Adv in Health Sci Educ (2018) 23:61–74
https://doi.org/10.1007/s10459-017-9774-1

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

Jeffrey J. H. Cheung[1,2] · Kulamakan M. Kulasegaram[1,3] · Nicole N. Woods[1,3] · Carol-anne Moulton[1,4] · Charlotte V. Ringsted[5] · Ryan Brydges[1,6,7]

Received: 11 November 2016 / Accepted: 14 April 2017 / Published online: 22 April 2017

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-min 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 no 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.

✉ Jeffrey J. H. Cheung
jeffreyjcheung@gmail.com

1 The Wilson Centre, University Health Network, University of Toronto, Toronto, Canada

2 SickKids Learning Institute, Hospital for Sick Children, Toronto, Canada

3 Department of Family and Community Medicine, University of Toronto, Toronto, Canada

4 Department of Surgery, University of Toronto, Toronto, Canada

5 Centre for Health Science Education, Faculty of Health, Aarhus University, Århus, Denmark

6 Department of Medicine, University of Toronto, Toronto, Canada

7 Allan Waters Family Simulation Centre, St. Michael's Hospital, Toronto, Canada

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.

Keywords Basic science · Cognition · Instructional design · Integration · Transfer · Procedural skills · Simulation · Path analysis

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 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 and VanLehn 2012; Needham and 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 (Kulasegaram et al. 2013, 2015). 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; Kulasegaram et al. 2013, 2015). Such integrated instruction appears to enhance novice learners' retention of clinical reasoning performance (Baghdady et al. 2009, 2013, 2014a, b; Woods et al. 2005, 2006, 2007a), 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 and Allard 1993). Much research shows that encouraging learners to engage in these cognitive problem-solving operations improves skill acquisition and transfer (Guadagnoli and Lee 2004; Lee et al. 1994). These benefits have been demonstrated via the superiority of random versus blocked practice (Shea and Morgan 1979; Shea and Zimny 1983, 1988) delayed versus concurrent feedback schedules (Lee et al. 1990; Winstein and Schmidt 1990), and modelling performances with augmented feedback versus without (McCullagh and Caird 1990). Learners engaged in these beneficial elaborative conceptual processes (Shea and Morgan 1979) are thought to create more meaningful and memorable representations of the movement task (Schmidt and Lee 2005). Those representations are believed to improve subsequent retention and transfer for learning both psychomotor and cognitive skills (Lee et al. 1994; Schmidt and Bjork 1992). 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. 2011, 2012, 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 and 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 et al. 2012, 2015; Rittle-Johnson and Schneider 2015). 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.

Methods

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 et al. 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 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 et al. 2011).

Learning materials

We modified pre-existing videos (Haji 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 Fig. 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 min and the How + Why group video was 17:39 min.

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 min, 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 min, 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

