session, rather than twice (at both sessions).

The study findings reaffirm the relationship between conceptual knowledge and transfer and reveal cognitive integration as the likely learning mechanism supporting skill transfer. Using a modified mediation analyses for three group comparisons, the study shows the same positive indirect effect of the integrated instruction on transfer when compared to the procedural only group. Further, this indirect effect was significantly larger for the integrated for causation group compared to the integrated in sequence group. The results were not replicated for retention, which may be due to changes in the protocol, specifically, the delivery of the procedural and conceptual knowledge tests.

This study was submitted to the Journal of General Internal Medicine for publication in a special issue on medical education. Though it was accepted, the editorial board decided to publish the article in another to-be-determined issue. The journal targets an audience of primary health care specialists and provides scholarly research and perspective pieces on improving patient care, primary care education, general internal medicine, and hospital medicine. The discussion points of the article focused on the importance of instructional design that facilitates transfer, and the potential that conceptual understanding (i.e., integrating procedural and conceptual knowledge) holds in advancing simulation instructional design. Part of this discussion critiques the current practice-based models that value volume, time, and effort over considerations of the ingredients of expertise that will prepare trainees to be adaptive and transfer their skills. Though the study advances understanding of cognitive integration in simulation-based procedural skills training, it has yet to demonstrate a total effect of the intervention; which suggests there may be methodological and pedagogical modifications to enhance instruction that aims to “integrate for causation” – thereby improving cognitive integration.

4.2 Abstract

Background: Curricular constraints require being selective about the type of content trainees practice in their formal training. Teaching trainees procedural knowledge about “how” to perform steps of a skill along with conceptual knowledge about “why” each step is performed can support skill retention and transfer (i.e., the ability to adapt knowledge to novel problems). However, how best to organize How and Why content for procedural skills training is unknown.

Objectives: We examined the impact of different approaches to integrating why and how content on trainees’ skill retention and transfer of simulation-based lumbar puncture (LP).

Design and Participants: We randomized medical students (N = 66) to practice LP for 1-hour using one of three videos. One video presented only the how content for LP (Procedural Only). Two other videos presented how and why content (e.g., anatomy) in two ways: Integrated in Sequence, with why content followed by How content, or Integrated for Causation, with how and why content integrated throughout.

Main Measures: Pairs of blinded raters scored participants’ retention and transfer LP performances on a global rating scale (GRS), and written tests assessed participants’ procedural and conceptual knowledge.

Key Results: Simple mediation regression analyses showed that participants receiving an integrated instructional video performed significantly better on transfer through their intervention’s positive impact on conceptual knowledge (all $p<0.01$). Further, the Integrated for Causation group performed significantly better on transfer than the Integrated in Sequence group ($p<0.01$), again mediated by improved conceptual knowledge. We observed no mediation of participants’ skill retention (all $p>0.01$).

Conclusions: When teaching supports cognitive integration of how and why content, trainees are able to transfer learning to new problems because of their improved conceptual understanding. Instructional designs for procedural skills that integrate How and Why content can help educators optimize what trainees learn from each repetition of practice.

4.3 Introduction

When teaching trainees clinical skills, research and curriculum documents emphasize high-volumes of training as *the* path to expertise (Ericsson, 2004). Research on how educators can optimize such training has made a strong case for instructional design features like deliberate practice (Ericsson, 2006b; McGaghie et al., 2014), mixed-practice (Hatala, Brooks, & Norman, 2003), spaced-practice (Moulton et al., 2006), and retrieval practice (i.e., test-enhanced learning) (Larsen, Butler, Lawson, & Roediger III, 2012; Larsen et al., 2008), which have all been linked consistently to better learning outcomes. While this literature informs educators on how to structure training for effective skill learning (e.g., provide timely feedback, include a range of difficulties and contextual variations, space practice across time, and encourage retrieval practice by consistent testing), it often neglects the role content (i.e., the material trainees have been assigned to learn about a skill) plays in expertise development. When teaching trainees to perform a bedside invasive procedure for example, what content would an educator present about anatomy, equipment, sterility, patient safety, and communication? How and when should an educator relate these different content areas to trainees' procedural actions?

Educators must inevitably make decisions about what they teach, what they do not teach, and how they choose to relate the selected types of content (or not). Moreover, in the finite number of hands-on training sessions allotted to any curriculum, educators need to maximize how well they prepare trainees to generalize their skills to address problems encountered in novel contexts, otherwise known as the transfer of learning (Kulasegaram & McConnell, 2016). Informing educators on how best to select and organize different types of content has the potential to optimize teaching approaches, and to help learners reap the most from each repetition during training.

One aspect of content, conceptual knowledge, has consistently been shown to underlie expertise development and skill transfer. Conceptual knowledge refers to the generalizable principles that transcend specific contexts of a task or procedure (e.g., the type of clinical environment, or particular features of a patient case), and is described as "knowing why" (Baroody et al., 2007). Such knowledge differs from procedural knowledge, which refers to the specifics of executing a task or procedure proficiently, and is described as "knowing how". For example, basic sciences comprise the conceptual knowledge underlying clinical reasoning, whereas clinical knowledge of the constellation of signs and symptoms of specific disease states comprise the procedural knowledge required to engage in clinical reasoning. In deconstructing expertise, researchers have found that experts rely on their understanding of basic science pathophysiology to solve non-routine clinical cases (de Bruin, Schmidt, & Rikers, 2005; Mylopoulos & Woods, 2009; Rikers, Loyens, & Schmidt, 2004; Rikers, Loyens, te Winkel, Schmidt, & Sins, 2005). In studying how expertise develops, researchers have found that trainees learning to make clinical diagnoses have better diagnostic skill retention and transfer when teaching involves integrating basic science and clinical knowledge (Baghdady et al., 2014a, 2014b, 2013, 2009; Castillo et al., 2018; Kulasegaram et al., 2017; Kulasegaram et al., 2015, 2014; Woods, 2007; Woods et al., 2005, 2007b; Woods, Howey, Brooks, & Norman, 2006; Woods, Neville, et al., 2006). Hence, identifying and selecting the relevant conceptual and procedural knowledge appears to be a key decision point when educators choose which content to teach.

Research shows that it matters how educators organize conceptual and procedural knowledge in teaching material. A series of studies demonstrate that integrating conceptual and procedural knowledge using “science based causal explanations” led to superior diagnostic skill retention and transfer compared to teaching that merely presented these two types of content in close spatial and temporal proximity (Baghdady et al., 2013; Kulasegaram et al., 2017; Kulasegaram et al., 2015). That is, teaching material helped trainees achieve improved outcomes when it explicitly integrated conceptual and procedural knowledge in a way that encouraged them to make causal connections between the two, a process referred to as *cognitive integration* (Kulasegaram et al., 2013). Hence, evidence has accrued to show that how educators organize procedural knowledge (i.e., how content) and conceptual knowledge (i.e., why content) matters for trainees’ outcomes.

Organizing content to promote cognitive integration represents an instructional design principle that may generalize to procedural skills. In our previous study,(Cheung et al., 2017) we designed a video to illustrate the causal explanations between how and why content for novice medical trainees learning lumbar puncture (LP) on a part-task simulator (i.e., a mannequin representing only a patient’s lower back). For example, we taught learners to angle their needle at a 15 degree towards the simulator’s umbilicus (how content) *because* the underlying anatomy of the spinous processes are also angled at 15 degrees (why content). We compared the impact of this integrated instructional video to a video containing only procedural knowledge and found preliminary evidence that the integrated instructional video improved participants' procedural skill retention and transfer. We did not clarify, however, the extent to which these benefits resulted from simply including conceptual knowledge in the video (vs. not including it), or from how we had organized the how and why content in the video (i.e., using causal explanations to support cognitive integration).

In this study, we aim to replicate and extend our previous research. Using the controlled setting of simulation-based LP training, we experimentally tested the impact of three conditions on LP skill retention and transfer: procedural and conceptual knowledge 'Integrated for Causation' (i.e., how and why content interleaved throughout and linked by causal explanations), both knowledges 'Integrated in Sequence' (i.e., conceptual knowledge first, followed by procedural knowledge), and procedural knowledge presented alone. Further, we examined the potential mediating role that trainees' conceptual knowledge plays in supporting their skill retention and transfer.

4.4 Methods

4.4.1 Participants

After receiving institutional ethics approval, we recruited 66 medical students from the University of Toronto. Inclusion criteria included being a pre-clerkship (year 1 and 2) MD student; participants were excluded if they had previous LP training. We based this sample size on our previous study using the same procedural skill, student population, and similar educational interventions (Brydges et al., 2012; Cheung et al., 2017).

4.4.2 Learning Materials

We developed three instructional videos for LP based on previous educational materials (Cheung et al., 2017; Haji, Cheung, Woods, Regehr, de Ribaupierre, et al., 2016). The videos present procedural and conceptual knowledge for LP in varying combinations. The procedural knowledge component (How content) demonstrates how to appropriately execute the steps necessary to complete an LP. The conceptual knowledge component (Why content) demonstrates key concepts underlying the technical performance of LP, specifically, spinal anatomy, equipment function and design, sterility, and patient safety. Our first video presented How content through a step-by-step LP demonstration on a simulated mannequin (Procedural Only). Our other two videos integrated How and Why content using two different organizational approaches. One organized the How and Why content in sequential order, presenting conceptual knowledge first followed by the same procedural knowledge as the first video (Integrated in Sequence). In the other video, we organized the How and Why content in an interleaved fashion (Integrated for Causation), a design intended to help trainees establish cause and effect relationships between the procedural steps and their related concepts. The Procedural Only video was 13:48, the Integrated in Sequence video was 23:04, and the Integrated for Causation video was 21:01. See Box 4-1 for an example of how content is presented in each video.

Box 4-1. Instructional Materials
All three instructional videos contain the same How content (i.e., same demonstration of LP), but present this content at different times. For example, for the procedural step of inserting the spinal needle, all videos provide the following verbal instruction:

"Now pick up the spinal needle from the tray and remove the sheath placing it back in the tray. Ensure the stylet is firmly inside, and that the bevel of the needle is facing upwards or downwards, towards the patient's side…" **Procedural Only** [7:32-7:44]; **Integrated in Sequence** [16:48-17:00]; **Integrated for Causation** [10:37-10:49]

The integrated videos present Why content in addition to How content. They differ in the *temporal* and *causal* relationships they are designed to establish between these two content types.

In the Integrated in Sequence video, conceptual explanations for why the stylet should be firmly inside the needle and why the bevel oriented in the prescribed manner are provided at the beginning of the video. Trainees viewing this video are tasked with connecting this conceptual knowledge with the relevant procedural knowledge provided later in the video [16:48-17:00]:

> *"The outer membrane of the thecal sac consists of a tough connective tissue called dura mater. The fibres of the dural tissue run longitudinal and parallel to the spine."* [0:56-1:07]
>
> *"During invasive procedures that compromise the dural tissue, excessive trauma can result in prolonged CSF leakage, which can further lead to severe headaches due to the loss of CSF cushioning and supporting the brain…"* [2:10-2:22]
>
> *"…if epithelial tissue is introduced into the subarachnoid space, it can lead to the growth of a cyst…"* [2:53-2:58]
>
> *"The stylet blocks the shaft of the needle preventing the formation of skin plugs that may clog the needle."* [3:37-3:42]
>
> *"The opening of the needle tip is where you to observe the bevel, an angled cutting edge. Tissue trauma can be reduced by aligning this cutting-edge parallel to the fibres of the tissue, allowing the fibres to be spread apart rather than cut."* [3:49-4:03]

By contrast, the Integrated for Causation video presents the same conceptual explanations in close temporal proximity with the procedural instruction. Further, the connections between the conceptual and procedural are made explicitly using cause-and-effect language. Trainees viewing this video experience both How and Why content in a manner intended to promote cognitive integration.

> *"When inserting the spinal needle, having the stylet in place will prevent skin tissue from entering the hollow shaft of the needle, forming a skin plug, which if introduced into the subarachnoid space can lead to the growth of a cyst."* [10:49-11:03]
>
> *"When the bevel faces the patient's side, the sharp cutting edge of the needle will be parallel with the fibres of the dura that cover the thecal sac and run longitudinal and parallel with the spine. Thus, allowing these fibres to be spread apart rather than cut. This will reduce the size of the tear made in the dura and consequently the amount of CSF that leaks out of the subarachnoid space after the procedure. Which can lead to severe headaches for the patient due to the loss of CSF that cushions and supports the brain."* [11:04-11:33]

4.4.3 Procedure

We used stratified randomization (by study year and sex) to allocate 22 participants into each of the three groups: Procedural Only, Integrated in Sequence, and Integrated for Causation. All participants then attended a self-regulated simulation-based LP training session and a follow-up session one-week later. Each participant completed the study protocol individually, with the lead author present at each training and retention session.

At the training session, after providing written informed consent and completing a demographic questionnaire, participants had 25-minutes to review their assigned instructional video. Participants were made aware they were randomized to one of three video interventions but were unaware of how these videos differed in content or organization. Immediately after, participants had 1-hour to practice a simulated scenario of LP on a part-task model (Lumbar Puncture Simulator II, Kyoto Kagaku Co., Ltd, Kyoto, Japan). During practice, participants had access to their instructional video (via laptop) and could alternate between practicing the scenario and reviewing the video. After practice, participants were tested on the same scenario without access to the instructional video (post-test).

One week later, participants returned for the follow-up session, requiring they complete a retention test, on the same scenario from the week prior, followed by a transfer test, on a new LP scenario. After the transfer test, participants completed written tests of procedural knowledge and conceptual knowledge. The lead author administering the training and follow-up sessions could not be blinded to participants' group allocation. To minimize potential bias, all external feedback provided to participants regarding their LP performances was withheld until the end of the study. Thus, participants' only source of LP content came from the instructional videos and their own self-directed practice. Figure 4-1 presents the study design.

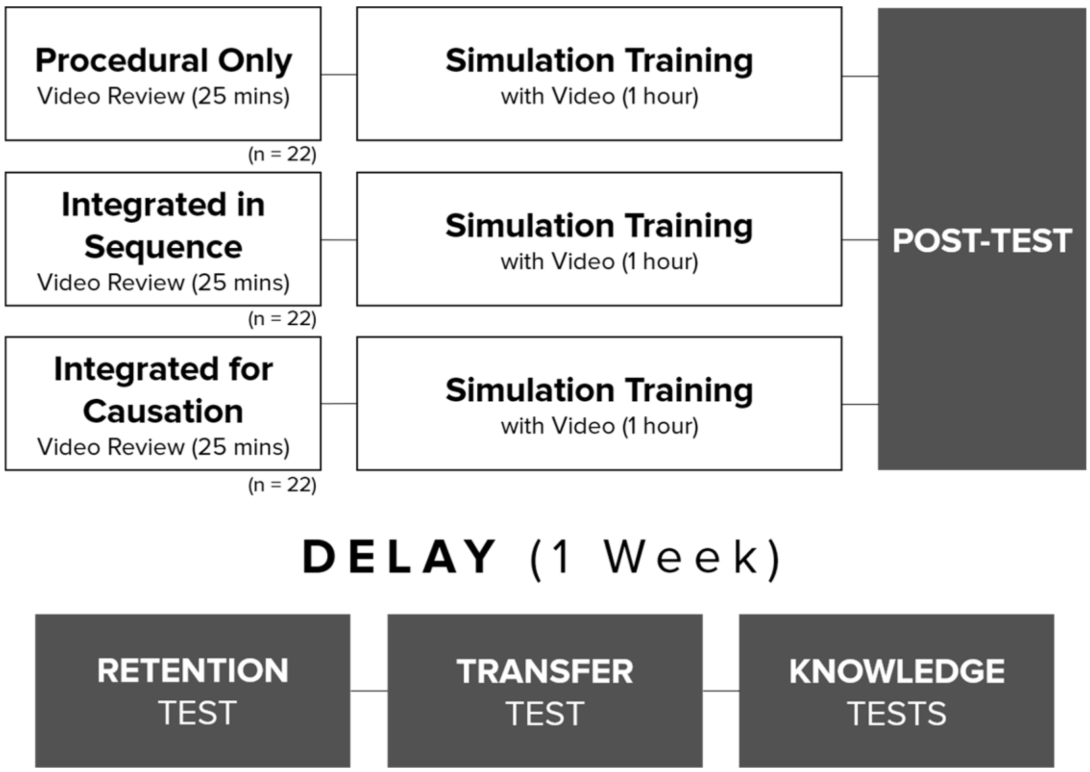


Figure 4-1. Study Flow Diagram

4.4.4 Procedural and Conceptual Knowledge Tests

We adopted both written knowledge tests from our previous study (Cheung et al., 2017). We used the exact same procedural knowledge test, requiring participants to sort 13 key steps of the LP procedure into their appropriate order. We modified the conceptual knowledge test, adding six new short-answer items, based on consultations with procedural specialists across Canada.

4.4.5 LP Simulation Scenarios

The simulation scenarios presented during training, post-test, and retention test involved a healthy patient requiring an LP to rule out multiple sclerosis. The simulator was positioned in the lateral decubitus position with an anatomical spine insert that represented normal anatomy. The transfer test scenario involved a sick, older, obese patient, suspected of having meningitis who could not tolerate lying on his side, thus requiring the LP to be performed with him sitting upright. The simulator was rotated into an upright posture, and the anatomical spine insert represented obese anatomy (i.e., thicker tissue). Both scenarios were taken from our previous study (Cheung et al., 2017), which were originally adapted from Haji, Cheung, Woods, Regehr, de Ribaupierre, et al., (2016).

4.4.6 Outcome Measures and Analyses

To assess the procedural knowledge tests, the first author reviewed and scored each test out of a maximum of 13 points. To assess the conceptual knowledge tests, two neurology residents (PGY4 and PGY5), blinded to participant group allocation, reviewed the tests and scored each out of a maximum of 20 points. After conducting an item analysis of the conceptual knowledge test, we removed four items with a difficulty index >0.85 (more than 85% of all participants answered the item correctly), resulting in a maximum score of 13 points. The procedural knowledge test and final conceptual knowledge test questions are included as supplemental content (see Appendix I and II respectively). To compare group performances on the written procedural and conceptual knowledge tests, we used Welch's one-way analysis of variance (ANOVA) and Games-Howell Post-Hoc Test as needed; these tests allowed us to account for the lack of homogeneity of variance (Levene's test $p < 0.05$).

To assess LP performances, we video recorded all participants' tests. One camera captured a wide-angle shot of the overall procedure and a second camera focused on participants' hands as they manipulated the LP equipment. We merged these two recordings to produce a single split-screen video of each test performance for rater assessment. The raters included six neurology residents (two PGY3, three PGY4, and one PGY5) blinded to group allocation of all recorded performances. The raters attended a two-hour rater orientation session aimed at familiarizing them with the Global Rating Scale (GRS) used to assess procedural competence (Martin et al., 1997). Researchers have collected strong validity evidence for using the GRS to assess simulated procedural skills (Ilgen et al., 2015). and thus we used it as our primary outcome. During the orientation session, raters scored three randomly selected LP performances (out of the 198 total) and discussed their scoring on each until they came to a consensus score. Raters concluded the session by coming to a consensus that a score of three out of five on the GRS denoted the participant was "capable of performing the procedure (or dimension of the GRS) independently without compromising patient safety."

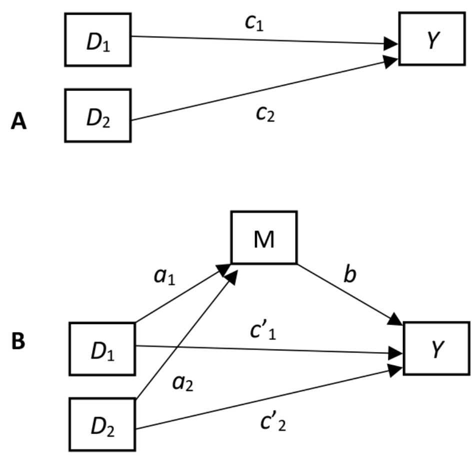
After rater orientation, we randomly allocated the raters into three pairs to score a pilot sample of 15 randomly selected LP performances. We assessed inter-rater reliability by calculating an intra-class correlation coefficient. One rater pair showed poor reliability (ICC < 0.60) and we removed their data from the analyses. We reassigned this pair's 15 videos along with the 150 remaining to be scored by the two other rater pairs, both of which demonstrated high reliability (ICCs > 0.80) in the pilot sample. We performed all reliability analyses using G String IV version 6.3.8 (Bloch, 2013), and performed comparative and correlational analyses using SPSS Version 22.

4.4.7 Mediation Analyses

To capture the relationship between our experimental conditions, participants' conceptual knowledge, and their LP GRS performance, we conducted simple mediation analyses(Leppink, 2015; Rucker et al., 2011) using the PROCESS macro for SPSS (Hayes, 2013). Using indicator coding, we compared group performances by computing two mediation models for retention, and transfer test GRS scores (Hayes, 2013). In the first model, the Procedural Only group was coded as the control, allowing us to compare the two integrated groups to the Procedural Only group; in the second model, the Integrated in Sequence group was coded as the control, allowing us to compare the two integrated instruction groups. We did not compute mediation models for LP GRS performance at post-test because our previous study results revealed no significant correlation between conceptual knowledge and post-test LP GRS performance (Cheung et al., 2017).

These analyses enabled us to examine how the three interventions influenced the groups' retention and transfer test scores in three ways (see Figure 4-2): (1) the relative total effects on those outcomes (c_1 and c_2 in Figure 4-2A); (2) the relative direct effects (c'_1 and c'_2 in Figure 4-2B), after controlling for participants' conceptual knowledge test scores (M); and (3) the relative indirect effects (a_1b and a_2b in Figure 4-2B), when including the conceptual knowledge (M) as a mediator of intervention effects in the model. To calculate and compare the relative indirect effects of each intervention (Figure 4-2B: D_1 & D_2 acting on Y through M), the PROCESS macro computed a bias-corrected bootstrap 99% confidence interval using 10000 bootstrap samples of the product between path $a_k \times b$. To account for family-wise error from multiple comparisons, we set our alpha to 0.01. Using this methodology, non-zero confidence intervals denote statistical significance.

