

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



Jeffrey J. H. Cheung, MSc, PhD(c)^{1,2}, Kulamakan M. Kulasegaram, PhD^{1,3}, Nicole N. Woods, PhD^{1,3}, and Ryan Brydges, PhD^{1,2,4}

¹The Wilson Centre, University Health Network and University of Toronto, Toronto, Canada; ²Allan Waters Family Simulation Centre, St. Michael's Hospital, Toronto, Canada; ³Department of Family and Community Medicine, University of Toronto, Toronto, Canada; ⁴Department of Medicine, University of Toronto, Toronto, Canada.

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 h 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.

KEY WORDS: basic science; clinical skills training; instructional design; simulation; cognition/problem solving; transfer.

J Gen Intern Med 34(6):969–77

DOI: 10.1007/s11606-019-04959-y

© Society of General Internal Medicine 2019

INTRODUCTION

When teaching trainees clinical skills, research and curriculum documents emphasize high volumes of training as *the* path to expertise.¹ Research on how educators can optimize such training has made a strong case for instructional design features like deliberate practice,^{2, 3} mixed-practice,⁴ spaced-practice,⁵ and retrieval practice (i.e., test-enhanced learning),^{6, 7} 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.⁸ 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.

Earlier versions of this paper were presented at the Association for Medical Education in Europe in Helsinki, Finland (September 2017), and at the Canadian Conference for Medical Education in Halifax, Nova Scotia, Canada (April 2018).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-019-04959-y>) contains supplementary material, which is available to authorized users.

Published online April 1, 2019

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.”⁹ 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.^{10–13} 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.^{14–26} 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.^{20, 24, 25} 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*.²⁷ 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,²⁸ 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° towards the simulator’s umbilicus (how content) *because* the underlying anatomy of the spinous processes are also angled at 15° (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.

METHODS

Participants

After receiving institutional ethics approval, we recruited 66 medical students from the University of Toronto. Inclusion criteria included being a pre-clerkship (years 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.^{28, 29}

Learning Materials

We developed three instructional videos for LP based on previous educational materials.^{28, 30} 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 1 for an example of how content is presented in each video).

Box 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]

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 1 week later. Each participant completed the study protocol individually, with the lead author present at each training and follow-up session.

At the training session, after providing written informed consent and completing a demographic questionnaire, participants had 25 min to review their assigned instructional video. Participants were made aware they were randomized to one of the three video interventions but were unaware of how these videos differed in content or organization. Immediately after, participants had 1 h 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 1 presents the study design.

Procedural and Conceptual Knowledge Tests

We adopted both written knowledge tests from our previous study.²⁸ 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.

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,²⁸ which were originally adapted from Haji et al.³⁰

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](#)