These mediation analyses allowed us to test our hypotheses that participants' conceptual knowledge would mediate their procedural skill retention and transfer, and that the mediation effect would be larger when participants received instruction Integrated for Causation versus Integrated in Sequence.



Indicator Coding

Model		Procedural Only	Integrated in Sequence	Integrated for Causation
1	D_1	0	1	0
	D_2	0	0	1
2	D_1	1	0	0
	D_2	0	0	1

Figure 4-2. Simple mediation analyses of intervention effects (*D*) on outcome *Y* (i.e., Global rating scale scores for retention or transfer performances) using indicator coding. Figure 2A reveals the relative total effects (c_1 and c_2) of D_1 and D_2; Figure 2B includes Conceptual Knowledge Test Score (*M*) as a mediator variable to reveal the effect of *M* on *Y* (*b*), and the relative indirect (a_1b and a_2b) and direct effects (c'_1 and c'_2) of D_1 and D_2 on *Y*. For comparison of relative effects, the Procedural Only and Integrated in Sequence groups were set as controls in Models 1 and 2 respectively.

4.5 Results

Inter-rater reliability was excellent for the conceptual knowledge test (ICC = 0.90) and the GRS (ICC = 0.89). All descriptive data are presented as means (*M*) and standard deviations (*SD*) and are summarized in Table 4-1.

4.5.1 Procedural and Conceptual Knowledge Test Performance

For the procedural knowledge test, there was no significant difference between the groups (Procedural Only: $M = 12.50$, $SD = .80$; Integrated in Sequence: $M = 12.27$, $SD = 1.78$; Integrated for Causation: $M = 12.86$, $SD = .47$), $F(2,35.80) = 1.46$, $p = .10$. For the conceptual knowledge test, there was a significant difference between the groups, $F(2,34.71) = 25.82$, $p < .001$, $\eta_p^2 = .40$. Post-hoc analyses showed the Integrated for Causation group ($M = 5.98$, $SD = 2.13$) scored significantly higher on the conceptual knowledge test than both the Procedural Only group ($M = 2.69$, $SD = 0.77$), $t = 3.28$, $p < .001$, and the Integrated in Sequence group ($M = 4.14$, $SD = 1.84$), $t = 1.84$, $p < .05$. Further, the Integrated in Sequence group scored significantly higher than the Procedural Only group, $t = 1.44$, $p < .01$.

4.5.2 Mediation Models Linking Conceptual Knowledge and Lumbar Puncture Performance

Confirming the findings above, the mediation model (Figure 4-2B) also detected group differences in conceptual knowledge, as participants in both integrated instruction groups had higher conceptual knowledge scores relative to participants in the Procedural Only group (Model 1: $a_1 = 1.44$ and $a_2 = 3.28$; Tables 4-2 and 4-3). When comparing participants receiving the two types of integrated instruction (Model 2), the Integrated for Causation group scored higher in conceptual knowledge than the Integrated in Sequence group (Model 2: $a_2 = 1.84$, Tables 4-2 and 4-3).

The model allowed us to examine the relationship between all participants' (regardless of their assigned group) conceptual knowledge (*M*) and their GRS performance outcomes (*Y*) (Path '*b'* in Figure 2). We found that participants who had higher conceptual knowledge had higher but not significantly different retention scores (b = .579) (Table 4-2). Conversely, we found that participants who had higher conceptual knowledge had significantly higher transfer scores (b = 1.24) (Table 4-3).

4.5.3 Relative Total and Direct Effects: Effects of the Interventions Before and After Controlling for Participants' Conceptual Knowledge

Without adjusting for participants' conceptual knowledge, we found no significant differences in the relative total effects (c_1 & c_2 in Figure 4-2A) of any intervention on LP GRS retention (Table 4-2) and transfer (Table 4-3) performances, all $p > 0.05$. After controlling for participants' conceptual knowledge (*M* in Figure 4-2B), we similarly found no significant relative direct effects (c'_1 & c'_2 in Figure 4-2B) on LP GRS retention (Table 4-2) or transfer (Table 4-3) performances, all $p > 0.05$.

4.5.4 Relative Indirect Effects: Effects of the Interventions with Participants' Conceptual Knowledge as a Mediator

Relative to the Procedural Only group (Model 1), both integrated instruction interventions indirectly influenced and improved participants' transfer performance via their improved conceptual knowledge (Integrated in Sequence [a_1b = 1.76], 99% CI = .40 to 3.94; Integrated for Causation [a_2b = 4.08], 95% CI = 1.35 to 7.72).

Relative to the Integrated in Sequence group (Model 2), the indirect effect of the Integrated for Causation group was also non-zero (a_2b = 2.29], 95% CI = .38 to 5.81), indicating a significant indirect influence of that group's intervention on their transfer performance via their improved conceptual knowledge compared to the Integrated in Sequence group.

4.5.5 Summary of Effects

To illustrate the mediating link between the interventions and participants' conceptual knowledge, the model showed that, if we assumed equal scores on the conceptual knowledge test, relative to a participant in the Procedural Only group, a participant in the Integrated in Sequence group scored 1.76 more points on the transfer test GRS, and a participant in the Integrated for Causation group scored 4.08 more points. On a GRS out of five, this equates to an improvement of 0.30 (6.0%) and 0.68 (13.6%) points, respectively. Again, adjusting for conceptual knowledge test scores, participants in the Integrated for Causation group scored 2.29 points higher on the GRS at transfer test compared to a participant in the Integrated in Sequence group, equating to 0.39 (7.7%) points out of five.

We did not observe this indirect effect of the interventions, mediated through conceptual knowledge, on participants' LP GRS retention performance (Table 4-2). Specifically, the confidence intervals of the relative indirect effects contained zero, and thus were non-significant (Model 1: Integrated in Sequence [a_1b = .84], 99% CI = -.35 to 2.83; Integrated for Causation [a_2b = 1.90], 95% CI = -.89 to 5.80; Model 2: Integrated for Causation [a_2b = 1.07], 95% CI = -.36 to 4.30).

Table 4-1. Mean group scores ± standard deviation for written procedural and conceptual knowledge tests, and Global Rating Scale scores for lumbar puncture performance on retention and transfer tests.

Outcome Measure	**Procedural Only Group**	**Integrated in Sequence Group**	**Integrated for Causation Group**
Procedural Knowledge Test Score (out of 13)	12.50 ± 0.80	12.27 ± 1.78	12.86 ± 0.47
Conceptual Knowledge Test Score (out of 10)	2.69 ± 0.77	4.14 ± 1.84	5.98 ± 2.13
LP Retention Test GRS Score (out of 30)	17.00 ± 5.59	17.64 ± 4.26	17.82 ± 6.43
LP Transfer Test GRS Score (out of 30)	16.27 ± 5.89	15.52 ± 4.58	16.16 ± 5.53

Table 4-2. Regression Coefficients, Standard Errors, and Model Summary for Lumbar Puncture Retention. Model 1: D_1 and D_2 values are computed relative to the Procedural Only group (control); Model 2: D_1 and D_2 values are computed relative to the Integrated in Sequence group. Overall R^2statistics are shared for both models (they are identical models, but with different relative comparisons).

	***M* (Conceptual Knowledge Test Score)**				***Y* (Global Rating Scale Retention Performance)**							
Model 1		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (Integrated in Sequence)	a_1	1.44	.51	2.84*	c_1	.64	1.66	.38	c'_1	-.20	1.75	-.11
D_2 (Integrated for Causation)	a_2	3.28	.51	6.47*	c_2	.82	1.66	.49	c'_2	-1.08	2.12	-.51
M		-	-	-		-	-	-	b	.58	.41	1.42
Constant	i_1	2.70	.36	7.50*	i_2	17.00	1.17	14.50*	i_3	15.44	1.60	9.65*
Model 2		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (Procedural Only)	a_1	-1.44	.51	-2.84*	c_1	-.64	1.66	-.38	c'_1	.20	1.75	.11
D_2 (Integrated for Causation)	a_2	1.84	.51	3.63*	c_2	.18	1.66	.11	c'_2	-.88	1.81	-.49
M		-	-	-		-	-	-	b	.58	.41	1.42
Constant	i_1	4.14	.36	11.53*	i_2	17.64	1.17	15.0*	i_3	15.24	2.05	7.44*
		$R^2 = .40$ $F(2,63) = 21.03$, $p < .0001$*				$R^2 = .004$ $F(2,62) = .13$, $p = .875$				$R^2 = .04$ $F(3,62) = .76$, $p = .52$		

* denotes statistical significance with alpha < 0.05

Table 4-3. Regression Coefficients, Standard Errors, and Model Summary for Lumbar Puncture Transfer. Model 1: D_1 and D_2 values are computed relative to the Procedural Only group (control); Model 2**:** D_1 and D_2 values are computed relative to the Integrated in Sequence group. Overall R^2statistics are shared for both models (the are identical models, but with different relative comparisons).

	***M* (Conceptual Knowledge Test Score)**				***Y* (Global Rating Scale Transfer Performance)**							
Model 1		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (Integrated in Sequence)	a_1	1.44	.51	2.84*	c_1	-.750	1.62	-.464	c'_1	-2.54	1.60	-1.59
D_2 (Integrated for Causation)	a_2	3.28	.51	6.47*	c_2	-.114	1.62	-.070	c'_2	-4.19	1.94	-2.16
M		-	-	-		-	-	-	b	1.24	.37	3.33*
Constant	i_1	2.70	.36	7.50*	i_2	16.3	1.14	14.2*	i_3	12.9	1.46	8.85*
Model 2		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (Procedural Only)	a_1	-1.44	.51	-2.84*	c_1	.750	1.62	.46	c'_1	2.54	1.60	1.59
D_2 (Integrated for Causation)	a_2	1.84	.51	3.63*	c_2	.636	1.62	.39	c'_2	-1.65	1.65	-.10
M		-	-	-		-	-	-	b	1.24	.373	3.33*
Constant	i_1	4.14	.36	11.53*	i_2	15.5	1.14	13.6*	i_3	10.4	1.87	5.55*
	$R^2 = .40$ $F(2,63) = 21.03, p < .0001$*				$R^2 = .004$ $F(2,62) = .125, p = .883$				$R^2 = .16$ $F(3,62) = 3.80, p = .014$*			

* denotes statistical significance with alpha < 0.05

4.6 Discussion

We examined the role of integrating conceptual knowledge (why content) and procedural knowledge (how content) on skill retention and skill transfer of simulation-based LP. Our results demonstrate that, mediated by the positive impact on their conceptual knowledge, participants in both integrated instruction groups had better skill transfer (but not skill retention), compared to participants in the procedural only instruction group. Further, the transfer benefit of integrated instruction was significantly higher for participants when we interleaved how and why content and linked the two using causal explanations (i.e., Integrated for Causation), compared to when we presented conceptual knowledge first, followed by procedural knowledge (i.e., Integrated in Sequence). For all participants, greater conceptual knowledge was associated with higher LP GRS transfer scores, but not retention scores.

These data replicate our previous findings showing that integrating conceptual and procedural knowledge can improve participants' transfer of learning (Cheung et al., 2017). By comparing why content Integrated for Causation versus why content Integrated in Sequence, we show that teaching which explicitly promotes cognitive integration appears to help participants further mobilize their conceptual knowledge, which then enhances how well they transfer their learning (Kulasegaram et al., 2015). There are three key implications for educators who want to design education that promotes transfer of learning: 1) they need to consider trainees' level of conceptual knowledge; 2) the content they choose to teach matters and should include conceptual knowledge that explains the "why"; and 3) how they expose the relationships between the how and why content matters.

4.6.1 Implications for Cognitive Integration and Instructional Design

Contemporary instructional design recommendations frequently focus on how to deliver content (e.g., deliberate practice, provision of feedback, distributed practice, test-enhanced learning) (Cook et al., 2012; Motola et al., 2013; Van Merriënboer & Sweller, 2010), rather than on what content to deliver. Educators who focus only on practice structure without equal attention to content (especially conceptual knowledge) may not be optimizing the educational value of each repetition of practice (Eva, Neville, & Norman, 1998; Kulasegaram et al., 2017; Norman, 2009). Our results further establish the relationship between improving trainees' conceptual knowledge of a skill and improving their skill transfer. Our materials for teaching lumbar puncture illustrate principles educators can use to integrate conceptual knowledge into their unique instructional materials – namely teaching how and why content in close temporal proximity and employing causal explanations. When designing procedural skill learning activities, educators will likely benefit their trainees most by facilitating integration where it matters most: for skill transfer, at the level of trainees' cognition.

Similar to studies of cognitive integration in clinical reasoning (Kulasegaram et al., 2017; Kulasegaram et al., 2015), we found that carefully distinguishing how and why content, followed by selecting and organizing that content using principles from cognitive psychology results in instruction that helps trainees connect relevant clinical concepts with their procedural actions to support skill transfer. Our study replicates and extends the research done in clinical reasoning, where basic science knowledge serves as the conceptual knowledge that supports activities such as diagnoses. Our results show that basic science may now be considered a specific example of conceptual knowledge and the role of conceptual knowledge extends beyond reasoning tasks. Hence, selecting and organizing content to promote cognitive integration appears to benefit learning of both clinical reasoning (Bandiera Glen et al., 2017; Goldszmidt, Minda, Devantier, Skye, & Woods, 2012), and bedside invasive procedures.

4.6.2 Limitations and Future Directions

Our experimental study presents a mechanistic and theory-driven account of how integrated instruction relates to participants' conceptual knowledge, and to their skill transfer. Given this study is only the second in our program of research, further work can focus on testing how conceptual knowledge can best be delivered through various formats of instruction. For experimental control and efficiency, we used videos as the sole delivery format for integrated instruction. Other formats, such as via instructor feedback or through using hands-on simulator modules, may potentially enhance the observed skill transfer benefits. Educators might, for example, design simulator modules that allow trainees to experience conceptual knowledge in closer temporal, causal, and spatial proximity with their hands-on experience performing the procedural actions of a skill.

Though our present findings replicate results from previous work demonstrating the positive relationship between conceptual knowledge and skill transfer, they did not replicate the previously observed positive relationship between conceptual knowledge and skill retention (Cheung et al., 2017). We believe one issue is the timing of when our participants completed their procedural and conceptual knowledge tests. In our previous study, participants completed the tests both after their initial viewing of the instructional video and during the follow-up session, whereas participants in the present study completed the tests only at follow-up, which may have deprived them of the benefits of test-enhanced learning. This may have only affected participants' retention performance because test-enhanced learning is generally better for retention outcomes, rather than transfer outcomes (Eva, Brady, Pearson, & Seto, 2018). Future research could explore this finding by systematically examining how the timing of knowledge tests influence the mediating effect of that knowledge on performance.

4.6.3 Conclusion

Taken together, our findings suggest educators will benefit from considering content as they design procedural skills training, specifically how they can integrate relevant conceptual and procedural knowledge to support trainees' cognitive integration and skill transfer. Our results extend studies of clinical reasoning, demonstrating integrated instruction that encourages trainees to create linkages between how and why content also supports transfer of procedural skills.

Chapter 5

Material Concepts: Exploring Simulation as a Medium to Enhance Cognitive Integration and Skill transfer

5.1 Preamble

In this final study of the dissertation, the effects of cognitive integration established in the previous chapters, is bolstered through designing simulators themselves to serve as integrated instruction. The previous two studies demonstrated a positive relationship between conceptual knowledge and transfer (Chapter 3 and 4), and further, that this relationship was more pronounced when participants received integrated instruction designed for cognitive integration, i.e., “integrated for causation”, as opposed to other forms of non-causal integrated instruction, i.e., “integrated in sequence” (Chapter 4). Cognitive integration has proven to be a process that enhances simulation-based procedural skills transfer, and the delivery of causal explanations are optimal means to support it. To this point, however, the exploration of integrated instruction and its impact on conceptual knowledge and transfer has been restricted to video-based instruction.

Designing simulators themselves to serve as integrated instruction presents unique opportunities to engage trainees with causal explanations. Being a physical medium of instruction, simulators allows trainees to interact with the procedural and conceptual knowledge elements of a skill, engaging in an active exploration of the relationships between the two knowledges. However, given the emphasis of simulators on realism and fidelity, the use of simulators to achieve greater cognitive integration has yet to be explored, and is in fact a potential hinderance for integration. Take for example the relationship between needle technique and the underlying spinal anatomy. When trainees practice on a LP simulator, these spatial relationships are obscured by opaque simulated skin tissue, leaving trainees to intuit the hidden interactions that take place between their procedural actions with their needle and the invisible concepts of anatomy. What might happen if these relationships were literally made visible to trainees by using a transparent skin covering? Or perhaps even allowing trainees to peel this simulated tissue away as they inserted their needle?

In this study, trainees were liberated from the burden of being unable to see, feel, or experience the relationships between procedural actions and concepts due to simulation design, specifically the notion of fidelity. Is there any virtue in infidelity? Perhaps if the infidelity supports cognitive integration by making typically concealed procedural-conceptual relationships visible and tangible for trainees to experience? This idea was inspired originally by a participant in the previous study (Chapter 4) who attempted to remove the skin covering the spinal insert of the LP simulator. Though at the time it seemed like an attempt to breach the study protocol, it gave rise to a creative and exciting perspective on simulation instructional design, so, to this participant, thank you!

To operationalize simulator-based integration, previous study designs in Chapter 3 and 4 were modified to compare the impact of simulator-based integration to video-based integration on transfer. Simulator and video modules were design to deliver integrated instruction during a 1-hour simulation-based LP training scenario (same scenario delivered in the two previous studies). Both integrated instructional interventions were designed to provide relevant causal explanations in a just-in-time fashion – immediately prior to participants performing a critical procedural step of LP. To allow for meaningful comparisons with our previous study, the control group from the previous two studies (procedural only instruction) was included in the experimental design, as well as the same post-, retention, and transfer tests in the study protocol. A secondary goal of this research design was to delineate the impact of retention tests on transfer tests by counterbalancing the order of these tests during participants' follow-up visit in the study protocol.

The study in this chapter represents the first-time simulator/simulation instructional design has been manipulated to promote the integration of procedural and conceptual knowledges. The results replicated the previous two studies findings that greater conceptual knowledge was associated with better LP skill transfer, but the results showed no significant differences in conceptual knowledge or skill transfer between the two integrated groups. Further, there was a significant effect of transfer tests on subsequent retention test performance, but not the inverse. The discussion of these results focuses on: i) limitations of the study design that could be modified in future studies to more conclusively address the question of how simulator-based integration impacts learning retention and transfer; and ii) the educational possibilities for simulation instructional designs that meaningfully deviate from reality in the pursuit of transfer, as opposed to fidelity.

The paper in this chapter has yet to be submitted for publication but will target for a general journal in HPE. The topic of fidelity, realism, and authenticity are wide spanning and deep-seated ideals across a variety of areas in HPE (Norman, 2014), and will likely be of interest to a wide range of researchers and educators in the field. The research questions and lessons from our results point in a new direction for simulation-based instructional design that moves beyond voluminous and effortful practice. The direct focus on learning transfer, and its mechanisms, is also an under-studied area in the healthcare simulation literature. The simulator-based integrated instructional materials developed for the study serve as an example for educators that show creative and sophisticated instructional designs need not be technologically complex or realistic.

5.2 Abstract

Background: Instruction that encourages trainees to integrate conceptual "why" and procedural "how" knowledge improves their transfer of procedural skills. For training away from the bedside and direct supervision, questions remain on how to represent the causal relationship between clinical concepts and procedural actions (e.g., how patient anatomy relates to inserting a needle). Though the properties of simulation may offer unique strategies for integration, a zealous commitment to the notion of fidelity may inadvertently limit the effectiveness of simulation-based training by hindering learning mechanisms known to support skill transfer. We varied the modality and level of interactivity when presenting these causal relationships during simulation-based lumbar puncture (LP) training and measured impacts on participants' retention and transfer.

Methods: During a 1-hour session, we randomized 66 medical students to one of three instructional interventions: i) video-based procedural-only instruction, ii) integrated video-based instruction, and iii) integrated simulator-based instruction. One-week later, we tested participants' LP skill retention and transfer, and their conceptual knowledge on a written test.

Results: Simple mediation regression analyses revealed that participants receiving integrated instruction had superior LP retention and transfer skills via gains in conceptual knowledge (all $p < 0.01$). We found no significant performance differences between the integrated groups ($p > 0.01$). Participants receiving procedural-only instruction practiced significantly more LPs during training ($M = 2.36$) than participants receiving integrated video-based ($M = 1.82$) and simulator-based instruction ($M = 1.50$), $p < 0.05$.

Discussion and Conclusions: Trainees' ability to create cognitive connections between conceptual and procedural knowledge appears to improve when they interact with instructional materials highlighting the causal relationships between these knowledge types. Simulation experiences can be designed to make abstract clinical concepts visible using hands-on, interactive modules, which enhances trainees' conceptual knowledge, as well as, their skill retention and transfer. However, integrated instruction reduced participants' time to practice LP scenarios, which may have reduced the effectiveness of our efforts to promote such "cognitive integration". We suggest ways future research can better assess the potential benefits from such simulator-based integration.

5.3 Introduction

The common definition that healthcare simulation is a technique for replicating clinical reality (Gaba, 2004) implies that better learning outcomes will arise from simulations that better replicate reality. A review of the evidence however, demonstrates that "high-fidelity" simulations are not associated with trainees' improved application of skills to novel contexts, defined as their "transfer of learning" (Norman et al. 2012). This finding has prompted researchers to question whether simulation will ever improve learning, particularly when viewed only as a modality for delivering content through realistic experiences (Artino & Durning, 2012; Hamstra et al., 2014). To realize the full educational potential of simulation, educators and researchers must consider the learning processes that enhance trainees' transfer of learning and address how simulation design might be leveraged to enhance them.

Rather than pay fealty to fidelity by ensuring a procedure is practiced with a high degree of realism, research suggests that trainees' skill transfer may benefit from simulation experiences that encourage trainees to learn relevant conceptual knowledge. Conceptual knowledge consists of information about *why* the steps of a procedure are carried out in the demonstrated manner (e.g., teaching the underlying anatomy involved in a procedure). By contrast, procedural knowledge consists of information about *how* to effectively perform a procedure (e.g., demonstration of the steps of a bedside procedure). Instruction that integrates these two knowledges has been shown to support transfer of learning in various educational domains, including mathematics (Rittle-Johnson & Schneider, 2015), clinical reasoning (Castillo et al., 2018; Woods, 2007), and more recently in simulation-based procedural skills (Cheung, Kulasegaram, Woods, & Brydges, 2019; Cheung et al., 2018).

The benefit of integrated instruction for skill transfer is most prominent when instruction encourages trainees to create cognitive connections between procedural and conceptual knowledge, a process known as cognitive integration (Kulasegaram et al. 2013). One strategy for promoting cognitive integration involves using causal explanations to highlight the cause and effect relationship between procedural and conceptual knowledge. A series of experiments have demonstrated that instruction which integrates the pathophysiology (conceptual knowledge) as an explanation of the constellation of clinical signs and symptoms (procedural knowledge) enables trainees to better retain their diagnostic skill (Woods, 2007; Woods et al., 2005, 2007a;

Woods, Neville, et al., 2006), and to transfer their diagnostic reasoning skills either when presented with more difficult cases (Woods et al., 2007b), or when they are tasked with learning to diagnose new related diseases (Mylopoulos & Woods, 2014). Further, studies have also demonstrated that integrated instruction using causal explanations is superior to integrated instruction that solely creates temporal and spatial proximity between knowledges (Baghdady et al. 2009, 2013, 2014; Cheung et al. 2019; Kulasegaram et al. 2015; Lisk et al. 2016; Woods et al. 2006, 2007a). For example, presenting a block of text describing the pathophysiology of a disease (conceptual knowledge) followed by a block of text describing its clinical features (procedural knowledge) is less effective than providing a single block of text presenting both types of knowledge together using "science-based causal explanations" (Kulasegaram et al. 2015). Taken together, these findings suggest that instruction without explicit causal linkages between procedural and conceptual knowledge is often associated with trainees' having difficulty spontaneously engaging in cognitive integration, which limits their capacity to transfer their learning effectively.

Simulation as a modality for teaching presents unique opportunities to circumvent the limitations of reality in ways that may enhance cognitive integration, and further, transfer of learning. Unlike other educational modalities, simulation can literally make the relationships between procedural and conceptual knowledge visible to trainees (e.g., revealing underlying physical anatomy), allowing them to see and experience the relationships between procedural actions and abstract concepts. For instance, many medical procedures require trainees to identify anatomical landmarks (e.g., ribs) and insert needles through layers of tissue to access targeted areas. In this case, cognitive integration may be bolstered by simulations designed to allow learners to peer through anatomy (i.e., using transparent tissues), enabling them to see how procedural knowledge elements (e.g., their needle grip, position, angle, and depth) interact with conceptual knowledge elements (e.g., a patient's positioning during the procedure and the anatomical location of bones, ligaments, nerves, and blood vessels that should be avoided). By enabling such overt integration, such integrated instruction in simulation may represent a novel instructional approach to support the transfer of complex skills by encouraging learners' cognitive integration. To date, however, experiments have only examined the impact of integrated instruction delivered through text, audio, lectures, and videos, and not through simulation itself.

To explore the potential for simulation to enhance cognitive integration and support clinical expertise development, we compared the impact of three instructional interventions on Lumbar Puncture (LP) skill retention and transfer: i) video-based procedural only instruction (control group); ii) just-in-time video-based integration; and iii) just-in-time simulator-based integration. We also compared how each intervention affected participants' procedural and conceptual knowledge scores on written tests. We hypothesized a dose-response relationship between the extent the interventions supported cognitive integration and participants conceptual knowledge, retention, and transfer scores. Namely, we anticipated that the simulator-based integration would be superior to video-based integration, and that both integrated instructional groups would be superior to the control. Further, our previous research led us to anticipate that participants' gains in conceptual knowledge would mediate their improvements in skill retention and transfer.

5.4 Methods

5.4.1 Participants

Upon institutional ethical approval, we recruited 66 medical students to participate in the study. All participants were in their first or second year of medical school and had no formal training in LP. We used stratified randomization (Year of Study and Gender) to allocate participants into one of three instructional groups, each differing in their availability of conceptual knowledge, and the degree to which they were designed to supported cognitive integration. This included a i) procedural only group (n = 22), ii) video-based integration group (n = 22), and iii) simulator-based integration group (n = 22). Our sample size was based on our previous studies using the same learner population, skill, and outcome measures (Brydges et al., 2012; Cheung et al., 2019, 2018).

5.4.2 Learning Materials

To teach trainees to perform a LP, we adopted two instructional videos from our previous study: a procedural only video, and an integrated video (Cheung et al., 2019). The procedural only video presented procedural knowledge content about how to perform an LP (skill demonstration with procedural instructions). The integrated video presented the same procedural knowledge content but included conceptual knowledge in the form of causal explanations (why the procedural step is performed in the prescribed manner), which were interleaved throughout. To allow for hands-on self-regulated simulation-based LP training, we used a part-task trainer from Kyoto Kagaku and a training scenario derived from previous experimental studies of LP (Cheung et al., 2019, 2018; Haji, Cheung, Woods, Regehr, de Ribaupierre, et al., 2016).

We also developed a series of integrated instruction modules, aimed at further enhancing the cognitive integration of procedural and conceptual knowledge for LP by providing integrated instruction in a just-in-time fashion before participants performed steps of the procedure during simulation-based training. Working with staff clinicians who teach LP using simulation, we ensured the modules focused on steps of the LP procedure learners often encounter difficulty and where we hypothesized relevant conceptual knowledge would bolster understanding and subsequent performance. We created four modules that covered the following content areas:

1. Sterility
2. Three-way stopcock valve system

3. Cleaning (e.g., cleaning the intended puncture site for the LP)
4. Needle technique

These integrated learning modules were delivered either via video or via simulators. Video-based modules were short clips of video adopted from our previously used integrated instructional video. Simulator-based modules presented the same procedural/conceptual content through a hands-on activity designed to 'make concepts visible' for participants as they engaged with them. For example, in the Sterility module, to teach how to perform an LP while maintaining sterility, participants were instructed to retrieve a sterile Ping-Pong ball from within a sterile bowl with the help of a pair of sterile gloves – to make the concept of sterility visible to learners, we used finger paint to represent the contamination present on non-sterile equipment and participants' hands. Thus, rather than solely viewing a video demonstration with causal explanations, the simulator-based modules presented these causal relationships through a physical activity. Table 5-1 describes each module for both the video-based and simulation-based integration groups.

Table 5-1A. Description of the sterility modules. Administered prior to putting on gloves.

Video-Based Integration

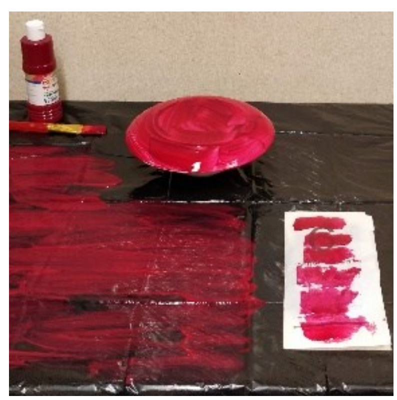
Describes how to put on sterile glove in a sterile fashion and why the steps are performed to prevent unwanted contamination. The segment explains why participants should only touch the inside portion of the glove with their non-sterile hands (because of contamination), and the need to equip sterile gloves prior to handling sterile tools inside the LP tray.

Clip duration: 0:54

Simulator-Based Integration

With the help of sterile gloves, learners are instructed to retrieve the Ping Pong ball from the bowl without getting any paint on it. All non-sterile surfaces are covered with finger paint (including the participants hands) to represent the concept of contamination.

Before contamination is made visible



After contamination is made visible

Table 5-1B. Description of the three-way stopcock modules. Administered prior to the preparing the manometer system.

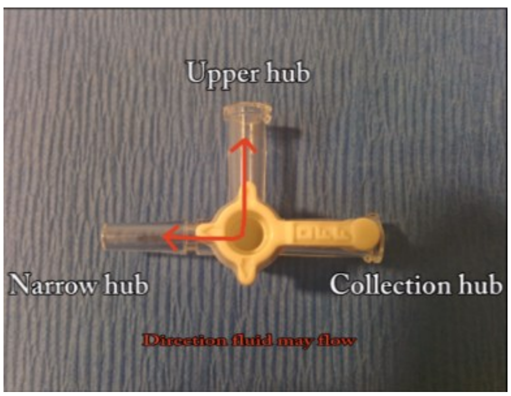
Video-Based Integration	**Simulator-Based Integration**
Describes the three-way stopcock as a valve system and demonstrates how the 'Off' dial can be adjusted to allow fluid to flow between two of the three hubs. All dial positions are presented.	Learners are instructed to move the air between all the syringes by adjusting the 'Off' dial of the three-way stopcock and squeezing the syringes. (E.g., move the air in the syringe from the left to the top, top to the right, and right to the left)
	
Clip duration: 1:00	

Table 5-1C. Description of the cleaning modules. Administered prior to picking up cleaning sponge stick from LP equipment tray.

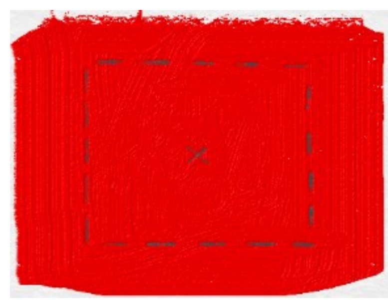
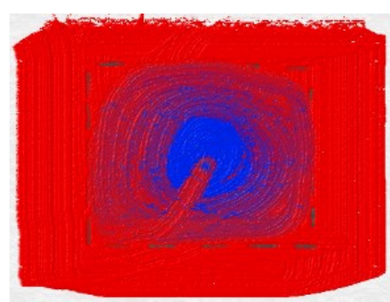
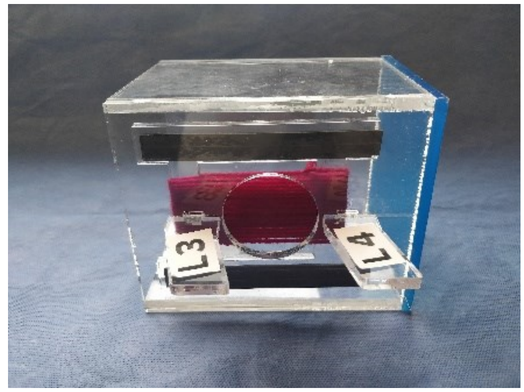
Video-Based Integration	**Simulator-Based Integration**
Describes the technique and importance of cleaning the puncture site to reduce the chance of infection. Introduces the concept of contamination and how the cleaning sponge stick is increasingly contaminated as it moves outward from the intended puncture site. Also explains that a wide area is cleaned to allow for the possibility of using an alternative puncture site if one should have difficulty with the initial puncture site. Clip duration - 1:11	On a touchscreen laptop, participants used a painting app to simulate cleaning of a puncture site denoted by the 'x'. Using their finger as a sponge stick dipped in 'blue' cleaning solution, they are instructed to clean from the puncture site outwards. Further, when they reach the outer edge of the dotted line, they are instructed to purposefully bring their finger back towards the 'x' – making visible the effects of cleaning inwards using a contaminated sponge stick  *Screen capture at beginning of cleaning task*  *Screen capture of completed cleaning task*

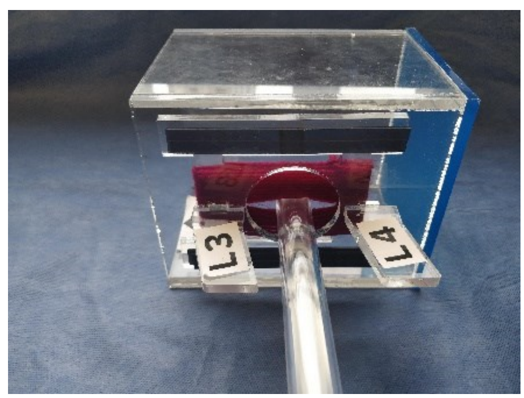
Table 5-1D. Description of the needle technique modules. Administered prior to using the LP needle.

Video-Based Integration	**Simulator-Based Integration**
Describes the bevel, needle shaft, and stylet components of a spinal needle and their relationship to patient anatomy (e.g., dural fibres and spinous processes) and the execution of the procedural steps of the LP. For example, the orientation of the bevel will either cut or spread apart dural fibres, the former of which can lead to excessive cerebrospinal fluid leakage and a severe headache for the patient. The video describes how to troubleshoot when they do not reach the thecal sac by pulling the needle to just below the surface of the skin and readjusting their angle. As well, the video describes the importance of a firm grip, needle control, and the possibility of performing the puncture either in a seated, crouching, or standing position as needed. Each step is accompanied by a causal explanation through the provision of conceptual knowledge. Clip duration - 3:57	Learners are given a large plastic needle and stylet to emphasize the hollow shaft of the needle, the bevel, and the needle's cutting edge.  *Needle (left) and stylet (right)*  *Needle with stylet inserted* A large plastic block representing the dural sac and spinous processes is used to illustrate

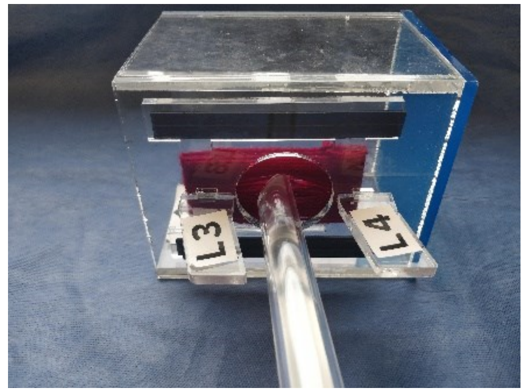
how the orientation of the large needle (bevel up or down), impacts dural fibres, which are represented by threads of yarn that can either be split apart when the bevel is parallel to the threads, or be damaged if not.



Representation of thecal sac, L3 and L4 spinous processes, and dural fibres (red yarn) that run parallel with the spine

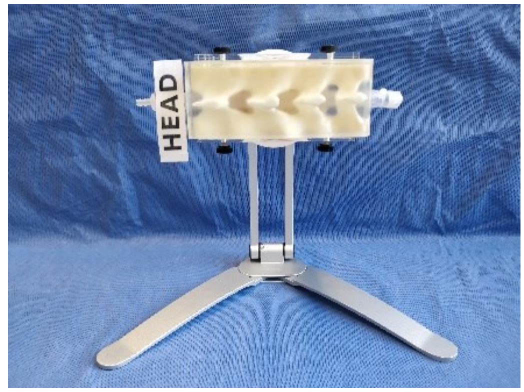


Parallel insertion causing dural fibres to spread apart



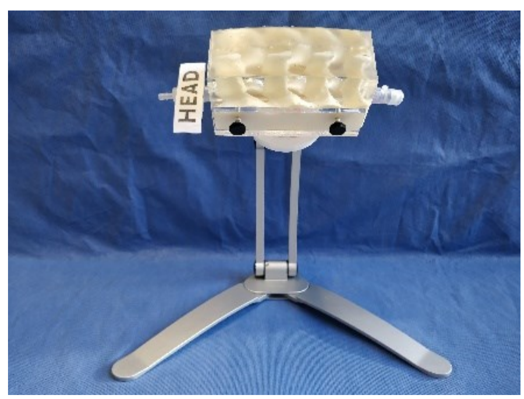
Perpendicular insertion causing trauma to the simulated dural fibres

Further, participants were tasked with using a real spinal needle to reach the thecal sac of a spinal block attached to a modified tablet stand. The spinal block is the same anatomical insert present in the lumbar puncture simulator used in the study training scenario. When participants were unsuccessful, the block was removed from the stand and participants could review their needle tract in the transparent tissue of the spinal block. After which they were given opportunity to adjust their technique and attempt another puncture.



Transparent spinal block on a modified tablet stand.

Learners were given a situation where the needle was hitting bone and instructed to redirect their needle (using the same puncture site) to reach the thecal sac. Finally, using the tablet stand, the angle of the block was adjusted to demonstrate the importance of body position (standing up/crouching) in certain situations.



The adjusted angle encourages learners to be aware of the relationship between their needle grip, angle, and body position.

5.4.3 Procedure

5.4.3.1 Initial Session

All participants individually attended two sessions in our laboratory: a training session and a follow-up session one-week later. At their training session, after providing informed consent and completing a demographic survey, participants reviewed one of two instructional videos for 25 minutes. Participants in the procedural only group reviewed the procedural only video, and participants in the integrated groups (video and simulator-based) reviewed the same integrated video.

After viewing their respective videos, participants began a 1-hour session of self-regulated simulation-based LP training. Participants were presented a scenario requiring them to perform a routine LP on a healthy patient in the lateral decubitus position. The scenario was ended when participants exceeded 3 needle passes (with 3 needle redirections for each needle pass) or if they successfully completed the LP procedure. If there was adequate time remaining in the 1-hour session, participants could attempt the same scenario again.

Participants in the procedural only group were immediately able to freely practice, alternating between performing the procedure in the scenario and reviewing their instructional video. During the first trial, participants in the integrated groups were stopped at predefined steps of the LP procedure to complete either their video-based or simulator-based learning module (i.e., just-in-time integrated instruction). For example, prior to putting on sterile gloves during the LP scenario, the research assistant would pause the scenario, and guide the participant to complete the Sterility module. For all trials after the first, participants in both integrated groups were able to allocate their time freely between performing the scenario or reviewing the integrated video, but could not return to the modules. At the end of the 1-hour, participants performed a post-test of the same LP scenario without access to any instructional materials.

5.4.3.2 Follow-up Session

At the follow-up session one-week after initial training, participants completed a retention and transfer test derived from our previous work (Cheung et al., 2019, 2018). The retention scenario was the same scenario as the training and post-test scenario. The transfer scenario consisted of a sick obese patient suspected of having meningitis and who could not tolerate being on their side, thus requiring the LP be performed with the patient sitting upright. In this scenario, we positioned the simulator in an upright position and inserted a spinal block with extra soft tissue to simulate obesity, making it more challenging for landmarking (via palpation) and needle technique. To assess for order effects of retention and transfer tests, testing order was counter-balanced across the groups.

Participants also completed written knowledge tests for procedural knowledge and conceptual knowledge, in that order. On the procedural knowledge test participants were tasked with sorting a list of procedural steps for LP in the appropriate order (scored out of 13). On the conceptual knowledge test, participants had to provide short-answers explanations to questions about why procedural steps were performed in the prescribed manner. Both tests were adopted verbatim from our previous study (Cheung et al., 2019).

All participants completed the study protocol individually with guidance from a research assistant who readout instructions and setup training and testing scenarios. During LP tests (post-retention and transfer), the research assistant video recorded participants. To preserve participant anonymity, all participants wore a gown and were given a pseudonym to use for each scenario. To avoid confounding effects of individualized instructor feedback, no feedback about participants' LP performance was provided until the end of the study.

5.4.4 Outcome Measures and Analyses

We computed all reliability analyses using G_String_IV (Bloch, 2013) and our descriptive and inferential statistics using SPSS Version 22 (IBM). To compare group outcome measures, we used Analysis of Variance (ANOVA) and Least Significant Difference post-hoc methods with our alpha set at 0.05. Where appropriate we applied Welch's ANOVA to account for lack of homogeneity of variance between groups (Levene's Test, $p < 0.05$).

5.4.4.1 Rater Training

To assess participants' test performances, three clinical experts in neurology (two PGY3 residents and one Fellow) were recruited to assess video-recordings, which provided a wide and close-up angle of the LP performance, presented in a split screen format. The wide angle was a stationary shot that captured the overall performance, while the close angle was operated by the research assistant and focused on the participants hands to capture nuanced LP needle technique.

Raters attended a 3-hour orientation session where they gained familiarity with the assessment process and tools. Raters first reviewed a task-specific LP checklist to orient them to the steps of the skill and then reviewed the primary study assessment tool of procedural competence, a global rating scale (GRS) (Martin et al., 1997). Raters then independently assessed three random video performances of LP from a previous study (Cheung et al., 2019), with discussion and deliberation after each performance until selecting a consensus score. The orientation session concluded with the raters agreeing that a 3 out of 5 score on the GRS meant participants were "capable of performing the procedure (or dimension of the GRS) independently without significantly compromising patient safety." As with our previous studies (Cheung et al., 2019, 2018), we selected the GRS as our primary outcome measure because of its strong validity evidence as a measure for procedural competence (Ilgen et al., 2015).

5.4.4.2 Participant Behaviours

During training, we recorded and compared the number of trials of the training scenario that all participants attempted. We also recorded the time participants in the simulation-based integration group spent completing the learning modules. We did not record the time participants spent reviewing their instructional videos.

5.4.4.3 Procedural and Conceptual Knowledge Tests

For the procedural knowledge test, the lead author (JC) independently scored each test out of 13. For the conceptual knowledge test, two expert raters (PGY3 and a Fellow, both in neurology) were given a scoring rubric and scored each test out of 13 points. We assessed inter-rater reliability for conceptual knowledge test scores by calculating an intraclass correlation coefficient (ICC), and found excellent relative agreement between raters, ICC (2,2) = 0.93. We compared the average scores for each group for both tests.

5.4.4.4 LP Skill Retention and Transfer

We decided to only examine impact of our interventions on participants' retention and transfer test performances based on our previous findings that post-test performance was not significantly impacted by integrated. Our analysis also evaluated the impact of testing order (i.e., whether participants completed the retention or transfer test first) on retention and transfer test GRS scores.