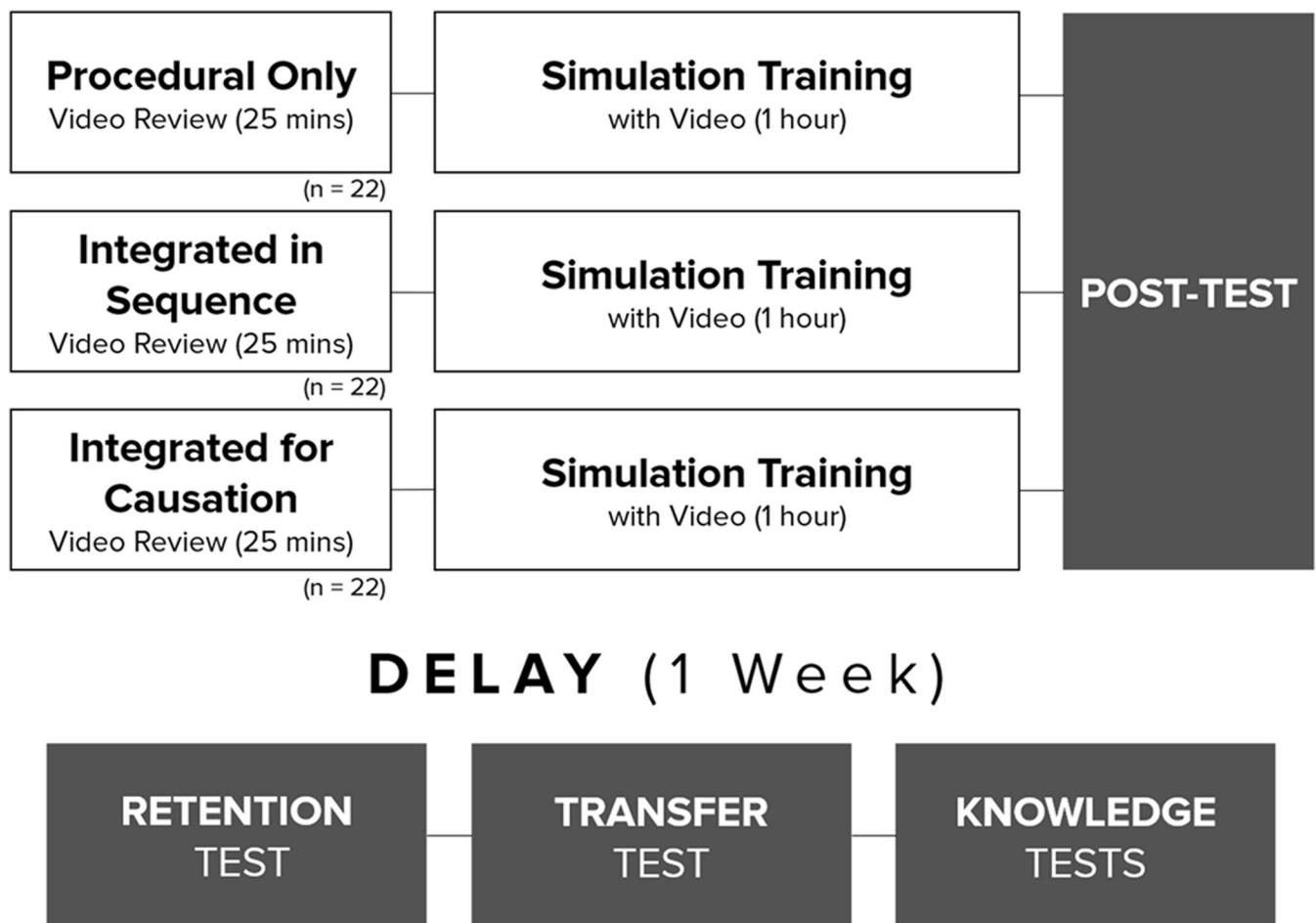


Figure 1 Study flow diagram.

content. To compare group performances on the written procedural and conceptual knowledge tests, we used Welch's one-way analysis of variance (ANOVA) and the 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 2-h rater orientation session aimed at familiarizing them with the global rating scale (GRS) used to assess procedural competence.³¹ Researchers have collected strong validity evidence for using the GRS to assess simulated procedural skills,³² 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 6.3.8.³³ and performed comparative and correlational analyses using SPSS Version 22.

Mediation Analyses

To capture the relationship between our experimental conditions, participants' conceptual knowledge, and their LP GRS performance, we conducted simple mediation analyses^{34, 35} using the PROCESS macro for SPSS.³⁶ Using indicator coding, we compared group performances by computing two mediation models for retention, and transfer test GRS scores.³⁶ In the first model, the Procedural Only group was coded as the