After the rater orientation, each video was scored by a random pair of raters. We used block-randomization to create an equal number of pairings across the three possible rater pair combinations $\binom{3}{2}$, which resulted in 44 common videos allotted to each rater pair. We assessed inter-rater reliability by calculating the intra-class correlation coefficient (ICC) for GRS retention and transfer scores. Each rater pair demonstrated high reliability, ICC $(2,2)_{\text{pair 1}} = 0.87$; ICC $(2,2)_{\text{pair 2}} = 0.90$; ICC$(2,2)_{\text{pair 3}} = 0.78$. Raters were blinded to group allocation and had no knowledge of the study design.

5.4.5 Mediation Analyses

To evaluate the relationship between our instructional interventions, participants' conceptual knowledge, and their LP performances (retention and transfer), we used simple mediation analyses with indicator coding via the PROCESS macro for SPSS (Hayes, 2013).

Using indicator coding, we computed two related regression models that represented path values relative to a defined control group. In one (Model 1 in Table 5-2), we set the procedural only group as the control for comparison to the video and simulator-based integration groups (D_1 and D_2, respectively). In a second coding scheme (Model 2 in Table 5-2), we set video-based integration group as our control for comparison to the procedural only and simulator-based integration groups (D_1 and D_2, respectively).

Table 5-2. Indicator coding for simple mediation analyses. Model 1 sets the procedural only group as a control, allowing for comparisons to the two integrated groups. Model 2 sets the video-based integration group as a control, allowing for comparisons to the two integrated groups.

		Indicator Coding		
Model		Procedural only	Video-based integration	Simulation-based Integration
1	D_1	0	1	0
	D_2	0	0	1
2	D_1	1	0	0
	D_2	0	0	1

Using simple mediation with indicator coding, we assessed the following relative effects of the study interventions (D_1 and D_2, Figure 5-1) on retention and transfer GRS performance (Y): i) the relative total effect (Path c_1 and c_2 in Figure 5-1A); ii) the relative indirect effects when conceptual knowledge test scores are included in the model as a mediator (Paths a_1 x b and a_2 x b in Figure 5-1B); and iii) the relative direct effects when the effect of conceptual knowledge is controlled for as a covariate (Paths c'_1 and c'_2 in Figure 5-1B). To determine relative indirect effects, we used the PROCESS macro to calculate 99% bias-corrected confidence intervals using 10,000 bootstrap samples; non-zero confidence intervals denote significance. Further, the analysis allows us to estimate the effect of our intervention on conceptual knowledge (Path a_1 and a_2 in Figure 5-1B) as well as the effect of conceptual knowledge, our mediator (M), on LP GRS performance (Y) (Path b in Figure 5-1B). We did not conduct mediation analysis with procedural knowledge test scores as a mediator based on our previous findings and theoretical framework.

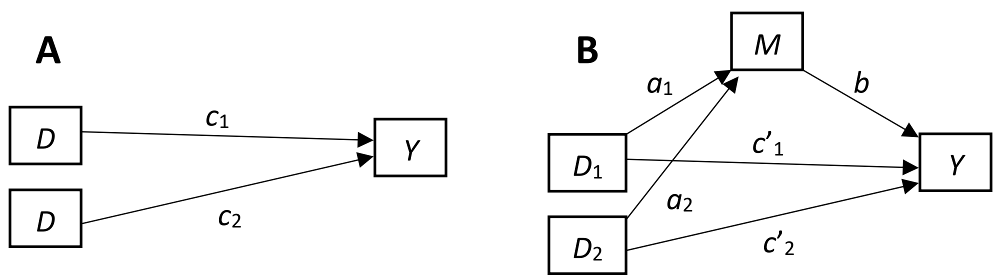


Figure 5-1. Mediation Analyses with Indicator Coding. A) Paths c_1 and c_2 represent the relative total effects of interventions D_1 and D_2 on outcome Y relative to a control group (not shown). B) M is included as a mediating variable, through which the interventions impact Y, and covariate. Paths a_1 and a_2 represent the effect of interventions D_1 and D_2 on outcome Y, paths c'_1 and c'_2 represent the relative direct effect of interventions D1 and D2 on outcome Y, after controlling for M as a covariate, path b represents the relationship of M and Y, and paths a_1 x b and a_2 x b represent the relative indirect effects of D_1 and D_2 on outcome Y, through their effects on M. [Figure adopted with permission from Cheung et al., 2019 with as per Chapter 4]

5.5 Results

Three participants rescheduled their follow-up session, which resulted in a delay of more than seven days from their initial training session. We found no evidence that their performance outcomes were outliers and decided to keep their data in all analyses. Mean (M) and standard deviation (SD) values for all outcome variables are summarized in Table 5-3.

5.5.1 Participant Behaviours – Differences in Time Allocation During Training

There was a significant difference in the number of trials attempted by group, $F(2,41.11) = 10.25$, $p < .001$. Post-hoc comparisons revealed that participants in the procedural only group attempted significantly more LPs during training ($M = 2.36$, $SD = .73$) compared to participants in the video-based integration group ($M = 1.82$, $SD = .50$), $p = 0.017$, 95% CI [.09, 1.00], and simulator-based integration groups ($M = 1.50$, $SD = .51$), $p < 0.001$, 95% CI [.40, 1.33].

Participants in the simulator-based integration group spent an average of 15 minutes and 23 seconds ($SD = 172.28$ seconds) completing their modules. For participants in the video-based integration group, though we did not measure time spent reviewing video-based learning modules, the length of these videos required a minimum time of 7 minutes and 2 seconds.

5.5.2 Procedural and Conceptual Knowledge Test Performance

For the procedural knowledge test, we found no significant differences between the groups (procedural only: $M = 12.64$, $SD = .95$; video-based Integration: $M = 12.59$, $SD = .67$; simulator-based integration: $M = 12.59$, $SD = .73$), $F(2,63) = .03$, $p = .98$. For the conceptual knowledge test, we found a significant difference between the groups, $F(2,40.91) = 16.98$, $p < .001$, $\eta_p^2 = .29$. Post-hoc comparisons revealed that the simulator-based integration group ($M = 5.77$, $SD = 1.52$) and video-based integration group ($M = 5.49$, $SD = 2.26$) scored significantly higher on the conceptual knowledge test than the procedural only group ($M = 3.26$, $SD = 1.47$), $p < .001$, 95% CI [1.41, 3.61]; and $p = .001$, 95% CI [0.82, 3.63] respectively; the difference between the two integration groups was not significant, $p = .88$.

5.5.3 Testing Order Effects: The Impact of Retention on Transfer, and Transfer on Retention

We found LP retention GRS scores for participants who performed the transfer test prior to the retention test (*M* = 17.42, *SD* = 5.58) were significantly higher than participants who performed the retention test prior to the transfer test (*M* = 14.00, *SD* = 4.50), $F(1,65) = 5.49$, $p = .022$, $\eta_p^2 = .009$. By contrast, for participants' LP transfer GRS scores, we found no difference between participants who first performed the retention test (*M* = 14.41, *SD* = 4.21) compared to participants who first performed the transfer test (*M* = 15.21, *SD* = 4.33), $F(1,65) = .58$, $p = .45$.

5.5.4 Mediation Analyses

All model coefficients for the mediation analyses are summarized in two tables, one for GRS retention performance as the outcome variable (Table 5-4) and one for GRS transfer performance as the outcome variable (Table 5-5). Looking at the relationship between group effects (D_1 and D_2) and conceptual knowledge (*M*), our mediation analyses show that, compared to participants in the procedural only group (Model 1 in Table 5-4 and 5-5), those in the video-based and simulator-based integration groups had significantly higher conceptual knowledge test scores (Model 1: $a_1 = 2.23$, $a_2 = 2.51$, respectively), both $p < .001$, 99% CI_{a1} [.79, 3.66], 99% CI_{a2} [1.08, 3.94]. For our Model 2 comparison, we found no significant difference in conceptual knowledge test scores between the video-based and simulation-based integration groups (Model 2: $a_2 = 0.28$), $p = 0.60$. These findings align with in our ANOVA group comparison of conceptual knowledge test scores reported above.

The mediation models also characterize the relationship between conceptual knowledge (*M*) and LP retention and transfer GRS scores (*Y*), regardless of participants' group assignment (path *b* in Figure 5-1B). We found a significant positive relationship between participants' greater conceptual knowledge and both improved retention GRS scores ($b = 1.03$), $p < .01$, 99% CI [.79, 3.66] and improved transfer GRS scores ($b = 1.06$), $p < .001$, 99% CI [.32, 1.79].

5.5.4.1 Relative Total and Direct Effects: Before and After Controlling for Conceptual Knowledge

For both LP retention and transfer test GRS scores, we found no significant differences in the relative total effects (paths c_1 and c_2 in Figure 5-1A), and no significant differences in the relative direct effects of our interventions (Paths after controlling for conceptual knowledge, paths c'_1 and c'_2 in Figure 5-1B), all $p > 0.05$ for relative comparisons to the procedural only group (Model 1 in Tables 5-4 and 5-5), and to the video-based integrated group (Model 2 in Tables 5-3 and 5-4).

5.5.4.2 Relative Indirect Effects: Conceptual Knowledge as a Mediator of Retention and Transfer Performance

For LP skill retention (Table 5-4) and transfer (Table 5-5), compared to participants in the procedural only group (Model 1) we found a significant relative indirect effect of participants' conceptual knowledge on their retention GRS scores (video-based integration: a_1 x b = 2.29, 99% CI [.32, 5.54]; simulation-based integration - a_2 x b = 2.58, 99% CI [.52, 5.83]), and on their transfer GRS scores (video-based integration: a_1 x b = 2.36, 99% CI [.54, 4.94]; simulation-based integration - a_2 x b = 2.66, 99% CI [.77, 5.26].

Compared to the video-based integration group (Model 2), we did not observe a significant relative indirect effect of participants' conceptual knowledge on retention GRS scores (a_2 x b = .29, 99% CI [-1.33, 2.35]) or transfer GRS scores (a_2 x b = .30, 99% CI [-1.35, 2.18]) for those participants in the simulator-based integration group.

Table 5-3. Mean group scores ± standard deviation for written procedural and conceptual knowledge tests, and Global Rating Scale scores for lumbar puncture performance on retention and transfer tests.

Outcome Measure	**Procedural Only Group**	**Video-Based Integration Group**	**Simulator-Based Integration Group**
Number of LP Trials Attempted at Training	2.36 ± .73	1.82 ± .50	1.50 ± 0.51
Procedural Knowledge Test Score (out of 13)	12.64 ± 0.95	12.59 ± 0.67	12.59 ± 0.73
Conceptual Knowledge Test Score (out of 13)	3.26 ± 1.47	5.49 ± 2.26	5.77 ± 1.52
LP Retention Test GRS Score (out of 30)	16.16 ± 4.69	15.43 ± 5.01	16.30 ± 6.12
LP Transfer Test GRS Score (out of 30)	14.77 ± 3.61	14.64 ± 4.13	15.02 ± 5.08

Table 5-4. Regression Coefficients, Standard Errors, and Model Summary for LP Retention test GRS scores. Regression values are presented for two models produced using indicator coding (reflecting a change in our control group used for comparison): i) Model 1, where D_1 and D_2 values are relative to the procedural only group (control); and ii) Model 2, where D_1 and D_2 values are relative to the video-based integration group.

	M **(Conceptual Knowledge Test Score)**				***Y*** **(Global Rating Scale Retention Performance)**							
Model 1 (Procedural Only as Control)		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (vs. Video-based integration)	a_1	2.23	.54	4.13*	c_1	-.73	1.60	-.45	c'_1	-3.02	1.71	-1.77
D_2 (vs. Simulation-based integration)	a_2	2.51	.54	4.66*	c_2	.14	1.60	.09	c'_2	-2.44	1.75	-1.40
M		-	-	-		-	-	-	b	1.03	.35	2.91*
Constant	i_1	3.26	.38	8.55*	i_2	16.16	1.13	14.28*	i_3	12.81	1.57	8.14*
Model 2 (Video-based Integration as Control)		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (vs. Procedural only)	a_1	-2.23	.54	-4.13*	c_1	.73	1.60	.45	c'_1	3.02	1.71	1.71
D_2 (vs. Simulation-based integration)	a_2	.28	.54	.53	c_2	.86	1.60	.54	c'_2	.57	1.51	.38
M		-	-	-		-	-	-	b	1.03	.35	2.91*
Constant	i_1	5.49	.38	14.39*	i_2	15.43	1.13	13.63*	i_3	9.79	2.22	4.42*
	$R^2 = .54$ $F(2,63) = 13.00$, $p <$ .0001*				$R^2 = .073$ $F(2,63) = .17$, $p = .85$				$R^2 = .12$ $F(3,62) = 2.94$, $p =$ .04*			

* denotes statistical significance with alpha < 0.05

Table 5-5. Regression Coefficients, Standard Errors, and Model Summary for LP Transfer test GRS scores. Regression values are presented for two models produced using indicator coding (reflecting a change in our control group used for comparison): i) Model 1, where D_1 and D_2 values are relative to the procedural only group (control); and ii) Model 2, where D_1 and D_2 values are relative to the video-based integration group.

	***M* (Conceptual Knowledge Test Score)**				**Y (Global Rating Scale Transfer Performance)**							
Model 1 (Procedural Only as Control)		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (vs. Video-based integration)	a_1	2.23	.54	4.13*	c_1	-.14	1.30	-.10	c'_1	-2.49	1.33	-1.87
D_2 (vs. Simulation-based integration)	a_2	2.51	.54	4.66*	c_2	.25	1.30	.19	c'_2	-2.41	1.36	-1.76
M		-	-	-		-	-	-	b	1.06	.28	3.84*
Constant	i_1	3.26	.38	8.55*	i_2	14.77	.92	16.04*	i_3	11.32	1.23	9.23*
Model 2 (Video-based Integration as Control)		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***		**Coeff.**	***SE***	***t***
D_1 (vs. Procedural only)	a_1	-2.23	.54	-4.13*	c_1	.14	1.30	.10	c'_1	2.49	1.33	1.87
D_2 (vs. Simulation-based integration)	a_2	.28	.54	.53	c_2	.39	1.30	.30	c'_2	.09	1.18	.07
M		-	-	-		-	-	-	b	1.06	.28	3.84*
Constant	i_1	5.49	.38	14.39*	i_2	14.63	.92	15.89*	i_3	8.83	1.73	5.11*
	$R^2 = .54$ $F(2,63) = 13.00$, $p < .0001$*				$R^2 = .038$ $F(2,63) = .04$, $p = .95$				$R^2 = .19$ $F(3,62) = 4.94$, $p = .003$*			

* denotes statistical significance with alpha < 0.05

5.6 Discussion

We sought to explore the potential benefits of integrated instruction delivered via hands-on simulator modules, hypothesizing that by designing simulators that make abstract concepts visible and material to trainees, we would further support cognitive integration compared to video-based integrated instruction and thus, would further improve participants' transfer of learning. Our mediation analyses, we found that integrated instruction resulted in increased conceptual knowledge, and that this increased conceptual knowledge was associated with improvements in both LP skill retention and transfer. However, our results did not show that simulator-based integration improved conceptual knowledge beyond the improvements associated with video-based integration. Though our findings do not align with our primary hypothesis, they do replicate our previous work showing that conceptual knowledge supports participants' skill retention and transfer, and further, they highlight cognitive integration as a potential learning mechanism which integrated instruction facilitates.

5.6.1 Implications for Cognitive Integration

Our findings add to our program of research aimed at exploring how instructional content and its delivery impacts trainees' understanding, and further, how this impacts skill retention and transfer. Our findings replicate previous work establishing the positive role integration of conceptual knowledge plays in simulation-based clinical skill retention and skill transfer (Cheung et al., 2019, 2018). Both integrated groups outperformed participants who received procedural only instruction; reiterating that the benefits of integrated instruction, as initially described in clinical reasoning, also apply to psychomotor skills taught in simulation. We also provide additional evidence that basic science is a specific instance of conceptual knowledge (Cheung et al., 2019). By demonstrating that integration of conceptual knowledge from various content areas (e.g., anatomy, sterility, patient safety, communication) enhances skill learning, we add evidence to the claim that integrating basic science improves clinical reasoning due to its ability to promote conceptual understanding (Bandiera Glen et al., 2017; Goldszmidt et al., 2012).

Further, by physically manipulating the modality of simulation itself to enhance trainees' cognitive integration while learning a psychomotor skill, we expand on previous research in clinical reasoning that has to date only manipulated texts, audio, and images. We suggest that our innovative design of simulator-based integrated instruction modules likely did not maximize the potential benefits for participants' cognitive integration. Our module design represents an example of how educators can use simulators to push the bounds of how instructional design can support cognitive integration. Continued innovation with simulators – both physical and virtual – offers new possibilities for revealing the deep conceptual structure of clinical problems and concepts to trainees, which has shown to be an important factor in improving trainees' skill transfer across education of various professional groups (Norman et al. 2007), mathematics (Kaminski, Sloutsky, & Heckler, 2008), skills training (C. H. Shea, Wright, Wulf, & Whitacre, 2000), and clinical reasoning (Kulasegaram et al. 2017).

5.6.2 Implications for Simulation Instructional Design

Current reviews of effective instructional design features when using healthcare simulation emphasize the repeated practice of procedural knowledge and provide little guidance on what elements of conceptual knowledge may help bolster learning (Cook et al., 2012; Issenberg et al., 2005; McGaghie et al., 2010; Sawyer et al., 2015). Our research findings show that conceptual knowledge plays a positive mediating role in enhancing retention and transfer, suggesting a refocusing for simulation instructional design to support mechanisms of learning, like cognitive integration. We suggest that our program of research(Cheung et al., 2019, 2018) collectively implies that simulation-based instructional designs emphasizing cognitive integration as a learning process to support, will help educators shift away from focusing on fidelity, or on using simulation as an opportunity to practice procedural steps 'correctly' and 'authentically'. Indeed, our work suggests that simulations that are most effective for learning may not be simulations that are most effective for replicating reality.

In using simulation-based assessments to enhance learning, our results suggest that transfer tests are more desirable than retention tests. We found that participants who completed their transfer test had improved GRS scores in their subsequent LP retention test compared to participants who completed their retention test first. We did not find the inverse of retention tests promoting subsequent transfer test performance. We reason that unlike retention tests, which test participants' ability to replicate and recall previous information, transfer tests require learners to adapt their previously acquired knowledge and skill, and thus, may bolster the mechanisms of test-enhanced learning (Larsen et al., 2008). Though transfer performance was typically worse than retention, this may be a “desirable difficulty” (Bjork, 1994) if enhanced future learning is prioritized over immediate performance (Schmidt & Bjork, 1992b). For education practice, this finding suggests that simulation-based assessments in curricula may be leveraged as learning opportunities most optimally when they assess for transfer of learning rather than retention. For research practice in simulation, it suggests that researchers may be better off focusing on a singular construct in their outcome measures (i.e., retention or transfer), as having multiple assessments in a study (even with counter-balancing) may reduce measurement precision.

5.6.3 Limitations and Future Directions

Though we replicated the positive benefits of integrated instruction observed in our previous work, we did not find simulator-based integration to be more effective than video-based integration. We believe the lack of difference may be explained by several factors. First, there may have been a misalignment between the benefits conferred by our integrated simulation-based learning modules and our assessments. For example, in the written conceptual knowledge test, a question asks participants: “Why do you clean the patients back in an outward fashion from the centre of the insertion site?” Scored out of a maximum of 1 point, there were a variety of answers would merit the participant full marks without requiring deep conceptual understanding that we hypothesized would arise from engaging in the simulation-based integration learning modules more than video-based integration learning modules. Similarly, our transfer test was not optimally aligned to assess the impact of gains in conceptual understanding for concepts beyond needle technique and anatomy. Though participant time and energy were devoted to learning modules that focused on understanding sterility, cleaning the puncture site, and the manometer - the transfer test only manipulated anatomy and patient positioning that were covered in the needle technique module. Thus, time spent on the other modules that promoted understanding of sterility, cleaning, and how a three-way stopcock functions may not have benefited skill transfer as operationalized in our study.

A second factor explaining the lack of superiority of simulation-based integration to video-based integration is participants’ time on task. The integrated modules took a significant amount of time away from practice during the training scenario. On average, participants in the simulation-based integration group had about 25% (15 minutes) of their 1-hour of practice taken up by their modules, which likely explains why participants in the procedural only group had significantly more attempts at the LP training scenario than participants in our integrated groups.

Future work to address these issues should aim to create greater alignment between assessments of conceptual knowledge, assessments of transfer, and the integrated instruction modules. We suggest that educators and researchers must characterize the conceptual knowledge trainees are acquiring from integrated instruction, and design comprehensive conceptual knowledge and transfer tests that assess trainees' ability to adapt their acquired knowledge and skill. We only spoke with clinical experts to design our modules, and we recommend that a more robust analysis will likely require exploring how trainees interact with the materials over time, as well as their perspectives on the knowledge they acquire from simulations designed to promote cognitive integration. As well, we recommend that future experimental studies standardize time on task across study groups. Researchers could accomplish this by controlling participants' number of trials during training rather than the total time of training.

5.6.4 Conclusions

Our findings confirm the relationship between greater conceptual knowledge and simulation-based skill retention and transfer, and the role integrated instruction can play in mediating this effect when designed to support cognitive integration. We believe that further innovation in the design of simulators for uniquely enhancing cognitive integration represents a potentially fruitful avenue of research in simulation-based training and assessment. By considering how simulation can be used to support cognitive integration and testing its impacts on learning outcomes, we have presented novel strategies for educators and researchers to explore in the pursuit of optimizing instructional design and skill transfer in health professions education. Taken together, this work suggests there is much to gain by considering how simulators can be used to support mechanisms already known to enhance learning and transfer, such as cognitive integration, rather than how simulators can be used to imitate reality.

5.7 Acknowledgements

The authors would like to thank Dr. Luke Devine, and Thomas Sun for their support in the development of our study learning modules. This study was made possible by the generous funding support from the Currie Fellowship program at The Wilson Centre, University Health Network that supported the first author's PhD studies (JJHC) and the Royal College Medical Education Research Grant that funded research expenses.