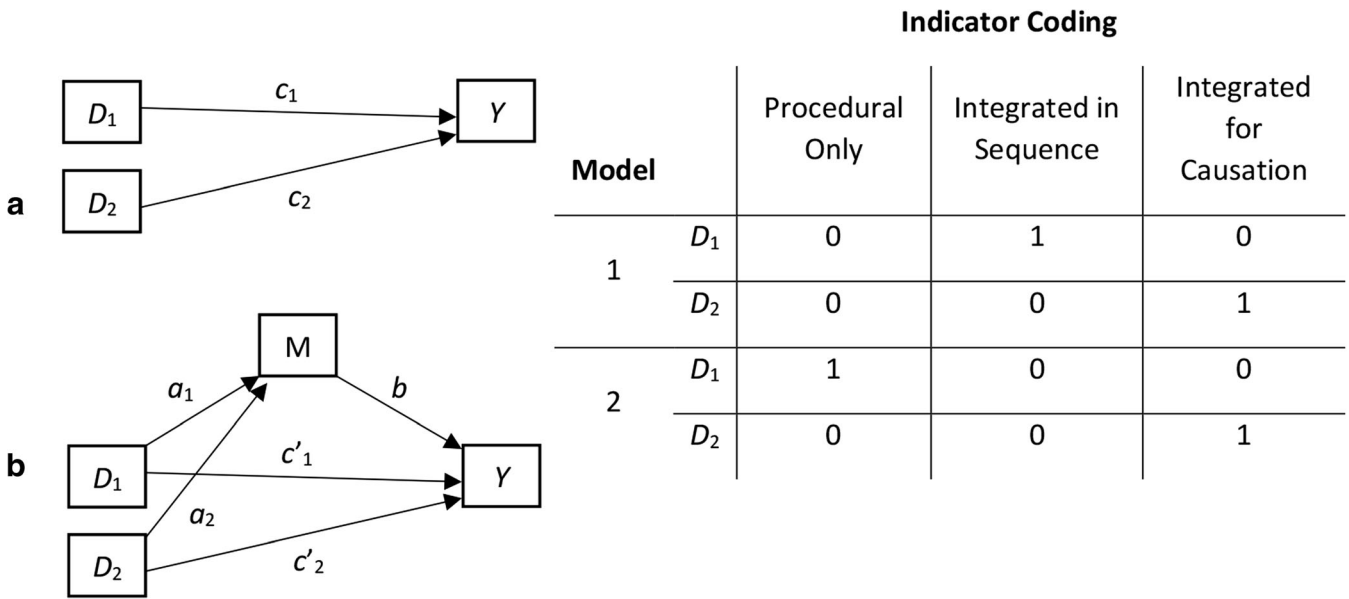


Figure 2 Simple mediation analyses of intervention effects (D) on outcome Y (i.e., global rating scale scores for retention or transfer performances) using indicator coding. (A) reveals the relative total effects (c_1 and c_2) of D_1 and D_2 ; (B) includes conceptual knowledge test score (M) as a mediator 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.

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.²⁸

These analyses enabled us to examine how the three interventions influenced the groups' retention and transfer test scores in three ways (see Fig. 2): (1) the relative total effects on those outcomes (c_1 and c_2 in Fig. 2(a)); (2) the relative direct effects (c'_1 and c'_2 in Fig. 2(a)), after controlling for participants' conceptual knowledge test scores (M); and (3) the relative indirect effects (a_1b and a_2b in Fig. 2(b)), 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 (Fig. 2(b): D_1 and D_2 acting on Y through M), the PROCESS macro computed a bias-corrected bootstrap 99% confidence interval using 10,000 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.

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).

Procedural and Conceptual Knowledge Test Performance

For the procedural knowledge test, there was no significant difference between the groups (Procedural Only: $M = 12.50$, $SD = 0.80$; Integrated in Sequence: $M = 12.27$, $SD = 1.78$; Integrated for Causation: $M = 12.86$, $SD = 0.47$), $F(2,35.80) = 1.46$, $p = 0.10$. For the conceptual knowledge test, there was a significant difference between the groups, $F(2,34.71) = 25.82$, $p < 0.001$, $\eta^2_p = 0.40$. 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 ($M = 2.69$, $SD = 0.77$), $t = 3.28$, $p < 0.001$, and the Integrated in Sequence group ($M = 4.14$, $SD = 1.84$), $t = 1.84$, $p < 0.05$. Further, the Integrated in Sequence group scored significantly higher than the Procedural Only group, $t = 1.44$, $p < 0.01$.

Mediation Models Linking Conceptual Knowledge and Lumbar Puncture Performance

Confirming the findings above, the mediation model (Fig. 2(b)) 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 1

Table 1 Regression Coefficients, Standard Errors, and Model Summary for Lumbar Puncture Retention