Chapter 6
General Discussion

6.1 Summary of Findings

The goal of this dissertation was to explore the potential role of integrating conceptual knowledge on simulation-based procedural skills retention and transfer. The three papers in this dissertation tested the hypothesis that greater conceptual understanding of a procedural skill would enhance skill retention and transfer, and that this would be optimally achieved through integrated instruction that supports trainees' cognitive integration. These studies clarify several questions pertaining to the relationship of integrated instruction, conceptual knowledge, and the retention and transfer of simulation-based procedural skills. Specifically, as described in Chapter 2, these studies were designed to address the following research aims:

1. To characterize the relationship between conceptual knowledge and skill transfer (Chapters 3, 4, and 5).

2. To examine the impact of integrated instruction (i.e., integrating conceptual knowledge with procedural knowledge) on skill retention and transfer (Chapters 3, 4, and 5).

3. To test whether the process of cognitive integration mediates the effectiveness of skill transfer in simulation-based procedural skills training. That is, integrated instruction that supports cognitive integration is more effective than integrated instruction that does not (Chapter 4 and 5).

4. To design and test novel simulation-based instruction to enhance cognitive integration and transfer (Chapter 5).

Accompanying these aims were hypotheses that were tested throughout the three studies:

1. Greater conceptual knowledge (but not procedural knowledge) would be associated with improved skill retention and transfer.

2. Integrated instruction would improve skill retention and transfer through its positive effect on trainees' conceptual knowledge.

3. Compared to integrated instruction that merely presented both procedural and conceptual knowledge, integrated instruction that supports cognitive integration of these two

knowledges would better enhance conceptual understanding, and thus, skill retention and transfer.

4. Procedural and conceptual knowledge serve the same cognitive function as clinical and basic sciences in clinical reasoning; that is, basic science represents a subset of conceptual knowledge and clinical science represents a subset of procedural knowledge.

5. Simulation itself (e.g., simulators) can be designed to accentuate the causal linkages between procedural and conceptual knowledge and thereby support skill retention and transfer.

Hypothesis 1 was explored in all three studies, which all demonstrated a significant positive relationship between participants' conceptual knowledge and their retention and transfer performances. Two of three studies (Chapter 3 and 4) showed no significant relationship between procedural knowledge and retention or transfer; however, the final study did (Chapter 5). It is unclear why this relationship was present in the final study. The finding may relate to changes in our procedural protocol, namely the reduction in time allotted for self-regulated LP practice in lieu of video and simulator-based modules. By increasing the variability of practice time and thus the potential learning of the procedural knowledge fundamentals for LP skill, this manipulation may have increased the sensitivity of the procedural knowledge test to detect individuals who had poorer or superior performances in both retention and transfer. These findings were not discussed in the manuscript version of the article (Chapter 5) as they were not part of the primary analyses. However, they provide evidence that the just-in-time interventions used in Chapter 5 resulted in quantifiably different learning experiences as compared to the previous two studies (Chapter 3 & 4), where there was no such relationship between procedural knowledge and LP retention or transfer performance. Another possible explanation for this finding is measurement error of the procedural knowledge test; namely the lack of variance due to ceiling effect, which results in a higher false positive significance rate (Austin & Brunner, 2003).

Hypothesis 2 was explored and confirmed in all three studies through the use of mediation analysis. All studies found a significant indirect effect for integrated instruction on skill transfer. Two of three studies (Chapter 3 and Chapter 5) also found this effect on skill retention. Taken together these findings show that, through its positive impact on trainees' conceptual knowledge, integrated instruction facilitates procedural flexibility (i.e., transfer) and likely procedural skill reproduction as well (i.e., retention).

Hypothesis 3 and 4 were first tested in Chapter 4, which compared instruction integrated for causation (i.e., designed to promote cognitive integration) to instruction integrated in sequence (i.e., including both procedural and conceptual knowledge without making explicit their causal relationships). Mediation analyses demonstrate that participants who received the instruction integrated for causation had higher conceptual knowledge scores than those who received the instruction integrated in sequence; which in turn enhanced transfer performance. This was not the case for participants' retention performance, where there was no significant difference in total or indirect effects of integrated instruction compared to the control or each other. This may have resulted from changes in our experimental protocol that reduced the exposure to the procedural and conceptual knowledge tests. Rather than receiving the tests twice (as in the Chapter 3 study), participants received a single exposure to the written tests at the end of the study protocol after they completed their LP retention and transfer tests. The results confirmed the superiority of integrated instruction designed for cognitive integration and extended findings from clinical reasoning literature to psychomotor skills. The findings establish cognitive integration as a desirable cognitive activity that supports the learning process in skills other than clinical reasoning. Further, the data reveal that cognitive integration operates on procedural and conceptual knowledges, of which the clinical and basic science knowledges are respectively subsets of.

By comparing simulator-based integration modules to video-based integration modules, the experiment in Chapter 5 also aimed to test Hypotheses 3 and 4. Simulator-based integration modules were designed to amplify cognitive integration through increased interactivity and vividness of causal explanations. In doing so, the study i) explores possibilities for delivering effective integrated instruction through simulation instructional design (i.e., integrated instruction that may more effectively support cognitive integration); and ii) further tests the nature of knowledges that makes for effective integration (i.e., beyond the basic and clinical sciences examined in diagnostic reasoning). However, given methodological limitations in the study design, it is unclear if simulator-based integration had its intended effect. Though the theoretical framework gives reason to assume simulator-based integration was superior in supporting cognitive integration than video-based integration, the simulator-based interventions also took more time for participants to complete, which reduced the amount of practice they had during their 1-hour of simulation-based LP training. Thus, it is challenging to draw any firm conclusions regarding the superiority of simulator-based integration to video-based integration given the non-equivalence in the amount of LP practice. Nonetheless, trainees' conceptual knowledge was positively associated with both LP skill transfer and retention, supporting Hypothesis 1. Furthermore, when comparing the integrated instructional groups (both video and simulator-based) to the control group, the results again reveal a positive indirect effect of integrated instruction on skill retention and transfer that is mediated by conceptual knowledge; supporting Hypothesis 2.

The learning materials developed in Chapter 5 also provide some insights to Hypothesis 5. The comparable performances of participants across all groups, despite the reduction of almost 25% in overall training time for the integrated groups, suggests that conceptual knowledge gains offset some of the disadvantages of reduced hands-on procedural LP practice. Further, though learner perceptions of learning interventions provide an incomplete understanding of instructional effectiveness, the simulator-based integration modules were well received by participants in the study – who saw potential in their utilization. Once again, firm conclusions are challenged by the limitations to the study methodology. However, what is clear from the study is that there are exciting possibilities for novel and creative instructional designs that leverage simulator-based integration, which should be refined and tested in future work.

In summation, the findings of this dissertation contribute the following to our understanding: i) they characterize the role of conceptual knowledge in the transfer and retention of simulation-based procedural skills; ii) they establish the benefits of the integration of conceptual knowledge for psychomotor skills taught in simulation; iii) they reveal that integrated instruction is most effective when it facilitates cognitive integration; in other words, cognitive integration is the mechanism through which integrated instruction improves skill retention and transfer; iv) in the context of integration in clinical reasoning, they demonstrate that basic and clinical sciences are subsets of more general types of knowledge, specifically conceptual and procedural knowledge respectively; and v) they provide examples of novel instructional design strategies for facilitating cognitive integration that educators and researchers may apply and study.

6.2 Implications for Simulation Instructional Design and Research

6.2.1 The Role of Conceptual Knowledge in Procedural Skills Training

The findings of this dissertation suggest causally integrated conceptual and procedural knowledge is an instructional design feature that supports skill transfer. Integrated instruction, and the role of conceptual knowledge in particularly, is currently missing from instructional design guidelines for simulation-based training. All three studies revealed a positive association between trainees' conceptual knowledge and skill transfer. Two of the three studies also showed a positive association between conceptual knowledge and skill retention. Instructional design features that increase trainees' conceptual knowledge of procedural skills, such as integrated instruction, may be important to incorporate into simulation-based training.

To optimally improve transfer, effective integration of conceptual knowledge must happen from cognition up, not curriculum down. The results from Chapter 5 demonstrate that the effectiveness of integrated instruction for simulation-based procedural skills training was dependent on how well the instructional materials encouraged trainees to create causal connections between procedural and conceptual knowledge, i.e., how well integrated instruction facilitated cognitive integration. Separating these two knowledges by presenting them in sequence (conceptual knowledge first, followed by procedural knowledge) reduced improvement in trainees' conceptual knowledge, and further reduced the benefit of integrated instruction on trainees' transfer performance. These findings mirror the original findings from clinical reasoning, which demonstrate that effective integration (that improves learning outcomes) is a cognitive activity. Specifically, it is cognitive integration that makes the addition of conceptual knowledge useful for trainees, improving skill retention and transfer (Kulasegaram et al., 2013). To facilitate cognitive integration in simulation-based procedural skills training, evidence is accumulating to suggest that integrated instruction should be designed to encourage trainees to make connections between their procedural actions when performing a skill and the underlying concepts that explain *why* they are being performed. Current efforts to achieve 'curricular integration' of simulation into curricula are a good start (Cook, Hamstra, et al., 2013), but to reap the full benefit of integration, educators and researchers must begin to consider 'cognitive integration' *within* simulation-based training experiences at the session level, in addition to the integration that is designed at the curricular level (Goldman & Schroth, 2012).

According to the theoretical premise of cognitive integration as explored in this dissertation, its skill retention and transfer benefits should also extend beyond simulators and simulation scenario design, to how trainees are prepared for simulation, given feedback, and debriefed. An example of evidence in support of this hypothesis can be found from a body of experimental studies exploring how trainees are prepared for simulation-based training. These studies demonstrate that activities that elaborate trainees' conceptual knowledge can enhance skill retention outcomes; for example, instructional activities that have trainees compare and contrast video-recordings of errorful and flawless performances (Cheung et al., 2016; Domuracki et al., 2015) and activities that encourage collaboration and interaction with peers using web-based learning platforms (Cheung et al., 2016, 2013; Grierson, Barry, Kapralos, Carnahan, & Dubrowski, 2012). In contrast, though simulation feedback and debriefing have received much attention and proven to 'work', understanding of the elements that make feedback effective are only beginning to be elucidated (Cheng et al., 2014; Cook et al., 2011; Fanning & Gaba, 2007; Raemer et al., 2011). Similar to the simulation instructional design literature more generally, the majority of current feedback and debriefing recommendations focus on their structure and organization (e.g., duration, educator presence and characteristics, timing, content area, use of video) (Cheng et al., 2014). Hence, current recommendations providing little guidance for educators about the types of knowledge that feedback and debriefing should imparting to learners, or even how debriefing relates the transfer of learning (Rivière, Jaffrelot, Jouquan, & Chiniara, 2019). The findings from this dissertation suggest that similar to the design of simulation and simulator instruction, feedback and debriefing in HPE may benefit from an emphasis on the procedural and conceptual knowledges they contain, and careful consideration of how their presentation can be coordinated to facilitate cognitive integration.

The predictive value of conceptual knowledge for procedural skills retention and transfer also has implications for simulation-based assessment and how educators incorporate simulation into curricula. Compared to simulation-based assessments, which take large amounts of time to design, coordinate, conduct, and assess – conceptual knowledge assessments are a relatively faster and more cost-effective means for educators to repeatedly test trainees and gain a sense of their procedural flexibility. Leveraging non-simulation assessments of conceptual knowledge (e.g., written tests) may help educators more efficiently gauge trainees' readiness for subsequent procedural skills training in simulation. Further, using conceptual knowledge assessments can support conceptual knowledge development through the benefits of test-enhanced learning, which can be defined as a form of *retrieval practice* whereby trainees engage in the active mental processes of recall that result in the construction of more robust memory structures for the recalled information (Larsen et al., 2012, 2008; McDaniel, Roediger, & McDermott, 2007). Though testing procedural skills in simulation have shown to improve skill retention (Kromann, Jensen, & Ringsted, 2009), conceptual knowledge assessments may be more cost-effective and benefits specifically for transfer. A study in clinical reasoning has already demonstrated that tests of trainees' basic science knowledge (a subset of conceptual knowledge) improve trainees' diagnostic skill retention (Baghdady et al., 2014b).

Finally, these findings suggest that greater coordination of simulation and non-simulation instruction within curricula may enhance cognitive integration and thus skill transfer for procedural skills. By considering when, where, and how conceptual knowledge is taught and assessed – educators can create more robust opportunities for trainees to acquire and integrate the relevant procedural and conceptual knowledges required to become adaptive experts. This applies not only to the instructional design of simulation and non-simulation sessions, but also assessments. For example, prior to a simulation-based LP training session, teaching and testing conceptual knowledge about spinal anatomy, tool usage, and patient safety can help an educator gauge their trainees' preparation for simulation but also may serve to improve trainees' subsequent LP learning outcomes.

6.2.2 Beyond Deliberate Practice

In contrast to instructional guidelines that emphasize repeated deliberate practice, the findings of this dissertation suggest that trainees' expertise development depends on *what* types of knowledge are included in instruction, and how these knowledges are then related to one another. For simulation educators, this diverts attention from creating simulation experiences that are focused on having trainees faithfully reproduce the procedural steps of a skill, to creating simulation experiences that encourage trainees to integrate the relevant knowledges that will develop expertise and flexibility with the procedural skill. Hence, in addition to ensuring trainees can practice its procedural steps in simulation, my work suggests that there is value in also teaching the underlying conceptual knowledge that is common across a variety of procedural skills. For example, though the individual steps of each bedside procedure may be different, the concepts of sterility and landmarking are present across many bedside procedures and represent types of conceptual knowledge that can help trainees retain and transfer their simulation-based procedural skills.

Given the constraints on curricula time that limit the feasibility of voluminous deliberate practice, integrated instruction presents a potential means to accelerate expertise development and prepare trainees to more effectively retain and transfer their clinical skills. Educators cannot teach trainees everything, and thus, approaching simulation-based instructional design with the aim of developing integrated knowledge serves as a framework to guide educators in making difficult decisions about what to teach and what to exclude in their curricula while optimizing what counts – transfer. By integrating conceptual knowledge to help trainees make meaning of their procedural actions during simulation-based procedural skills training, educators may set their trainees on a course toward becoming an adaptive expert; preparing them to innovate novel solutions for novel problems they will inevitably face in their training and practice as health professionals. Considering the knowledges that enable adaptive expertise and transfer offers educators a means to work within the limitations of curriculum, moving beyond deliberate practice.

6.2.3 Infidelity and Transfer

Though a chorus of authors have described the minimal relationship between simulation fidelity/realism and transfer (Hamstra et al., 2014; Norman et al., 2012), this dissertation offers the first program of research to suggests that there may be virtue in purposefully deviating from realism to enhance both skill retention and transfer outcomes. The results in Chapter 5 reveals the potential for simulation design to demonstrate positive deviance from the realities of clinical practice when these deviations support trainees' cognitive integration. The simulator-based integration modules explored in Chapter 5 represent instructional elements that "are not 'realistic' uses of simulators but may be effective for learning" (Dieckmann et al., 2007). Hence, in addition to contributing further evidence of the poor relationship between fidelity and transfer (i.e., the modules would be judged as quite low fidelity by most), this work presents a framework to educators and researchers for manipulating simulation fidelity to leverage cognitive processes, namely cognitive integration, that enhances desirable learning outcomes.

Though the integrated instructional videos and modules developed in this study are specific to LP, the principles of designing integrated instruction that supports cognitive integration are applicable across all procedural skills. The video and simulator-based instructional modules provide worked-examples of strategies for encouraging cognitive integration by presenting causal relationships between procedural and conceptual knowledge in ways that make these connections explicit, timely, visible, and even material. These examples and the principles of their design serve to expand the toolkit of instructional design features that simulation educators have to work with to better improve trainees' learning outcomes.

Taking the premise of this dissertation to its logical conclusion, a lack of fidelity in simulation no longer becomes a weakness, but rather a potential strength. The ability for simulators and simulations to be unrealistic, affords simulation educators an opportunity to create unreal learning opportunities that would be challenging to achieve in real-world practice. Just as error-ful practice can provide unique opportunities for trainees to learn, so too can unrealistic simulations. For example, research on observational practice prior to simulation-based training has shown observations of errorful demonstrations are more effective than errorless demonstrations (Cheung et al., 2016; Domuracki et al., 2015). Similarly, performing a task under unrealistic and inauthentic conditions may provide opportunities for trainees to make connections between different knowledges involved in a skill that would be difficult under more authentic learning conditions, simulated or real. During a real-life LP, for instance, the chance to scrutinize the relationship between needle technique and a real patients' underlying anatomy might improve learning, but would not be feasible under real clinical conditions.

Try as we may, by definition, a simulation of reality will never capture all aspects of reality. The very definition of simulation as an imitation of aspects of reality precludes this. However, this need not limit the ingenuity, creativity, and sophistication of simulation educators seeking to enhance learning outcomes from healthcare simulation. High-fidelity and realism may have been one avenue to vent these creative energies, but now developing integrated instruction to help trainees make visible key concepts in material and coherent ways can be a new frontier in simulation design, education, and research – one that may suggest realism is a restraint rather than an ideal to strive for.

6.2.4 Mediation Analyses in Experimental Research

The mediation analyses employed throughout the dissertation represent a unique contribution to the field of healthcare simulation and HPE research. These analyses allow for mechanisms of simulation-based instructional interventions (i.e., indirect effects through a mediator variable) to be statistically evaluated, as well as the total effects on learning outcomes. They emphasize understanding of the learning processes and mechanisms of an intervention, as opposed to a sole focus on outcomes (i.e., whether an intervention 'worked'). Mediation analyses allow for nuances of interventions to be captured, but also require that researchers have an *a priori* learning mechanism and mediator variable (conceptual knowledge in the case of this dissertation) built into their study design. Such analyses help encourage theory-driven inquiry about the active ingredients that make a simulation-based intervention effective, about how those ingredients may be measured as mediator variables, and further, they provide statistical assessments of the impact of multiple interacting variables. For example, with regards to the evaluation of simulation realism, mediation analyses can assist empirical research in clarifying how various modes of realism (physical, semantical, and phenomenal) relate to educational learning outcomes. Several mediator variables could explain why working to align realism with education goals might improve learning outcomes – such as improved learner engagement (Rudolph et al., 2007), or improved motivation (Artino & Durning, 2012). Mediation analyses and the theory-driven thinking that underpins the approach can address empirical questions about if, why, how, and when these facets of simulation realism may influence the learning process (e.g., could trainees' learning outcomes be better because of the increased motivation that they experience as a result of being part of a more *phenomenally* realistic simulation?). Further work is required to clarify the active ingredients of *physical*, *semantical,* and *phenomenal* realism that impact trainees learning and their resulting educational outcomes. The descriptive frameworks for simulation realism introduce language to describe instructional 'levers' that educators can manipulate (physical, semantical, phenomenal) – but more important to consider is what these levers are actively doing to impact trainees' learning and their simulation-based training outcomes – mediation analyses may support such empirical clarification research.

An advantage of mediation analyses for the purposes of simulation research is its relative efficiency. Rather than having to sample so they can ensure studies are statistically powered to detect small overall effects of simulation-based interventions on learning outcomes, researchers can explore indirect effects of interventions through mediators; which provides greater statistical power (Hayes, 2013). Hence, smaller samples sizes can be used in addressing questions of mediation effects. Given the resource intensive nature of simulation research and calls for researchers to engage in more rigorous and controlled comparisons of active educational interventions (Cook, 2010, 2012; Norman, 2003), mediation offers researchers a more powerful tool to more effectively study the factors of simulation instructional design that affect its educational effectiveness. Further, as discussed above, mediation analyses encourage thinking at the level of learning processes, which can help educators consider the ‘hidden’ influences their simulation-based instructional designs may have on trainees, help them use theory to define and measure these influences, and help them iteratively refine their simulation-based interventions (Haji, Da Silva, Daigle, & Dubrowski, 2014).

6.3 Implications for Expertise and Knowledge Integration in Health Professions Education

6.3.1 Adaptive Expertise and Cognitive Integration

The integrated instruction delivered in each of the three studies reaffirms models of adaptive expertise, which suggest that procedural fluency, in this case assessed by transfer tests, will arise from a combination of procedural and conceptual knowledge. These findings provide evidence that the model of adaptive expertise predominantly studied in the context of mathematics education (Baroody & Dowker, 2003; Rittle-Johnson & Schneider, 2015), translates to clinical procedural skills taught using simulation. Furthermore, by devising and testing novel ways of integrating procedural and conceptual knowledge through the medium of simulation, our work presents an alternative to merely sequencing these knowledges temporally across a teaching session. Rather, we showed benefits for trainees' retention and transfer outcomes when a coordinated learning activity (e.g., simulator-based integration modules) highlighted the relationships between these two knowledge types. For instructional designs that intend to support adaptive expertise development, rather than address whether procedural or conceptual knowledge should be taught first or second (Rittle-Johnson et al., 2015), the more appropriate question of content organization becomes: will the causal relationships between procedural and conceptual knowledge be clear to the trainee? It is not only a matter of ordering the types of knowledge in instruction, but about helping trainees appreciate their cause and effect relationships. Hence, our results create stronger linkages between adaptive expertise and research on cognitive integration (Mylopoulos & Woods, 2014, 2017; Steenhof, Woods, Van Gerven, & Mylopoulos, 2019), marrying cognitive theory described in clinical reasoning with educational theory on the knowledges that enable adaptive expertise.

The use of interactive physical simulators to enhance cognitive integration provides a novel and unique approach to delivering integrated instruction for many categories of clinical skills. For many skills, procedural or conceptual knowledge may be best integrated through these interactive physical modules as opposed to text or audio-visual mediums. For example, when performing an LP, subtle differences in the initial angle and direction of a needle can lead to large deviations from the targeted space, which lies several centimeters beyond the point at which the needle breaks the skin. Verbally describing the nuances of angle, depth, and force are likely suboptimal in articulating these relationships as compared to physical demonstrations or physical practice with an integrated simulator-based learning module. Further, by operationalizing instructional strategies that support the integration of conceptual and procedural knowledge using physical simulators, this dissertation opens new avenues for research to examine the nature and role of procedural and conceptual knowledge for many clinical skills that involve the physical interaction and manipulation of objects in the environment.