	<i>M</i> (conceptual knowledge test score)			<i>Y</i> (global rating scale retention performance)								
	Coeff.	<i>SE</i>	<i>t</i>	Coeff.	<i>SE</i>	<i>t</i>	Coeff.	<i>SE</i>	<i>t</i>			
Model 1												
<i>D</i> ₁ (Integrated in Sequence)	<i>a</i> ₁	1.44	0.51	2.84*	<i>c</i> ₁	0.64	1.66	0.38	<i>c'</i> ₁	-0.20	1.75	-0.11
<i>D</i> ₂ (Integrated for Causation)	<i>a</i> ₂	3.28	0.51	6.47*	<i>c</i> ₂	0.82	1.66	0.49	<i>c'</i> ₂	-1.08	2.12	-0.51
<i>M</i>	-	-	-	-	-	-	-	-	<i>b</i>	0.58	0.41	1.42
Constant	<i>i</i> ₁	2.70	0.36	7.50*	<i>i</i> ₂	17.00	1.17	14.50*	<i>i</i> ₃	15.44	1.60	9.65*
Model 2												
<i>D</i> ₁ (Procedural Only)	<i>a</i> ₁	-1.44	0.51	-2.84*	<i>c</i> ₁	-0.64	1.66	-0.38	<i>c'</i> ₁	0.20	1.75	0.11
<i>D</i> ₂ (Integrated for Causation)	<i>a</i> ₂	1.84	0.51	3.63*	<i>c</i> ₂	0.18	1.66	0.11	<i>c'</i> ₂	-0.88	1.81	-0.49
<i>M</i>	-	-	-	-	-	-	-	-	<i>b</i>	0.58	0.41	1.42
Constant	<i>i</i> ₁	4.14	0.36	11.53*	<i>i</i> ₂	17.64	1.17	15.0*	<i>i</i> ₃	15.24	2.05	7.44*
		<i>R</i> ² = 0.40				<i>R</i> ² = 0.004				<i>R</i> ² = 0.04		
		<i>F</i> (2,63) = 21.03, <i>p</i> < 0.0001*				<i>F</i> (2,62) = 0.13, <i>p</i> = 0.875				<i>F</i> (3,62) = 0.76, <i>p</i> = 0.52		

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*² statistics are shared for both models (they are identical models, but with different relative comparisons) *denotes statistical significance with *alpha* < 0.05

and 2). 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 1 and 2).

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 Fig. 2). We found that participants who had higher conceptual knowledge had higher but not significantly different retention scores (*b* = 0.579) (Table 1). Conversely, we found that participants who had higher conceptual knowledge had significantly higher transfer scores (*b* = 1.24) (Table 2).

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 and *c*₂ in Fig. 2(a)) of any intervention on LP GRS retention (Table 1) and transfer (Table 2) performances, all *p* > 0.05.

After controlling for participants' conceptual knowledge (*M* in Fig. 2(b)), we similarly found no significant relative direct effects (*c'*₁ and *c'*₂ in Fig. 2(b)) on LP GRS retention (Table 1) or transfer (Table 2) performances, all *p* > 0.05.

Relative Indirect Effects: Effects of 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*₁*b* = 1.76], 99% CI = 0.40 to 3.94; Integrated for Causation [*a*₂*b* = 4.08], 99% 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*₂*b* = 2.29], 99% CI = 0.38 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.

Table 2 Regression Coefficients, Standard Errors, and Model Summary for Lumbar Puncture Transfer

	<i>M</i> (conceptual knowledge test score)			<i>Y</i> (global rating scale transfer performance)								
	Coeff.	<i>SE</i>	<i>t</i>	Coeff.	<i>SE</i>	<i>t</i>	Coeff.	<i>SE</i>	<i>t</i>			
Model 1												
<i>D</i> ₁ (Integrated in Sequence)	<i>a</i> ₁	1.44	0.51	2.84*	<i>c</i> ₁	-0.750	1.62	-0.464	<i>c'</i> ₁	-2.54	1.60	-1.59
<i>D</i> ₂ (Integrated for Causation)	<i>a</i> ₂	3.28	0.51	6.47*	<i>c</i> ₂	-0.114	1.62	-0.070	<i>c'</i> ₂	-4.19	1.94	-2.16
<i>M</i>	-	-	-	-	-	-	-	-	<i>b</i>	1.24	0.37	3.33*
Constant	<i>i</i> ₁	2.70	0.36	7.50*	<i>i</i> ₂	16.3	1.14	14.2*	<i>i</i> ₃	12.9	1.46	8.85*
Model 2												
<i>D</i> ₁ (Procedural Only)	<i>a</i> ₁	-1.44	0.51	-2.84*	<i>c</i> ₁	0.750	1.62	0.46	<i>c'</i> ₁	2.54	1.60	1.59
<i>D</i> ₂ (Integrated for Causation)	<i>a</i> ₂	1.84	0.51	3.63*	<i>c</i> ₂	0.636	1.62	0.39	<i>c'</i> ₂	-1.65	1.65	-0.10
<i>M</i>	-	-	-	-	-	-	-	-	<i>b</i>	1.24	0.373	3.33*
Constant	<i>i</i> ₁	4.14	0.36	11.53*	<i>i</i> ₂	15.5	1.14	13.6*	<i>i</i> ₃	10.4	1.87	5.55*
		<i>R</i> ² = 0.40				<i>R</i> ² = 0.004				<i>R</i> ² = 0.16		
		<i>F</i> (2,63) = 21.03, <i>p</i> < 0.0001*				<i>F</i> (2,62) = 0.125, <i>p</i> = 0.883				<i>F</i> (3,62) = 3.80, <i>p</i> = 0.014*		

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*² statistics are shared for both models (they are identical models, but with different relative comparisons) *denotes statistical significance with *alpha* < 0.05

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 1). Specifically, the confidence intervals of the relative indirect effects contained zero and thus were non-significant (Model 1: Integrated in Sequence [$a_1b = 0.84$], 99% CI = -0.35 to 2.83; Integrated for Causation [$a_2b = 1.90$], 99% CI = -0.89 to 5.80; Model 2: Integrated for Causation [$a_2b = 1.07$], 99% CI = -0.36 to 4.30).

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.²⁸ By comparing why content Integrated for Causation versus why content Integrated in Sequence, we show that teaching that explicitly promotes cognitive integration appears to help participants further mobilize their conceptual knowledge, which then enhances how well they transfer their learning.²⁴ 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.

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)^{37–39}, 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.^{25, 40, 41} 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,^{24, 25} 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^{42, 43} and bedside invasive procedures.

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.²⁸ 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.⁴⁴ Future research could explore this finding by systematically examining how the timing of knowledge tests influence the mediating effect of that knowledge on performance.

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.

Acknowledgements: *The authors extend their thanks and appreciation to the Currie Fellowship program at the Wilson Centre, University Health Network and the Canada Graduate Scholarship program at the Natural Sciences and Engineering Research Council for funding JJHC's PhD studies. We thank the staff at the Surgical Skills Centre, Mount Sinai Hospital, Toronto, for sharing equipment and resources, Thomas Sun at Sun Innovations for his ongoing innovation and technical support of our projects, and the Wilson Centre for providing space to conduct the experiment.*

Corresponding Author: *Jeffrey J. H. Cheung, MSc, PhD(c); The Wilson Centre, University Health Network and University of Toronto, Toronto, Canada (e-mail: jeffreyjcheung@gmail.com).*

Funding Information *We are very grateful for the funding from the Bank of Montreal Chair in Health Professions Education Research and from the Department of Medicine, University of Toronto.*

Compliance with Ethical Standards:

We received institutional ethics approval from the University of Toronto prior to participant recruitment. All participants provided informed consent prior to engaging in the study protocol.