6.3.2 Redefining Basic and Clinical Science

By providing evidence that basic science and clinical science function as subsets of the broader knowledge categories of conceptual and procedural knowledge, this dissertation presents a new typology for scholars and educators to conceptualize and operationalize knowledge integration in their own contexts. It is not only basic science that can assist trainees in becoming adaptive experts capable of skill transfer, but any form of knowledge that serves the function of conceptual knowledge, that is, knowledge that presents trainees with an organizing framework to understand *why* a skill or task is performed in the recommended fashion. Basic science appears to work not because it is a 'basic science', but rather, because it serves as conceptual knowledge that connects procedural knowledge elements of clinical skills so that they make sense in the minds of trainees (Woods et al., 2007a).

This more inclusive interpretation of basic science echoes calls in the clinical reasoning literature for its definition to be expanded to include areas of knowledge outside of the traditional biological sciences (Bandiera et al., 2013; Bandiera Glen et al., 2017; Goldszmidt et al., 2012; Kuper & D'Eon, 2011). However, rather than expanding the definition of basic science, conceptual knowledge may be a more appropriate and precise terminology to employ. In addition to opening the discourse of integration to other forms of conceptual knowledge (for example, from the social sciences and humanities), using the broader term conceptual knowledge would avoid mental gymnastics involved with recategorizing the social sciences and humanities under the flag of 'basic science'. And at the same time, it would allow more common interpretation of basic sciences (i.e., as physical, chemical, and biological sciences) to be preserved.

6.3.3 Simulation as a Bridge Between Theory and Practice

Expertise in HPE involves complex problem-solving, where multiple solutions may exist for any given problem and the amount of interrelated knowledge that must be comprehended is high (van

Merriënboer & Sweller, 2005). If developing such expertise requires the effective integration of procedural and conceptual knowledge, there may be few educational modalities better to achieve this integration than simulation. In addition to the advantages of repeated low-stakes practice without immediate risk to patient safety, simulation may confer an advantage when it comes to supporting the cognitive integration of clinical skills, which often require a combination of psychomotor and cognitive skills. There may be critical interactions that exists between these skills that are difficult to convey in either the clinical environment, where practical procedural skills are typically learned (e.g., where procedural equipment is stored and how to obtain it in a timely manner through teamwork), or the classroom environment, where cognitive skills are typically taught (e.g., diagnosis from clinical presentation). Hence, integrated instruction delivered through simulation can give life to the interrelatedness of these skills and their underlying knowledges. Simulation, when designed for the purpose of integration, could serve as a bridge between the various theories and practices trainees are currently acquiring from disparate educational environments.

Whether integrated instruction using simulation will be more efficient or effective than the current systems of clinical skills training is unclear. However, what is clear is the need for more systematic and accountable training. Rather than leaving this to chance and opportunity, educators and researchers can work toward creating HPE curricula that prepare trainees to appreciate the relationships between conceptual knowledge and procedural knowledge, or more colloquially, between theory and practice. This dissertation suggests that simulation may play an important part in achieving this aspiration.

6.4 Limitations

"Easy-to-do science is what those in physics, chemistry, geology, and some other fields do. Hard-to-do science is what the social scientists do and, in particular, it is what we educational researchers do. In my estimation, we have the hardest-to-do science of them all! We do our science under conditions that physical scientists find intolerable. We face particular problems and must deal with local conditions that limit generalizations and theory building—problems that are different from those faced by the easier-to-do sciences"

- David C. Berliner (2002)

Several limitations challenge our interpretation of our findings. To avoid redundancy, this section will aim to synthesize the limitations mentioned to this point, with a focus on the general challenges of working in the liminal space between psychomotor and cognitive skills, and between the realities of educational practice and educational/cognitive psychology theory.

Education research is a messy, complex business. The studies in this dissertation operationalized constructs from education and clinical reasoning in the context of simulation-based instructional design. In doing so, decisions were made which resulted in necessary trade-offs between theoretical generalizability and practical relevance. These trade-offs are apparent in the design of the specific instructional interventions and assessments. While the control and alignment of instructional interventions, study protocols, and assessments across the three studies can be considered a strength, these constraints also limited our exploration of integrated instruction, conceptual knowledge, and skill transfer.

In terms of the major limitations of the instructional interventions and study protocols, all the experiments restricted participants' simulation-based LP training to one-hour of self-regulated practice with free access to review an assigned instructional video. These decisions attempted to balance experimental control with practical relevance and logistical constraints. Whether the one-hour was an adequate duration for trainees to benefit fully from integrated instruction is unclear. Perhaps an extra half hour would have significantly enhanced the effectiveness of participants receiving integrated instruction across the studies. Furthermore, the studies did not account for the varying time that trainees spent reviewing the instructional videos during their LP practice. The time assigned to viewing the videos was particularly noticeable in study 3 (Chapter

5), in which the just-in-time integrated video-based and simulator-based modules during the one-hour led to noticeable differences in the number of full LP attempts participants in the different experimental conditions were able to attempt during training. Further, being fully self-regulated without external feedback from an instructor, it is unclear whether some of the mundane struggles that participants experienced in learning to perform an LP created unnecessary noise in the study results. For example, several participants were uncertain about where they were landmarking on the simulator, spending several minutes palpating erroneous landmarks on the simulator, far from the desired region. For the sake of experimental control, the study protocol prohibited the research assistant from guiding individuals toward the appropriate site for landmarking. However, this may have also created unintended complications in assessing the efficacy of the various instructional interventions.

The choices made regarding the assessments also created limitations to the study implications and generalizability. In particular, the iterative nature of the studies restrained major adjustments to our assessments of conceptual knowledge and transfer across studies. Though attempts were made to bolster the sensitivity of the written conceptual knowledge test in study two (Chapter 4), modifications were additions rather than restructuring of the testing format, which consisted of short answer questions asking participants to explain the rationale behind particular steps of the LP procedure (Appendix II). Greater sensitivity in the conceptual knowledge test was of particular concern in the third study (Chapter 5), where the findings revealed no differences in conceptual knowledge between the video-based and simulator-based integration groups. Furthermore, in operationalizing transfer, the same assessment scenario was used throughout the dissertation. This scenario differed from the training, post-test, and retention test scenarios by changing the position of the simulated patient (seated rather than laying on their side) and their anatomy (obese rather than normal tissue thickness). These differences represent only one dimension of conceptual manipulation – anatomy/needle technique. Thus, it is unclear whether other transfer test designs may have also been more sensitive in detecting differences between integrated and unintegrated instructional groups. These limitations in the conceptual knowledge and transfer assessments may have resulted in reduced power to detect the effects of the various manipulations of integrated instruction tested in this dissertation.

Another limitation was the exclusion of data regarding participant experiences. None of the studies formally probed participants about their experiences with the interventions or assessment scenario and study protocol. Though interventions were piloted with novice students to allow for improvement and refinement of data collection, formal collection of data about how trainees interacted with their instruction could provide rich insights to why features of integrated instruction supported learning or why it might not. These insights would be particularly useful in any further refinements of integrated simulator-based instruction, which, as discussed in Chapter 5, may have suffered from a misalignment with our conceptual knowledge and transfer assessments.

Chapter 7
Conclusions

This dissertation advances the study of simulation-based instructional design by clarifying how *what* is taught in simulation impacts trainees' ability to transfer their learning. Applying theory from psychology, education, and clinical reasoning, three studies explored the relationship between integrated instruction, conceptual knowledge, and the simulation-based learning outcomes of skill retention and transfer for LP, a prototypical bedside procedure. In doing so, this dissertation addresses gaps in current instructional design evidence, which offer little guidance about how the nature of the content delivered using simulation-based training might optimally support retention and transfer. First, evidence from Chapter 3 demonstrates integrated instruction, which causally links procedural and conceptual knowledge, improves novice trainees' conceptual knowledge, and further, that gains in trainees' conceptual knowledge result in improved skill retention and transfer. Second, findings from Chapter 4 reveal that integrated instruction is most effective for improving trainees' conceptual knowledge and skill transfer when it is designed to encourage trainees to create causal, mechanistic connections between procedural and conceptual knowledge, a process termed cognitive integration. Third, Chapter 5 provides an example of how simulator design can be reimagined to support cognitive integration by purposefully leveraging simulation infidelity to further emphasize to trainees the connections between procedural and conceptual knowledge.

Taken together, the studies establish an important role for conceptual knowledge instruction in simulation-based procedural skills training. Integrated instruction, designed to support cognitive integration, optimizes the acquisition of conceptual knowledge, which further supports trainees' skill retention and transfer. Thus, this dissertation extends research findings from clinical reasoning on the cognitive integration of basic science and clinical science. By demonstrating the efficacy of cognitive integration in the realm of psychomotor skills taught in the context of healthcare simulation, this work also suggests that basic science and clinical science knowledge are subsets of conceptual and procedural knowledge, respectively. The more general categorization of conceptual and procedural knowledge has the potential to expand the domains of clinical skills training for which cognitive integration might serve as an underlying learning mechanism.

Though advances in simulation expand the teaching possibilities for healthcare educators, curricula cannot teach trainees all possible permutations of clinical problems; and thus, all components of curricula must instead focus on preparing trainees using educational experiences that are linked to improved transfer. Educators must inevitably make decisions about what to teach and what to assume trainees will figure out when the time comes. These decisions should be guided by an understanding of the procedural and conceptual knowledge components of clinical expertise that prepare trainees to be adaptive experts. The evidence gathered in this dissertation suggests that integrated instruction containing these knowledges should be designed to encourage trainees to engage in cognitive integration. Further, the examples of integrated instruction presented in this dissertation present simulation infidelity as a simulation-specific strategy to facilitate trainees' conceptual knowledge, their cognitive integration, and their learning transfer. Such integrated instruction represents a novel instructional design feature for future researcher to elaborate upon and for simulation educators to refine, both with the aspiration of uncovering how simulation can further enhance trainees' skill transfer.

Chapter 8
Future Directions

8.1 Revisiting Simulator-Based Integration

Future research should continue exploring and clarifying the mechanisms that facilitate cognitive integration in the context of healthcare simulation. The logical first step in this program of research would revisit simulator-based integration as explored in Chapter 5. Though findings in Chapter 5 support the relationship between conceptual knowledge and skill retention and transfer; they did not demonstrate the hypothesized superiority of simulator-based integration compared to video-based integration. As discussed in Chapter 5, there are a number of methodological limitations in the study design that could be addressed in future studies to replicate or challenge this finding.

A clear methodological choice would be to change the practice schedule for the different groups by using a trial-based protocol, rather than the one-hour block of LP practice time. Trial-based study interventions have been operationalized in previous studies of LP where the number of trials each participant attempts is standardized while time is unstandardized (but measured) (Haji, Cheung, Woods, Regehr, de Ribaupierre, et al., 2016). This manipulation would help equalize the time on task (i.e., practicing the LP training scenario), which was significantly imbalanced between the groups in Chapter 5.

A second adjustment to the study design would be to replace the initial integrated instructional video with a procedural only instructional video for all groups. This would make the comparison of integrated (video-based or simulator-based) instruction more precise, restricting integrated conceptual knowledge to only the just-in-time content delivered during simulation-based LP training. Further the time spent on these modules could be measured for both groups and used as a potential covariate in statistical comparisons.

A third, and more drastic change would be to embed the simulator-based integration modules within the LP training scenario itself, meaning that rather than being directed away from the training scenario to complete a learning module, the modules would be built *into* the LP training scenario/simulator itself. For example, the needle technique module could be adapted into the actual LP simulator (used to perform the training scenario) by simply removing the skin covering the simulator; and the sterility module could potentially be adopted by applying finger paint to the actual simulator.

Though all these ideas will need refinement prior to execution, there is certainly the potential to create a stronger alignment between the simulator-based integration modules and procedural practice of how to perform the LP training scenario. And in the process of refinement, this work would further clarify the mechanisms of cognitive integration and strategies for simulation-based instruction to enhance it.

8.2 Assessing Conceptual Knowledge and Transfer

The success of any educational intervention can only be as effective as the quality of its outcome measures. To fully appreciate the effects of integrated instruction, there may be alternative approaches to assessing our primary outcome measures of trainees' conceptual knowledge and skill transfer. Conceptual knowledge might be assessed using different types of written testing such as multiple-choice questions or concept maps that are developed to capture a wider range of conceptual knowledge (Hay, 2007). There may even be opportunities to explore the more general relationship between prior knowledge (in various domains of knowledge) on trainees' skill retention and transfer of a variety of skills. Recent work has demonstrated, for example, that medical trainees from a humanities or social science undergraduate background score higher on assessments of communication skill (Hirshfield, Yudkowsky, & Park, 2019). The same research principle could be applied to procedural skill performance within curricula using cohort study designs. For example, to assess the impact of an undergraduate block on anatomy of the spine on lumbar puncture, a study could compare the LP performance of trainees who have yet to complete the anatomy block with another cohort of trainees who have. Framing conceptual knowledge in this broader sense creates opportunities to ask many questions that consider knowledge taught across all components of the curriculum as factors affecting trainees' ability to learn and transfer skills taught during simulation-based training.

Another outcome measure that could be refined or enhanced would be the transfer test. There are other scenario designs that could be used to test trainees' capacity to demonstrate skill transfer. As discussed in Chapter 6.4 on Limitations, the transfer test utilized in all three studies of this dissertation manipulated only one dimension of conceptual knowledge, which was anatomy and its impact on needle technique. Future research studies could design simulated LP transfer scenarios that vary along more dimensions of conceptual knowledge. By increasing the alignment between the assessment of transfer and the hypothesized benefits of conceptual knowledge, such transfer tests might be more sensitive in showing that trainees with greater conceptual knowledge (i.e., received integrated instruction) perform better than trainees who received only procedural only instruction, improving the statistical power of the study to detect a significant group difference. Further, transfer can also be conceptualized as a building of knowledge and skill as trainees acquire new related skills, termed preparation for future learning (PFL) (Bransford & Schwartz, 1999; Mylopoulos et al., 2016). Assessments of PFL (i.e., the double transfer design) could be used to determine how well integrated instruction in simulation-based training prepares trainees to learn new skills in the future. For example, how well can trainees learn to perform other needle-based procedures like paracentesis, after having first learned LP with integrated versus unintegrated instructional materials? Future studies in simulation could begin to explore these relationships of transfer across skills, as well as within skills.

8.3 Participant Perceptions of Integrated Instruction

Another line of inquiry that future research could examine is the perception of trainees as they engage with integrated instruction. Formal data collection of trainees' perspectives and interpretations of these instructional materials can inform efforts to scale integrated instruction across curricula, and also provide valuable feedback for iterative refinements of simulation interventions and study protocols. Important questions to address pertain to the acceptability of simulator-based integration, particularly if such instructional designs were to look and feel significantly different from "the real thing"; and the perceived value of conceptual knowledge as trainees perform the transfer assessment scenario. Though trainee perceptions of educational value have shown to align poorly with actual educational value (Uttl, White, & Gonzalez, 2017), trainees' response processes to such interventions can play a large role in both the effectiveness of integrated instruction in supporting cognitive integration, and their overall uptake and demand.

8.4 Educator Use of Integrated Instruction

Just as instructional materials change what is possible for trainees to practice and observe, so too do the possibilities for educators to express and demonstrate ideas. For experimental control, the three studies in the dissertation restricted any external feedback from an educator. This decision precluded any formal observations of the interaction between an educator and student in environments where instruction was integrated versus unintegrated. A future line of research could work to clarify how integrated instruction impacts educator teaching and feedback practices. A study could explore the quantity, method, and types of explanations educators provide trainees with access to varying types of instructional material. For example, how might the inclusion of an anatomical spinal model in the room augment the way an educator describes how to perform an LP? Further, a research study could observe experienced versus inexperienced educators in simulation-based procedural skills, noting differences again in the quantity, method, and types of explanations (or procedural knowledge instruction) that are presented. Or perhaps even the instructional materials that are selected by these educators based on their experience teaching particular skills. Theoretically, such studies could also address questions about the ideal "depth" or quantity of conceptual explanations to foster understanding and facilitate skill retention and transfer. They could illustrate how experienced educators make decisions about when and how to teach conceptual knowledge, as well as how much. The insights from such studies would help clarify not only how integrated instructional methods impact trainees, but also how educators utilize integrated instruction, and how their educational practices can be altered by including instructional artefacts that can be used in the service of integration.

References

Aggarwal, R., & Darzi, A. (2006). Technical-skills training in the 21st century. *The New England Journal of Medicine*, *355*(25), 2695–2696. https://doi.org/10.1056/NEJMe068179

Ahya, S. N., Barsuk, J. H., Cohen, E. R., Tuazon, J., McGaghie, W. C., & Wayne, D. B. (2012). Clinical performance and skill retention after simulation-based education for nephrology fellows. *Seminars in Dialysis*, *25*(4), 470–473. https://doi.org/10.1111/j.1525-139X.2011.01018.x

American Board of Internal Medicine. (2016). Internal Medicine Policies. Retrieved from http://www.abim.org/certification/policies/internal-medicine-subspecialty-policies/internal-medicine.aspx

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 932–945. http://dx.doi.org.myaccess.library.utoronto.ca/10.1037/0278-7393.23.4.932

Artino, A. R., & Durning, S. J. (2012). ‘Media will never influence learning’: But will simulation? *Medical Education*, *46*(7), 630–632. https://doi.org/10.1111/j.1365-2923.2012.04270.x

Austin, P. C., & Brunner, L. J. (2003). Type I Error Inflation in the Presence of a Ceiling Effect. *The American Statistician*, *57*(2), 97–104. Retrieved from JSTOR.

Baghdady, M. T., Carnahan, H., Lam, E. W. N., & Woods, N. N. (2014a). Dental and Dental Hygiene Students' Diagnostic Accuracy in Oral Radiology: Effect of Diagnostic Strategy and Instructional Method. *Journal of Dental Education*, *78*(9), 1279–1285.

Baghdady, M. T., Carnahan, H., Lam, E. W. N., & Woods, N. N. (2014b). Test-enhanced learning and its effect on comprehension and diagnostic accuracy. *Medical Education*, *48*(2), 181–188. https://doi.org/10.1111/medu.12302

Baghdady, M. T., Carnahan, H., Lam, E. W., & Woods, N. N. (2013). Integration of basic sciences and clinical sciences in oral radiology education for dental students. *Journal of Dental Education*, *77*(6), 757–763.

Baghdady, M. T., Pharoah, M. J., Regehr, G., Lam, E. W. N., & Woods, N. N. (2009). The Role of Basic Sciences in Diagnostic Oral Radiology. *Journal of Dental Education*, *73*(10), 1187–1193.

Bandiera, G., Boucher, A., Neville, A., Kuper, A., & Hodges, B. (2013). Integration and timing of basic and clinical sciences education. *Medical Teacher*, *35*(5), 381–387. https://doi.org/10.3109/0142159X.2013.769674

Bandiera Glen, Kuper Ayelet, Mylopoulos Maria, Whitehead Cynthia, Ruetalo Mariela, Kulasegaram Kulamakan, & Woods Nicole N. (2017). Back from basics: Integration of science and practice in medical education. *Medical Education*, *52*(1), 78–85. https://doi.org/10.1111/medu.13386

Baroody, A. J. (2003). The Development of Adaptive Expertise and Flexibility: The Integration of Conceptual and Procedural Knowledge. In A. J. Baroody & A. Dowker (Eds.), *The*

Development of Arithmetic Concepts and Skills: Constructing Adaptive Expertise (pp. 1–33). Retrieved from http://site.ebrary.com/id/10084525

Baroody, A. J., & Dowker, A. (Eds.). (2003). *The Development of Arithmetic Concepts and Skills: Constructing Adaptive Expertise*. Retrieved from http://site.ebrary.com/id/10084525

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An Alternative Reconceptualization of Procedural and Conceptual Knowledge. *Journal for Research in Mathematics Education*, *38*(2), 115–131. https://doi.org/10.2307/30034952

Barsuk, J. H., Cohen, E. R., Feinglass, J., Kozmic, S. E., McGaghie, W. C., Ganger, D., & Wayne, D. B. (2014). Cost Savings of Performing Paracentesis Procedures at the Bedside After Simulation-based Education. *Simulation in Healthcare*, *9*(5), 312. https://doi.org/10.1097/SIH.0000000000000040

Barsuk, J. H., Cohen, E. R., Feinglass, J., McGaghie, W. C., & Wayne, D. B. (2009). Use of Simulation-Based Education to Reduce Catheter-Related Bloodstream Infections Simulation Catheter Training to Reduce Infections. *Archives of Internal Medicine*, *169*(15), 1420–1423.

Barsuk, J. H., Cohen, E. R., McGaghie, W. C., & Wayne, D. B. (2010). Long-Term Retention of Central Venous Catheter Insertion Skills After Simulation-Based Mastery Learning: *Academic Medicine*, *85*, S9–S12. https://doi.org/10.1097/ACM.0b013e3181ed436c

Barsuk, J. H., Cohen, E. R., Potts, S., Demo, H., Gupta, S., Feinglass, J., … Wayne, D. B. (2014). Dissemination of a simulation-based mastery learning intervention reduces central line-associated bloodstream infections. *BMJ Qual Saf*, *23*(9), 749–756. https://doi.org/10.1136/bmjqs-2013-002665

Barsuk, J. H., McGaghie, W. C., Cohen, E. R., Balachandran, J. S., & Wayne, D. B. (2009). Use of simulation-based mastery learning to improve the quality of central venous catheter placement in a medical intensive care unit. *Journal of Hospital Medicine*, *4*(7), 397–403. https://doi.org/10.1002/jhm.468

Bédard, J., & Chi, M. T. H. (1992). Expertise. *Current Directions in Psychological Science*, *1*(4), 135–139. https://doi.org/10.1111/1467-8721.ep10769799

Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Chicago ; La Salle: Open Court.