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*

REFERENCES

1. **Ericsson KA.** Deliberate Practice and the Acquisition and Maintenance of Expert Performance in Medicine and Related Domains. *Acad Med* 2004;79(10).
2. **Ericsson KA.** The influence of experience and deliberate practice on the development of superior expert performance. In: *The Cambridge Handbook of Expertise and Expert Performance*. 2006:683–703.
3. **McGaghie WC, Issenberg SB, Barsuk JH, Wayne DB.** A critical review of simulation-based mastery learning with translational outcomes. *Med Educ* 2014;48(4):375–385. <https://doi.org/10.1111/medu.12391>
4. **Hatala R, Brooks L, Norman G.** Practice makes perfect: The critical role of mixed practice in the acquisition of ECG interpretation skills. *Adv Health Sci Educ* 2003;8(1):17–26. <https://doi.org/10.1023/A:1022687404380>
5. **Moulton C-AE, Dubrowski A, MacRae H, Graham B, Grober E, Reznick R.** Teaching Surgical Skills: What Kind of Practice Makes Perfect?: A Randomized, Controlled Trial. *Ann Surg* 2006;124:66–75. <https://doi.org/10.1097/01.sla.0000234808.85789.6a>
6. **Larsen DP, Butler AC, Roediger III HL.** Test-enhanced learning in medical education. *Med Educ* 2008;42(10):959–966. <https://doi.org/10.1111/j.1365-2923.2008.03124.x>
7. **Larsen DP, Butler AC, Lawson AL, Iii HLR.** The importance of seeing the patient: test-enhanced learning with standardized patients and written tests improves clinical application of knowledge. *Adv Health Sci Educ* 2012;18(3):409–425. <https://doi.org/10.1007/s10459-012-9379-7>
8. **Kulasegaram KM, McConnell M.** When I say ... transfer-appropriate processing. *Med Educ* 2016;50(5):509–510. <https://doi.org/10.1111/medu.12955>
9. **Baroody AJ, Feil Y, Johnson AR.** An Alternative Reconceptualization of Procedural and Conceptual Knowledge. *J Res Math Educ* 2007;38(2):115–131. <https://doi.org/10.2307/30034952>
10. **Rikers RMJP, Loyens SMM, Schmidt HG.** The role of encapsulated knowledge in clinical case representations of medical students and family doctors. *Med Educ* 2004;38(10):1035–1043. <https://doi.org/10.1111/j.1365-2929.2004.01955.x>
11. **Rikers RMJP, Loyens S, te Winkel W, Schmidt HG, Sins PHM.** The Role of Biomedical Knowledge in Clinical Reasoning: A Lexical Decision Study. *Acad Med*. 2005;80(10):945–949. <https://doi.org/10.1097/00001888-200510000-00015>
12. **de Bruin AB, Schmidt HG, Rikers RM.** The role of basic science knowledge and clinical knowledge in diagnostic reasoning: a structural equation modeling approach. *Acad Med* 2005;80(8):765–773.
13. **Mylopoulos M, Woods NN.** Having our cake and eating it too: seeking the best of both worlds in expertise research. *Med Educ* 2009;43(5):406–413. <https://doi.org/10.1111/j.1365-2923.2009.03307.x>
14. **Woods NN, Brooks LR, Norman GR.** The value of basic science in clinical diagnosis: creating coherence among signs and symptoms. *Med Educ* 2005;39(1):107–112. <https://doi.org/10.1111/j.1365-2929.2004.02036.x>
15. **Woods NN, Neville AJ, Levinson AJ, Howey EH, Oczkowski WJ, Norman GR.** The value of basic science in clinical diagnosis. *Acad Med* 2006;81(10):S124–S127.
16. **Woods NN, Howey EHA, Brooks LR, Norman GR.** Speed kills? Speed, accuracy, encapsulations and causal understanding. *Med Educ* 2006;40(10):973–979. <https://doi.org/10.1111/j.1365-2929.2006.02556.x>
17. **Woods NN.** Science is fundamental: the role of biomedical knowledge in clinical reasoning: clinical expertise. *Med Educ* 2007;41(12):1173–1177. <https://doi.org/10.1111/j.1365-2923.2007.02911.x>
18. **Woods NN, Brooks LR, Norman GR.** The role of biomedical knowledge in diagnosis of difficult clinical cases. *Adv Health Sci Educ* 2007;12(4):417–426. <https://doi.org/10.1007/s10459-006-9054-y>
19. **Baghdady MT, Pharoah MJ, Regehr G, Lam EWN, Woods NN.** The Role of Basic Sciences in Diagnostic Oral Radiology. *J Dent Educ* 2009;73(10):1187–1193.
20. **Baghdady MT, Carnahan H, Lam EW, Woods NN.** Integration of basic sciences and clinical sciences in oral radiology education for dental students. *J Dent Educ* 2013;77(6):757–763.
21. **Baghdady MT, Carnahan H, Lam EWN, Woods NN.** Dental and Dental Hygiene Students' Diagnostic Accuracy in Oral Radiology: Effect of Diagnostic Strategy and Instructional Method. *J Dent Educ* 2014;78(9):1279–1285.

22. **Baghdady MT, Carnahan H, Lam EWN, Woods NN.** Test-enhanced learning and its effect on comprehension and diagnostic accuracy. *Med Educ* 2014;48(2):181–188. <https://doi.org/10.1111/medu.12302>
23. **Kulasegaram KM, Min C, Howey E, et al.** The mediating effect of context variation in mixed practice for transfer of basic science. *Adv Health Sci Educ* 2014;1–16. <https://doi.org/10.1007/s10459-014-9574-9>
24. **Kulasegaram KM, Manzone JC, Ku C, Skye A, Wadey V, Woods NN.** Cause and Effect: Testing a Mechanism and Method for the Cognitive Integration of Basic Science. *Acad Med* 2015;90:S63–S69. <https://doi.org/10.1097/ACM.0000000000000896>
25. **Kulasegaram KM, Chaudhary Z, Woods N, Dore K, Neville A, Norman G.** Contexts, concepts and cognition: principles for the transfer of basic science knowledge. *Med Educ* 2017;51(2):184–195. <https://doi.org/10.1111/medu.13145>
26. **Castillo J-M, Park YS, Harris I, et al.** A critical narrative review of transfer of basic science knowledge in health professions education. *Med Educ* 2018;52(6):592–604. <https://doi.org/10.1111/medu.13519>
27. **Kulasegaram KM, Martimianakis MA, Mylopoulos M, Whitehead CR, Woods NN.** Cognition Before Curriculum: Rethinking the Integration of Basic Science and Clinical Learning. *Acad Med* 2013;88(10):1578–1585. <https://doi.org/10.1097/ACM.0b013e3182a45def>
28. **Cheung JJH, Kulasegaram KM, Woods NN, Moulton C, Ringsted CV, Brydges R.** Knowing How and Knowing Why: testing the effect of instruction designed for cognitive integration on procedural skills transfer. *Adv Health Sci Educ* 2017. <https://doi.org/10.1007/s10459-017-9774-1>
29. **Brydges R, Nair P, Ma I, Shanks D, Hatala R.** Directed self-regulated learning versus instructor-regulated learning in simulation training. *Med Educ* 2012;46(7):648–656. <https://doi.org/10.1111/j.1365-2923.2012.04268.x>
30. **Haji FA, Cheung JJH, Woods N, Regehr G, de Ribaupierre S, Dubrowski A.** Thrive or overload? The effect of task complexity on novices' simulation-based learning. *Med Educ* 2016;50(9):955–968. <https://doi.org/10.1111/medu.13086>
31. **Martin JA, Regehr G, Reznick R, et al.** Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84(2):273–278.
32. **Ilgén JS, Ma IWY, Hatala R, Cook DA.** A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. *Med Educ* 2015;49(2):161–173. <https://doi.org/10.1111/medu.12621>
33. **Bloch R.** *G_String_IV*. Hamilton, ON; 2017. Available at: http://fhspend.mcmaster.ca/g_string/index.html. Accessed 6 January 2019.
34. **Leppink J.** On causality and mechanisms in medical education research: an example of path analysis. *Perspect Med Educ* 2015;4(2):66–72. <https://doi.org/10.1007/s40037-015-0174-z>
35. **Rucker DD, Preacher KJ, Tormala ZL, Petty RE.** Mediation Analysis in Social Psychology: Current Practices and New Recommendations: Mediation Analysis in Social Psychology. *Soc Personal Psychol Compass* 2011;5(6):359–371. <https://doi.org/10.1111/j.1751-9004.2011.00355.x>
36. **Hayes AF.** Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York: The Guilford Press; 2013.
37. **Motola I, Devine LA, Chung HS, Sullivan JE, Issenberg SB.** Simulation in healthcare education: A best evidence practical guide. *AMEE Guide No. 82. Med Teach.* 2013;35(10):e1511–e1530. <https://doi.org/10.3109/0142159X.2013.818632>
38. **Cook DA, Brydges R, Hamstra SJ, et al.** Comparative Effectiveness of Technology-Enhanced Simulation Versus Other Instructional Methods: A Systematic Review and Meta-Analysis. *Simul Healthc* 2012;7(5):308–320. <https://doi.org/10.1097/SIH.0b013e3182614f95>
39. **Van Merriënboer JJG, Sweller J.** Cognitive load theory in health professional education: design principles and strategies: Cognitive load theory. *Med Educ* 2010;44(1):85–93. <https://doi.org/10.1111/j.1365-2923.2009.03498.x>
40. **Eva KW, Neville AJ, Norman GR.** Exploring the etiology of content specificity: factors influencing analogic transfer and problem solving. *Acad Med* 1998;73(10 Suppl):S1–5.
41. **Norman G.** Teaching basic science to optimize transfer. *Med Teach* 2009;31(9):807–811. <https://doi.org/10.1080/01421590903049814>
42. **Goldszmidt M, Minda JP, Devantier SL, Skye AL, Woods NN.** Expanding the basic science debate: the role of physics knowledge in interpreting clinical findings. *Adv Health Sci Educ* 2012;17(4):547–555. <https://doi.org/10.1007/s10459-011-9331-2>
43. **Bandiera G, Kupern A, Mylopoulos M, et al.** Back from basics: integration of science and practice in medical education. *Med Educ* 2017;52(1):78–85. <https://doi.org/10.1111/medu.13386>
44. **Eva KW, Brady C, Pearson M, Seto K.** The pedagogical value of testing: how far does it extend? *Adv Health Sci Educ* 2018;1–14. <https://doi.org/10.1007/s10459-018-9831-4>

Publisher's Note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.