Berliner, D. C. (2002). Comment: Educational research: The hardest science of all. *Educational Researcher*, *31*(8), 18–20.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA, US: The MIT Press.

Bloch, R. (2013). G_String_IV (Version 6.3.8) [Windows]. Retrieved from http://fhsperd.mcmaster.ca/g_string/index.html

Boet, S., Bould, M. D., Fung, L., Qosa, H., Perrier, L., Tavares, W., … Tricco, A. C. (2014). Transfer of learning and patient outcome in simulated crisis resource management: A systematic review. *Canadian Journal of Anesthesia/Journal Canadien d’anesthésie*, *61*(6), 571–582. https://doi.org/10.1007/s12630-014-0143-8

Bordage, G. (2009). Conceptual frameworks to illuminate and magnify. *Medical Education*, *43*(4), 312–319. https://doi.org/10.1111/j.1365-2923.2009.03295.x

Boyd, V. A., Whitehead, C. R., Thille, P., Ginsburg, S., Brydges, R., & Kuper, A. (2018). Competency-based medical education: The discourse of infallibility. *Medical Education*, *52*(1), 45–57. https://doi.org/10.1111/medu.13467

Boyle, E., O'Keeffe, D. A., Naughton, P. A., Hill, A. D. K., McDonnell, C. O., & Moneley, D. (2011). The importance of expert feedback during endovascular simulator training. Retrieved from Journal of vascular surgery: Official publication, the Society for Vascular Surgery [and] International Society for Cardiovascular Surgery, North American Chapter website: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=21636241

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC, US: National Academy Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of Research in Education* (Vol. 24, pp. 61–100). Retrieved from http://www.jstor.org/stable/1167267

Broudy, H. S. (1977). Types of knowledge and purposes in education. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the Acquisition of Knowledge* (pp. 1–17). Hillsdale, NJ: Erlbaum.

Brydges, R., Carnahan, H., Rose, D., Rose, L., & Dubrowski, A. (2010). Coordinating Progressive Levels of Simulation Fidelity to Maximize Educational Benefit. *Academic Medicine*, *85*(5), 806–812. https://doi.org/10.1097/ACM.0b013e3181d7aabd

Brydges, R., Hatala, R., Zendejas, B., Erwin, P. J., & Cook, D. A. (2015). Linking Simulation-Based Educational Assessments and Patient-Related Outcomes: A Systematic Review and Meta-Analysis. *Academic Medicine*, *90*(2), 246–256. https://doi.org/10.1097/ACM.0000000000000549

Brydges, R., Nair, P., Ma, I., Shanks, D., & Hatala, R. (2012). Directed self-regulated learning versus instructor-regulated learning in simulation training. *Medical Education*, *46*(7), 648–656. https://doi.org/10.1111/j.1365-2923.2012.04268.x

Carter, B. N. (1952). The fruition of Halsted's concept of surgical training. *Surgery*, *32*(3), 518–527. https://doi.org/10.5555/uri:pii:0039606052902195

Case, R. (1991). The Anatomy of Curricular Integration. *Canadian Journal of Education / Revue Canadienne de l'éducation*, *16*(2), 215–224. https://doi.org/10.2307/1494974

Castillo, J.-M., Park, Y. S., Harris, I., Cheung, J. J. H., Sood, L., Clark, M. D., … Woods, N. (2018). A critical narrative review of transfer of basic science knowledge in health professions education. *Medical Education*, *52*(6), 592–604. https://doi.org/10.1111/medu.13519

Chamberland, M., Mamede, S., St-Onge, C., Rivard, M.-A., Setrakian, J., Lévesque, A., … Rikers, R. M. J. P. (2013). Students' self-explanations while solving unfamiliar cases:

The role of biomedical knowledge. *Medical Education*, *47*(11), 1109–1116. https://doi.org/10.1111/medu.12253

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

Cheng, A., Eppich, W., Grant, V., Sherbino, J., Zendejas, B., & Cook, D. A. (2014). Debriefing for technology-enhanced simulation: A systematic review and meta-analysis. *Medical Education*, *48*(7), 657–666. https://doi.org/10.1111/medu.12432

Cheung, J. J. H., Koh, J., Brett, C., Bägli, D. J., Kapralos, B., & Dubrowski, A. (2016). Preparation with web-based observational practice improves efficiency of simulation-based mastery learning. *Simulation In Healthcare: The Journal of the Society for Simulation in Healthcare*, *(Accepted) SIH-D-15-00083R2*.

Cheung, J. J. H., Koh, J., Mackinnon, K., Brett, C., Bägli, D., Kapralos, B., & Dubrowski, A. (2013). The use of web-based learning for simulation-based education and training of central venous catheterization in novice learners. *Studies in Health Technology and Informatics*, *184*, 71–77.

Cheung, J. J. H., Kulasegaram, K. M., Woods, N. N., & Brydges, R. (2019). Why Content and Cognition Matter: Integrating Conceptual Knowledge to Support Simulation-Based Procedural Skills Transfer. *Journal of General Internal Medicine*. https://doi.org/10.1007/s11606-019-04959-y

Cheung, J. J. H., Kulasegaram, K. M., Woods, N. N., Moulton, C., Ringsted, C. V., & Brydges, R. (2017). Knowing How and Knowing Why: Testing the effect of instruction designed for cognitive integration on procedural skills transfer. *Advances in Health Sciences Education*. https://doi.org/10.1007/s10459-017-9774-1

Cheung, J. J. H., Kulasegaram, K. M., Woods, N. N., Moulton, C., Ringsted, C. V., & Brydges, R. (2018). Knowing How and Knowing Why: Testing the effect of instruction designed for cognitive integration on procedural skills transfer. *Advances in Health Sciences Education*, *23*(1), 61–74. https://doi.org/10.1007/s10459-017-9774-1

Chi, M. T. (2006). Laboratory methods for assessing experts' and novices' knowledge. *The Cambridge Handbook of Expertise and Expert Performance*, 167–184.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices*. *Cognitive Science*, *5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The Nature of expertise*. Hillsdale, NJ: L. Erlbaum Associates.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing Deep Structure From the Interactions of Surface Features. *Educational Psychologist*, *47*(3), 177–188. https://doi.org/10.1080/00461520.2012.695709

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cohen, E. R., Feinglass, J., Barsuk, J. H., Barnard, C., O'Donnell, A., McGaghie, W. C., & Wayne, D. B. (2010). Cost Savings From Reduced Catheter-Related Bloodstream

Infection After Simulation-Based Education for Residents in a Medical Intensive Care Unit: *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *5*(2), 98–102. https://doi.org/10.1097/SIH.0b013e3181bc8304

Cook, D. A. (2010). One Drop at a Time: Research to Advance the Science of Simulation. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *5*(1), 1–4. https://doi.org/10.1097/SIH.0b013e3181c82aaa

Cook, D. A. (2012). If you teach them, they will learn: Why medical education needs comparative effectiveness research. *Advances in Health Sciences Education*, *17*(3), 305–310. https://doi.org/10.1007/s10459-012-9381-0

Cook, D. A., Andersen, D. K., Combes, J. R., Feldman, D. L., & Sachdeva, A. K. (2018). The value proposition of simulation-based education. *Surgery*. https://doi.org/10.1016/j.surg.2017.11.008

Cook, D. A., & Beckman, T. J. (2008). Reflections on experimental research in medical education. *Advances in Health Sciences Education*, *15*(3), 455–464. https://doi.org/10.1007/s10459-008-9117-3

Cook, D. A., Bordage, G., & Schmidt, H. G. (2008). Description, justification and clarification: A framework for classifying the purposes of research in medical education. *Medical Education*, *42*(2), 128–133. https://doi.org/10.1111/j.1365-2923.2007.02974.x

Cook, D. A., Brydges, R., Hamstra, S. J., Zendejas, B., Szostek, J. H., Wang, A. T., … Hatala, R. (2012). Comparative Effectiveness of Technology-Enhanced Simulation Versus Other Instructional Methods: A Systematic Review and Meta-Analysis. *Simulation in*

Healthcare: The Journal of the Society for Simulation in Healthcare, *7*(5), 308–320. https://doi.org/10.1097/SIH.0b013e3182614f95

Cook, D. A., Brydges, R., Zendejas, B., Hamstra, S. J., & Hatala, R. (2013). Mastery Learning for Health Professionals Using Technology-Enhanced Simulation: A Systematic Review and Meta-Analysis. *Academic Medicine*, *88*(8), 1178–1186. https://doi.org/10.1097/ACM.0b013e31829a365d

Cook, D. A., Hamstra, S. J., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., … Hatala, R. (2013). Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Medical Teacher*, *35*(1), e867–e898. https://doi.org/10.3109/0142159X.2012.714886

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., … Hamstra, S. J. (2011). Technology-Enhanced Simulation for Health Professions Education A Systematic Review and Meta-Analysis. *JAMA: The Journal of the American Medical Association*, *306*(9), 978–988. https://doi.org/10.1001/jama.2011.1234

de Bruin, A. B., Schmidt, H. G., & Rikers, R. M. (2005). The role of basic science knowledge and clinical knowledge in diagnostic reasoning: A structural equation modeling approach. *Academic Medicine*, *80*(8), 765–773.

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). *Types and qualities of knowledge*. *31*(2), 105–113.

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology*, *113*(4), 552–568. https://doi.org/10.1016/j.jecp.2012.06.009

Dieckmann, P., Gaba, D., & Rall, M. (2007). Deepening the Theoretical Foundations of Patient Simulation as Social Practice: *Simulation In Healthcare: The Journal of the Society for Simulation in Healthcare*, *2*(3), 183–193. https://doi.org/10.1097/SIH.0b013e3180f637f5

Dieckmann, P., Phero, J. C., Issenberg, S. B., Kardong-Edgren, S., Østergaard, D., & Ringsted, C. (2011). The first Research Consensus Summit of the Society for Simulation in Healthcare: Conduction and a synthesis of the results. *Simulation in Healthcare*, *6*(7), S1–S9.

Domuracki, K., Wong, A., Olivieri, L., & Grierson, L. E. M. (2015). The impacts of observing flawed and flawless demonstrations on clinical skill learning. *Medical Education*, *49*(2), 186–192. https://doi.org/10.1111/medu.12631

Dubrowski, A. (2005). Performance vs. learning curves: What is motor learning and how is it measured? *Surgical Endoscopy*, *19*(9), 1290.

Dubrowski, A., Park, J., Moulton, C., Larmer, J., & MacRae, H. (2007). A comparison of single- and multiple-stage approaches to teaching laparoscopic suturing. *The American Journal of Surgery*, *193*(2), 269–273. https://doi.org/10.1016/j.amjsurg.2006.07.013

Ericsson, K. A. (2004). Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. *Academic Medicine*, *79*(10). Retrieved from https://insights.ovid.com/pubmed?pmid=15383395

Ericsson, K. A. (2006a). An introduction to Cambridge handbook of expertise and expert performance: Its development, organization, and content. *The Cambridge Handbook of Expertise and Expert Performance*, 3–19.

Ericsson, K. A. (2006b). The influence of experience and deliberate practice on the development of superior expert performance. In *The Cambridge handbook of expertise and expert performance* (pp. 683–703). Retrieved from http://www.skillteam.se/wp-content/uploads/2011/12/Ericsson_delib_pract.pdf

Ericsson, K. A., Krampe, R. T., & Tesch-romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 363–406.

Eva, K W, Neville, A. J., & Norman, G. R. (1998). Exploring the etiology of content specificity: Factors influencing analogic transfer and problem solving. *Academic Medicine: Journal of the Association of American Medical Colleges*, *73*(10 Suppl), S1-5.

Eva, Kevin W., Brady, C., Pearson, M., & Seto, K. (2018). The pedagogical value of testing: How far does it extend? *Advances in Health Sciences Education*, 1–14. https://doi.org/10.1007/s10459-018-9831-4

Fanning, R. M., & Gaba, D. M. (2007). The Role of Debriefing in Simulation-Based Learning: *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *2*(2), 115–125. https://doi.org/10.1097/SIH.0b013e3180315539

Finnerty, E. P., Chauvin, S., Bonaminio, G., Andrews, M., Carroll, R. G., & Pangaro, L. N. (2010). Flexner Revisited: The Role and Value of the Basic Sciences in Medical

Education: *Academic Medicine*, *85*(2), 349–355. https://doi.org/10.1097/ACM.0b013e3181c88b09

Frank, J. R., Snell, L. S., Cate, O. T., Holmboe, E. S., Carraccio, C., Swing, S. R., … Harris, K. A. (2010). Competency-based medical education: Theory to practice. *Medical Teacher*, *32*(8), 638–645. https://doi.org/10.3109/0142159X.2010.501190

Gaba, D. M. (2004). The future vision of simulation in health care. *Quality and Safety in Health Care*, *13*(suppl_1), i2–i10. https://doi.org/10.1136/qshc.2004.009878

Gaba, D. M. (2011). Where Do We Come From? What Are We? Where Are We Going? *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *6*(4), 195–196. https://doi.org/10.1097/SIH.0b013e31822802d4

Goldman, E., & Schroth, W. S. (2012). Perspective: Deconstructing Integration. *Academic Medicine*, *87*(6), 729–734. https://doi.org/10.1097/ACM.0b013e318253cad4

Goldstone, R. L., & Son, J. Y. (2005). The Transfer of Scientific Principles Using Concrete and Idealized Simulations. *Journal of the Learning Sciences*, *14*(1), 69–110. https://doi.org/10.1207/s15327809jls1401_4

Goldszmidt, M., Minda, J. P., Devantier, S. L., Skye, A. L., & Woods, N. N. (2012). Expanding the basic science debate: The role of physics knowledge in interpreting clinical findings. *Advances in Health Sciences Education*, *17*(4), 547–555. https://doi.org/10.1007/s10459-011-9331-2

Grierson, L. E. M., Barry, M., Kapralos, B., Carnahan, H., & Dubrowski, A. (2012). The role of collaborative interactivity in the observational practice of clinical skills: Internetmediated educational networking. *Medical Education*, *46*(4), 409–416. https://doi.org/10.1111/j.1365-2923.2011.04196.x

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge Point: A Framework for Conceptualizing the Effects of Various Practice Conditions in Motor Learning. *Journal of Motor Behavior*, *36*(2), 212–224. https://doi.org/10.3200/JMBR.36.2.212-224

Haji, F. A. (2015). *Advancing Theory in Healthcare Simulation Instructional Design: The Effect of Task Complexity on Novice Learning and Cognitive Load*. University of Toronto, Toronto, ON.

Haji, F. A., Cheung, J. J. H., Woods, N. N., Regehr, G., deRibaupierre, S., & Dubrowski, A. (2016). Thrive or overload? The effect of task complexity on novices' simulation-based learning. *Medical Education*, *(Accepted) MED-2015-0552.R1*.

Haji, F. A., Cheung, J. J. H., Woods, N., Regehr, G., de Ribaupierre, S., & Dubrowski, A. (2016). Thrive or overload? The effect of task complexity on novices' simulation-based learning. *Medical Education*, *50*(9), 955–968. https://doi.org/10.1111/medu.13086

Haji, F. A., Da Silva, C., Daigle, D. T., & Dubrowski, A. (2014). From Bricks to Buildings: Adapting the Medical Research Council Framework to Develop Programs of Research in Simulation Education and Training for the Health Professions. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 1. https://doi.org/10.1097/SIH.0000000000000039

Hamstra, S. J., Brydges, R., Hatala, R., Zendejas, B., & Cook, D. A. (2014). Reconsidering Fidelity in Simulation-Based Training: *Academic Medicine*, *89*(3), 387–392. https://doi.org/10.1097/ACM.0000000000000130

Harden, R. M. (2000). The integration ladder: A tool for curriculum planning and evaluation. *MEDICAL EDUCATION-OXFORD-*, *34*(7), 551–557.

Hatala, R., Brooks, L., & Norman, G. (2003). Practice makes perfect: The critical role of mixed practice in the acquisition of ECG interpretation skills. *Advances in Health Sciences Education*, *8*(1), 17–26. https://doi.org/10.1023/A:1022687404380

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. W. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York, NY, US: W H Freeman/Times Books/ Henry Holt & Co.

Hattie, J., & Yates, G. C. R. (2013). The role of feedback. In *Visible Learning and the Science of How We Learn* (pp. 64–71). Routledge.

Hay, D. B. (2007). Using concept maps to measure deep, surface and non-learning outcomes. *Studies in Higher Education*, *32*(1), 39–57. https://doi.org/10.1080/03075070601099432

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, *76*(4), 408–420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: The Guilford Press.

Hays, R. (2013). Integration in medical education: What do we mean? *Education for Primary Care*, *24*(3), 151–152. https://doi.org/10.1080/14739879.2013.11494358

Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect size and Related Estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Hirshfield, L. E., Yudkowsky, R., & Park, Y. S. (2019). Pre-medical majors in the humanities and social sciences: Impact on communication skills and specialty choice. *Medical Education*, *53*(4), 408–416. https://doi.org/10.1111/medu.13774

Hodges, B. D. (2010). A tea-steeping or i-Doc model for medical education? *Academic Medicine: Journal of the Association of American Medical Colleges*, *85*(9 Suppl), S34-44. https://doi.org/10.1097/ACM.0b013e3181f12f32

Hodges, B. D., & Kuper, A. (2012). Theory and Practice in the Design and Conduct of Graduate Medical Education: *Academic Medicine*, *87*(1), 25–33. https://doi.org/10.1097/ACM.0b013e318238e069

Hodges, B. D., & Lingard, L. (2012). *The question of competence reconsidering medical education in the twenty-first century*. Retrieved from http://oclc-marc.ebrary.com/id/10614230

Holmboe, E. S. (2014). Realizing the Promise of Competency-Based Medical Education: *Academic Medicine*, 1. https://doi.org/10.1097/ACM.0000000000000515

Holyoak, K. J. (1991). Symbolic connectionism: Toward third-generation theories of expertise. In *Toward a general theory of expertise: Prospects and limits* (pp. 301–335). New York, NY, US: Cambridge University Press.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition*, *15*(4), 332–340. https://doi.org/10.3758/BF03197035

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist*, *52*(1), 35.

Ilgen, J. S., Ma, I. W. Y., Hatala, R., & Cook, D. A. (2015). A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Medical Education*, *49*(2), 161–173. https://doi.org/10.1111/medu.12621

Issenberg, S. B. (2006). The scope of simulation-based healthcare education. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, *1*(4), 203–208. https://doi.org/10.1097/01.SIH.0000246607.36504.5a

Issenberg, S. B., McGaghie, W. C., Hart, I. R., Mayer, J. W., Felner, J. M., Petrusa, E. R., … Ewy, G. A. (1999). Simulation Technology for Health Care Professional Skills Training and Assessment. *JAMA*, *282*(9), 861–866. https://doi.org/10.1001/jama.282.9.861

Issenberg, S. B., Mcgaghie, W. C., Petrusa, E. R., Lee Gordon, D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review*. *Medical Teacher*, *27*(1), 10–28. https://doi.org/10.1080/01421590500046924

Issenberg, S. B., Ringsted, C., Østergaard, D., & Dieckmann, P. (2011). Setting a Research Agenda for Simulation-Based Healthcare Education: A Synthesis of the Outcome From an Utstein Style Meeting. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *6*(3), 155–167. https://doi.org/10.1097/SIH.0b013e3182207c24

Issenberg, S. B., & Scalese, R. J. (2008). Simulation in Health Care Education. *Perspectives in Biology and Medicine*, *51*(1), 31–46. https://doi.org/10.1353/pbm.2008.0004

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The Advantage of Abstract Examples in Learning Math. *Science*, *320*(5875), 454–455. https://doi.org/10.1126/science.1154659

Kapur, M. (2014). Productive Failure in Learning Math. *Cognitive Science*, *38*(5), 1008–1022. https://doi.org/10.1111/cogs.12107

Kapur, M. (2016). Examining Productive Failure, Productive Success, Unproductive Failure, and Unproductive Success in Learning. *Educational Psychologist*, *51*(2), 289–299. https://doi.org/10.1080/00461520.2016.1155457

Karpov, Y. V., & Bransford, J. D. (1995). L.S. Vygotsky and the doctrine of empirical and theoretical learning. *Educational Psychologist*, *30*(2), 61–66. https://doi.org/10.1207/s15326985ep3002_2

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86.

Kromann, C. B., Jensen, M. L., & Ringsted, C. (2009). The effect of testing on skills learning. *Medical Education*, *43*(1), 21–27. https://doi.org/10.1111/j.1365-2923.2008.03245.x

Kulasegaram, Kulamakan M, Chaudhary, Z., Woods, N., Dore, K., Neville, A., & Norman, G. (2017). Contexts, concepts and cognition: Principles for the transfer of basic science knowledge. *Medical Education*, *51*(2), 184–195. https://doi.org/10.1111/medu.13145

Kulasegaram, Kulamakan M, & McConnell, M. (2016). When I say … transfer-appropriate processing. *Medical Education*, *50*(5), 509–510. https://doi.org/10.1111/medu.12955

Kulasegaram, Kulamakan Mahan, Manzone, J. C., Ku, C., Skye, A., Wadey, V., & Woods, N. N. (2015). Cause and Effect: Testing a Mechanism and Method for the Cognitive Integration of Basic Science. *Academic Medicine*, *90*, S63–S69. https://doi.org/10.1097/ACM.0000000000000896

Kulasegaram, Kulamakan Mahan, Martimianakis, M. A., Mylopoulos, M., Whitehead, C. R., & Woods, N. N. (2013). Cognition Before Curriculum: Rethinking the Integration of Basic Science and Clinical Learning. *Academic Medicine*, *88*(10), 1578–1585. https://doi.org/10.1097/ACM.0b013e3182a45def

Kulasegaram, Kulamakan Mahan, Min, C., Howey, E., Neville, A., Woods, N., Dore, K., & Norman, G. (2014). The mediating effect of context variation in mixed practice for transfer of basic science. *Advances in Health Sciences Education*, 1–16. https://doi.org/10.1007/s10459-014-9574-9

Kuper, A., & D'Eon, M. (2011). Rethinking the basis of medical knowledge. *Medical Education*, *45*(1), 36–43. https://doi.org/10.1111/j.1365-2923.2010.03791.x

Lammers, R. L., Temple, K. j., Wagner, M. J., & Ray, D. (2005). Competence of New Emergency Medicine Residents in the Performance of Lumbar Punctures. *Academic Emergency Medicine*, *12*(7), 622–628. https://doi.org/10.1197/j.aem.2005.01.014

Larsen, D. P., Butler, A. C., Lawson, A. L., & Roediger III, H. L. (2012). The importance of seeing the patient: Test-enhanced learning with standardized patients and written tests improves clinical application of knowledge. *Advances in Health Sciences Education*, *18*(3), 409–425. https://doi.org/10.1007/s10459-012-9379-7

Larsen, D. P., Butler, A. C., & Roediger III, H. L. (2008). Test-enhanced learning in medical education. *Medical Education*, *42*(10), 959–966. https://doi.org/10.1111/j.1365-2923.2008.03124.x

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest*, *46*(3), 328–344.

Lee, T. D., White, M. A., & Carnahan, H. (1990). On the role of knowledge of results in motor learning: Exploring the guidance hypothesis. *Journal of Motor Behavior*, *22*(2), 191–208.

Leppink, J. (2015). On causality and mechanisms in medical education research: An example of path analysis. *Perspectives on Medical Education*, *4*(2), 66–72. https://doi.org/10.1007/s40037-015-0174-z

Leung, W.-C. (2002). Competency based medical training: Review. *BMJ: British Medical Journal*, *325*(7366), 693.

Lin, Y., Cheng, A., Hecker, K., Grant, V., & Currie, G. R. (2018). Implementing economic evaluation in simulation-based medical education: Challenges and opportunities. *Medical Education*, *52*(2), 150–160. https://doi.org/10.1111/medu.13411

Lisk, K., Agur, A. M. R., & Woods, N. N. (2016). Exploring cognitive integration of basic science and its effect on diagnostic reasoning in novices. *Perspectives on Medical Education*, *5*(3), 147–153. https://doi.org/10.1007/s40037-016-0268-2

Long, D. M. (2000). Competency-based residency training: The next advance in graduate medical education. *Academic Medicine: Journal of the Association of American Medical Colleges*, *75*(12), 1178–1183.

Lucey, C. R. (2013). Medical Education: Part of the Problem and Part of the Solution. *JAMA Internal Medicine*, *173*(17), 1639. https://doi.org/10.1001/jamainternmed.2013.9074

Ma, I. W. Y., Zalunardo, N., Pachev, G., Beran, T., Brown, M., Hatala, R., & McLaughlin, K. (2012). Comparing the use of global rating scale with checklists for the assessment of central venous catheterization skills using simulation. *Advances in Health Sciences Education*, *17*(4), 457–470. https://doi.org/10.1007/s10459-011-9322-3

Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *The British Journal of Surgery*, *84*(2), 273–278.

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist*, *59*(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14

McCullagh, P., & Caird, J. K. (1990). Correct and Learning-Models and the Use of Model Knowledge of Results in the Acquisition and Retention of a Motor Skill. *Journal of Human Movement Studies*, *18*(3), 107–116.

McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, *14*(2), 200–206.

McGaghie, W. C., Draycott, T. J., Dunn, W. F., Lopez, C. M. R., & Stefanidis, D. (2011). Evaluating the Impact of Simulation on Translational Patient Outcomes. *Journal of the Society for Simulation in Healthcare August 2011*, *6*(7). https://doi.org/10.1097/SIH.0b013e318222fde9

McGaghie, W. C., Issenberg, S. B., Barsuk, J. H., & Wayne, D. B. (2014). A critical review of simulation-based mastery learning with translational outcomes. *Medical Education*, *48*(4), 375–385. https://doi.org/10.1111/medu.12391

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2006). Effect of practice on standardised learning outcomes in simulation-based medical education. *Medical Education*, *40*(8), 792–797. https://doi.org/10.1111/j.1365-2929.2006.02528.x

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003–2009: Simulation-based medical education research 2003–2009. *Medical Education*, *44*(1), 50–63. https://doi.org/10.1111/j.1365-2923.2009.03547.x

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2016). Revisiting 'A critical review of simulation-based medical education research: 2003–2009.' *Medical Education*, *50*(10), 986–991. https://doi.org/10.1111/medu.12795

Moorthy, K., Munz, Y., Sarker, S. K., & Darzi, A. (2003). Objective assessment of technical skills in surgery. *BMJ (Clinical Research Ed.)*, *327*(7422), 1032–1037. https://doi.org/10.1136/bmj.327.7422.1032

Motola, I., Devine, L. A., Chung, H. S., Sullivan, J. E., & Issenberg, S. B. (2013). Simulation in healthcare education: A best evidence practical guide. AMEE Guide No. 82. *Medical Teacher*, *35*(10), e1511–e1530. https://doi.org/10.3109/0142159X.2013.818632

Moulton, C.-A. E., Dubrowski, A., MacRae, H., Graham, B., Grober, E., & Reznick, R. (2006). Teaching Surgical Skills: What Kind of Practice Makes Perfect?: A Randomized, Controlled Trial. *Transactions of the ... Meeting of the American Surgical Association*, *124*, 66–75. https://doi.org/10.1097/01.sla.0000234808.85789.6a

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289–316. https://doi.org/10.1037/0033-295X.92.3.289

Mylopoulos, M., Brydges, R., Woods, N. N., Manzone, J., & Schwartz, D. L. (2016). Preparation for future learning: A missing competency in health professions education? *Medical Education*, *50*(1), 115–123. https://doi.org/10.1111/medu.12893

Mylopoulos, M., Lohfeld, L., Norman, G. R., Dhaliwal, G., & Eva, K. W. (2012). Renowned Physicians' Perceptions of Expert Diagnostic Practice: *Academic Medicine*, *87*(10), 1413–1417. https://doi.org/10.1097/ACM.0b013e31826735fc

Mylopoulos, M., & Woods, N. (2014). Preparing medical students for future learning using basic science instruction. *Medical Education*, *48*(7), 667–673. https://doi.org/10.1111/medu.12426

Mylopoulos, M., & Woods, N. N. (2009). Having our cake and eating it too: Seeking the best of both worlds in expertise research. *Medical Education*, *43*(5), 406–413. https://doi.org/10.1111/j.1365-2923.2009.03307.x

Mylopoulos, M., & Woods, N. N. (2017). When I say … adaptive expertise. *Medical Education*, *51*(7), 685–686. https://doi.org/10.1111/medu.13247

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory & Cognition*, *19*(6), 543–557.

Nestel, D., Brazil, V., & Hay, M. (2018). You can't put a value on that… Or can you? Economic evaluation in simulation-based medical education. *Medical Education*, *52*(2), 139–141. https://doi.org/10.1111/medu.13505

Norman, Geoff. (1996). From the Editors: Research, theory and science in medical education. *Advances in Health Sciences Education*, *1*(3), 175–177. https://doi.org/10.1007/BF00162914

Norman, Geoff. (2003). RCT = results confounded and trivial: The perils of grand educational experiments. *Medical Education*, *37*(7), 582–584.

Norman, Geoff. (2004). Editorial – Theory Testing Research Versus Theory-Based Research. *Advances in Health Sciences Education*, *9*(3), 175–178. https://doi.org/10.1023/B:AHSE.0000038310.81360.20

Norman, Geoff. (2009). Teaching basic science to optimize transfer. *Medical Teacher*, *31*(9), 807–811. https://doi.org/10.1080/01421590903049814

Norman, Geoff. (2014). Context, curriculum and competence. *Advances in Health Sciences Education*, *19*(5), 625–628. https://doi.org/10.1007/s10459-014-9565-x

Norman, Geoff, Dore, K., & Grierson, L. (2012). The minimal relationship between simulation fidelity and transfer of learning. *Medical Education*, *46*(7), 636–647. https://doi.org/10.1111/j.1365-2923.2012.04243.x

Norman, Geoff, Eva, K., Brooks, L., & Hamstra, S. (2006). Expertise in Medicine and Surgery. *The Cambridge Handbook of Expertise and Expert Performance*, 339–353.

Norman, Geoff, Norcini, J., & Bordage, G. (2014). Competency-Based Education: Milestones or Millstones? *Journal of Graduate Medical Education*, *6*(1), 1–6. https://doi.org/10.4300/JGME-D-13-00445.1

Norman, Geoffrey, Dore, K., Krebs, J., & Neville, A. J. (2007). The Power of the Plural: Effect of Conceptual Analogies on Successful Transfer. *Academic Medicine*, *82*(10), S16. https://doi.org/10.1097/ACM.0b013e3181405ad7

Norman, Geoffrey, Monteiro, S., & Salama, S. (2012). Sample size calculations: Should the emperor's clothes be off the peg or made to measure? *BMJ*, *345*, e5278. https://doi.org/10.1136/bmj.e5278

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(3), 398–415. https://doi.org/10.1037/0278-7393.17.3.398

Perkins, D. N., & Salomon, G. (1992). TRANSFER OF LEARNING. In T. Husén & T. N. Postlethwaite (Eds.), *The International Encylcopedia of Education* (2nd ed., p. 13). Oxford: Pergamon.

Porte, M. C., Xeroulis, G., Reznick, R. K., & Dubrowski, A. (2007). Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *The American Journal of Surgery*, *193*(1), 105–110. https://doi.org/16/j.amjsurg.2006.03.016

Raemer, D., Anderson, M., Cheng, A., Fanning, R., Nadkarni, V., & Savoldelli, G. (2011). Research Regarding Debriefing as Part of the Learning Process. *Journal of the Society for Simulation in Healthcare August 2011*, *6*(7). https://doi.org/10.1097/SIH.0b013e31822724d0

Regehr, G., Macrae, H., Reznick, R. K., & Szalay, D. (1998). Comparing the psychometric properties of checklists and global rating scales for assessing performance on an Osce-format examination. *Academic Medicine*, *73*(9), 993–997.

Reigeluth, C. M., & Carr-Chellman, A. A. (2009). *Instructional-design theories and models. Vol. 3, Vol. 3,*. Retrieved from http://public.eblib.com/EBLPublic/PublicView.do?ptiID=425294

Reznick, R. K., & MacRae, H. (2006). Teaching surgical skills—Changes in the wind. *New England Journal of Medicine*, *355*(25), 2664–2669.

Rikers, R. M. J. P., Loyens, S. M. M., & Schmidt, H. G. (2004). The role of encapsulated knowledge in clinical case representations of medical students and family doctors. *Medical Education*, *38*(10), 1035–1043. https://doi.org/10.1111/j.1365-2929.2004.01955.x

Rikers, R. M. J. P., Loyens, S., te Winkel, W., Schmidt, H. G., & Sins, P. H. M. (2005). The Role of Biomedical Knowledge in Clinical Reasoning: A Lexical Decision Study: *Academic Medicine*, *80*(10), 945–949. https://doi.org/10.1097/00001888-200510000-00015

Rittle-Johnson, B., & Schneider, M. (2015). Developing Conceptual and Procedural Knowledge of Mathematics. In R. C. Kadosh & A. Dowker (Eds.), *Oxford Handbook of Numerical Cognition* (pp. 1118–1134). Retrieved from http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199642342.001.0001/oxfordhb-9780199642342-e-014

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a One-Way Street: Bidirectional Relations Between Procedural and Conceptual Knowledge of Mathematics. *Educational Psychology Review*, *27*(4), 587–597. https://doi.org/10.1007/s10648-015-9302-x

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology*, *101*(4), 836–852. http://dx.doi.org.myaccess.library.utoronto.ca/10.1037/a0016026

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing procedures?: Developing procedural flexibility. *British Journal of Educational Psychology*, *82*(3), 436–455. https://doi.org/10.1111/j.2044-8279.2011.02037.x

Rivière, E., Jaffrelot, M., Jouquan, J., & Chiniara, G. (2019). Debriefing for the Transfer of Learning: The Importance of Context. *Academic Medicine*, *Publish Ahead of Print*. https://doi.org/10.1097/ACM.0000000000002612

Rosen, K. R. (2008). The history of medical simulation. *Journal of Critical Care*, *23*(2), 157–166. https://doi.org/10.1016/j.jcrc.2007.12.004

Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation Analysis in Social Psychology: Current Practices and New Recommendations: Mediation Analysis in Social Psychology. *Social and Personality Psychology Compass*, *5*(6), 359–371. https://doi.org/10.1111/j.1751-9004.2011.00355.x

Rudolph, J. W., Simon, R., & Raemer, D. B. (2007). Which Reality Matters? Questions on the Path to High Engagement in Healthcare Simulation: *Simulation In Healthcare: The Journal of the Society for Simulation in Healthcare*, *2*(3), 161–163. https://doi.org/10.1097/SIH.0b013e31813d1035

Salomon, G., & Perkins, D. N. (1989). Rocky Roads to Transfer: Rethinking Mechanism of a Neglected Phenomenon. *Educational Psychologist*, *24*(2), 113–142. https://doi.org/10.1207/s15326985ep2402_1

Salomon, G., & Perkins, D. N. (2015). School Learning for Transfer. In *International Encyclopedia of the Social & Behavioral Sciences* (pp. 96–100). https://doi.org/10.1016/B978-0-08-097086-8.92040-3

Sawyer, T., White, M., Zaveri, P., Chang, T., Ades, A., French, H., … Kessler, D. (2015). Learn, See, Practice, Prove, Do, Maintain: An Evidence-Based Pedagogical Framework for Procedural Skill Training in Medicine. *Academic Medicine*, 1. https://doi.org/10.1097/ACM.0000000000000734

Scalese, R. J., Obeso, V. T., & Issenberg, S. B. (2008). Simulation Technology for Skills Training and Competency Assessment in Medical Education. *Journal of General Internal Medicine*, *23*(1), 46–49. https://doi.org/10.1007/s11606-007-0283-4

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, *82*(4), 225.

Schmidt, R. A., & Bjork, R. A. (1992a). *NEW CONCEPTUALIZATIONS OF PRACTICE: Common Principles in Three*. Retrieved from http://www2.psychology.uiowa.edu/Classes/31330/MotorControl/SchmidtBjork1992.pdf

Schmidt, R. A., & Bjork, R. A. (1992b). New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training. *Psychological Science*, *3*(4), 207–218. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x

Schmidt, R. A., & Lee, T. D. (2005). *Motor control and learning: A behavioral emphasis* (4th ed.). Champaign, Il, US: Human Kinetics.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, *47*(6), 1525–1538. https://doi.org/10.1037/a0024997

Schwartz, D. L., & Bransford, J. D. (1998). A Time For Telling. *Cognition and Instruction*, *16*(4), 475–5223. https://doi.org/10.1207/s1532690xci1604_4

Schwartz, D. L., Chase, C. C., & Bransford, J. D. (2012). Resisting Overzealous Transfer: Coordinating Previously Successful Routines With Needs for New Learning. *Educational Psychologist*, *47*(3), 204–214. https://doi.org/10.1080/00461520.2012.696317

Schwartz, D. L., Lindgren, R., & Lewis, S. (2009). Constructivism in an age of non-constructivist assessments. In *Constructivist instruction: Success or failure?* (pp. 34–61). New York, NY, US: Routledge/Taylor & Francis Group.

Shea, C. H., Wright, D. L., Wulf, G., & Whitacre, C. (2000). Physical and observational practice afford unique learning opportunities. *Journal of Motor Behavior*, *32*(1), 27–36. https://doi.org/10.1080/00222890009601357

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(2), 179.

Shea, J. B., & Zimny, S. T. (1983). Context effects in memory and learning movement information. In *Advances in Psychology*: *Vol. 12*. *Memory and control of action* (pp. 345–366). Amsterdam: Elsevier.

Shea, J. B., & Zimny, S. T. (1988). Knowledge incorporation in motor representation. In O. G. Meijer & K. Roth (Eds.), *Complex movement behaviour: "The" motor-action controversy* (pp. 289–314). Amsterdam: Elsevier.

Sklar, D. P. (2013). Integrating Competencies: *Academic Medicine*, *88*(8), 1049–1051. https://doi.org/10.1097/ACM.0b013e31829ab5c8

Star, J. R. (2005). Reconceptualizing Procedural Knowledge. *Journal for Research in Mathematics Education*, *36*(5), 404–411.

Starkes, J. L., & Allard, F. (Eds.). (1993). *Cognitive issues in motor expertise*. Amsterdam ; New York: North-Holland.

Steenhof, N., Woods, N. N., Van Gerven, P. W. M., & Mylopoulos, M. (2019). Productive failure as an instructional approach to promote future learning. *Advances in Health Sciences Education*. https://doi.org/10.1007/s10459-019-09895-4

Teteris, E., Fraser, K., Wright, B., & McLaughlin, K. (2012). Does training learners on simulators benefit real patients? *Advances in Health Sciences Education*, *17*(1), 137–144. https://doi.org/10.1007/s10459-011-9304-5

The Royal College of Physicians and Surgeons of Canada. (2012). *Objectives of Training in the Subspecialty of General Internal Medicine*. Retrieved from http://www.royalcollege.ca/cs/groups/public/documents/document/y2vk/mdaw/~edisp/tztest3rcpsced000901.pdf

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related.

Studies in Educational Evaluation, *54*, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies: Cognitive load theory. *Medical Education*, *44*(1), 85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x

Whitehead, C., Kuper, A., & Webster, F. (2012). The conceit of curriculum. *Medical Education*, *46*(6), 534–536.

Whitehead, C. R., Austin, Z., & Hodges, B. D. (2012). Continuing the competency debate: Reflections on definitions and discourses. *Advances in Health Sciences Education*, *18*(1), 123–127. https://doi.org/10.1007/s10459-012-9407-7

Whitehead, Cynthia R., & Kuper, A. (2017). Faith-based medical education. *Advances in Health Sciences Education*, *22*(1), 1–3. https://doi.org/10.1007/s10459-016-9748-8

Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 677.

Woods, N. N. (2007). Science is fundamental: The role of biomedical knowledge in clinical reasoning: clinical expertise. *Medical Education*, *41*(12), 1173–1177. https://doi.org/10.1111/j.1365-2923.2007.02911.x

Woods, N. N., Brooks, L. R., & Norman, G. R. (2005). The value of basic science in clinical diagnosis: Creating coherence among signs and symptoms. *Medical Education*, *39*(1), 107–112. https://doi.org/10.1111/j.1365-2929.2004.02036.x

Woods, N. N., Brooks, L. R., & Norman, G. R. (2007a). It all makes sense: Biomedical knowledge, causal connections and memory in the novice diagnostician. *Advances in Health Sciences Education*, *12*(4), 405–415. https://doi.org/10.1007/s10459-006-9055-x

Woods, N. N., Brooks, L. R., & Norman, G. R. (2007b). The role of biomedical knowledge in diagnosis of difficult clinical cases. *Advances in Health Sciences Education*, *12*(4), 417–426. https://doi.org/10.1007/s10459-006-9054-y

Woods, N. N., Howey, E. H. A., Brooks, L. R., & Norman, G. R. (2006). Speed kills? Speed, accuracy, encapsulations and causal understanding. *Medical Education*, *40*(10), 973–979. https://doi.org/10.1111/j.1365-2929.2006.02556.x

Woods, N. N., Neville, A. J., Levinson, A. J., Howey, E. H., Oczkowski, W. J., & Norman, G. R. (2006). The value of basic science in clinical diagnosis. *Academic Medicine*, *81*(10), S124–S127.

Wulf, G. (2007). Self-controlled practice enhances motor learning: Implications for physiotherapy. *Physiotherapy*, *93*(2), 96–101. https://doi.org/10.1016/j.physio.2006.08.005

Zendejas, B., Wang, A. T., Brydges, R., Hamstra, S. J., & Cook, D. A. (2013). Cost: The missing outcome in simulation-based medical education research: A systematic review. *Surgery*, *153*(2), 160–176. https://doi.org/10.1016/j.surg.2012.06.025

Ziv, A., Wolpe, P. R., Small, S. D., & Glick, S. (2003). Simulation-based medical education: An ethical imperative. *Academic Medicine: Journal of the Association of American Medical Colleges*, *78*(8), 783–788.

Appendices

Appendix I

Procedural Knowledge Test

Place these steps in the appropriate order

1) ________________

2) ________________

3) ________________

4) ________________

5) ________________

6) ________________

7) ________________

8) ________________

9) ________________

10) ________________

11) ________________

12) ________________

13) ________________

A. Landmark, then insert anesthetic needle to freeze area

B. Open the lumbar puncture tray

C. Collect cerebrospinal fluid in vials

D. Drape the area

E. Position the patient

F. Landmark, then insert spinal needle

G. Clean the area around the puncture site

H. Palpate landmarks to identify puncture site

I. Put on sterile gloves

J. Measure opening pressure

K. Prepare your tray tools

L. Remove spinal needle and apply a bandage

M. Check for cerebrospinal fluid

Appendix II

Conceptual Knowledge Test[1]

1. Why do you angle your needle at a slight angle towards the patient's head? (1 point)
2. Why do you have the bevel of the needle facing the patient's side? Specifically, what would happen if you did not align the bevel this way? (3 points)
3. Why do you enter at the L3-L4 or L4-L5 area? (1 point)
4. Why do you clean the patients back in an outward fashion from the centre of the insertion site? (1 point)
5. Why do you insert and remove the needle with the stylet in place? Specifically, what would happen if you did not use a stylet? (3 points)
6. What might happen if you did not have the needle firmly between your thumb and index finger? (2 points)
7. When redirecting your needle, why is it important to first withdraw the needle to just below the surface of the skin? What would happen if you did not? (2 points)

[1] The conceptual knowledge test was modified between study 1 (Chapter 3) and study 2 (Chapter 4).