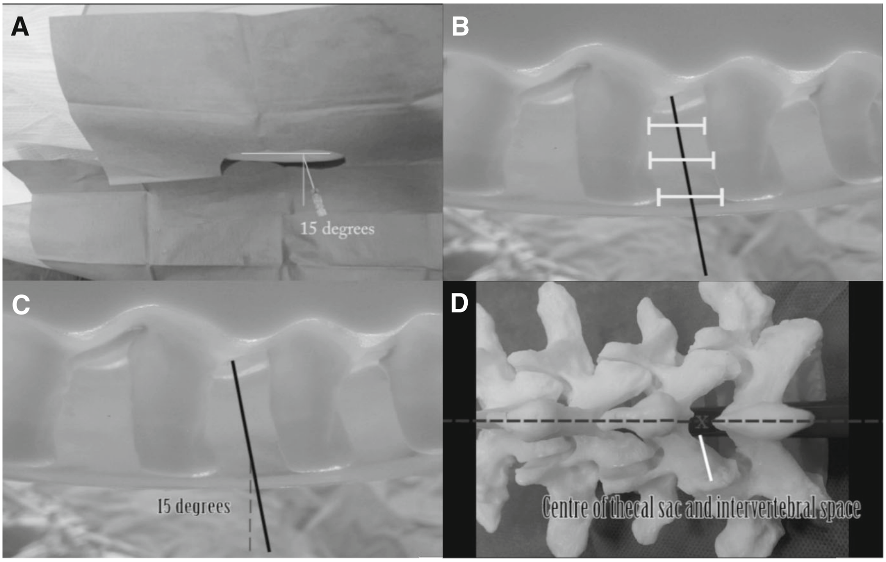

Fig. 1 Screen captures of instructional video demonstrating procedural *how* instruction and conceptual *why* explanations. **Procedural *how* instruction**: when inserting the spinal needle, both groups were provided these instructions in the video: "Confirm the site with your non-dominant hand and 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 allow 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**)]

enhance learning of motor skills compared to instructor-controlled practice (Cheung et al. 2016; Domuracki et al. 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 1-week later to complete the retention and transfer test scenarios. As a final step, participants completed the same procedural and conceptual knowledge tests. Figure 2 presents a summary of the study design.

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):

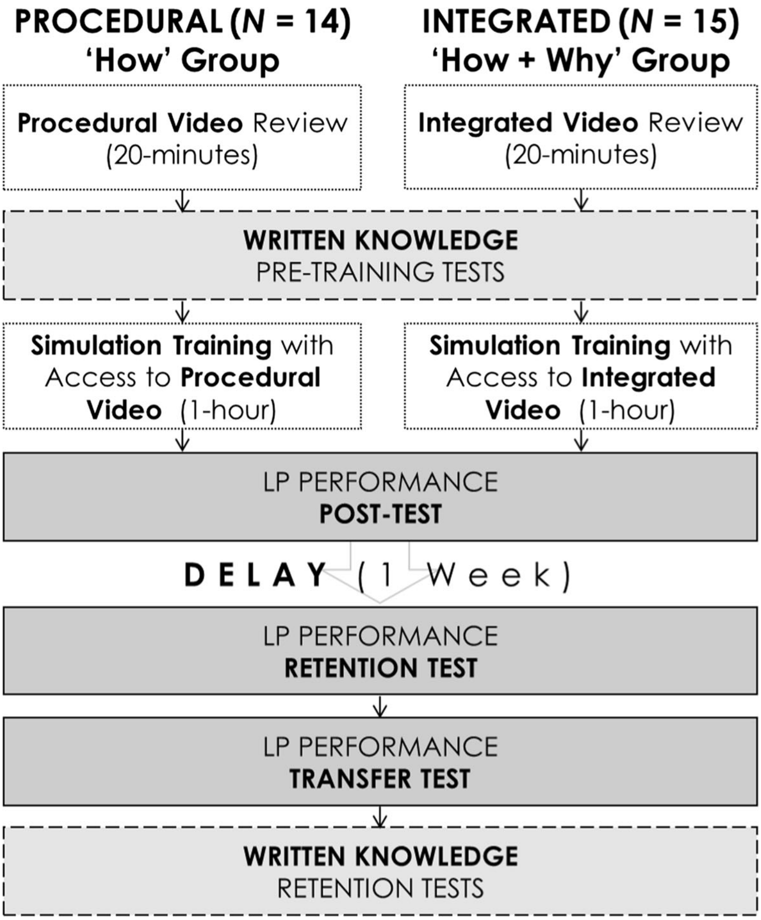

Fig. 2 Summary of study design and procedure with group sample sizes from final analysis

(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)

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 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.

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 et al. 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 validity evidence for GRS versus CL data in simulation-based studies (Brydges et al. 2012; Ilgen et al. 2015), we consider GRS scores on retention and transfer test as the primary outcomes. During a 1-h 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.

Statistical analyses

To compare conceptual and procedural knowledge test scores between groups and across pre-training and retention tests we used a 2 $\times$ 2 (group $\times$ 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 (Fig. 3A), as well as the direct and indirect effect of our group treatment when including conceptual knowledge scores as a mediator in the model (Fig. 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., $a \times b$).

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 ± SE. All outcome measures are summarized in Table 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.

LP performance tests

Path c in Fig. 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

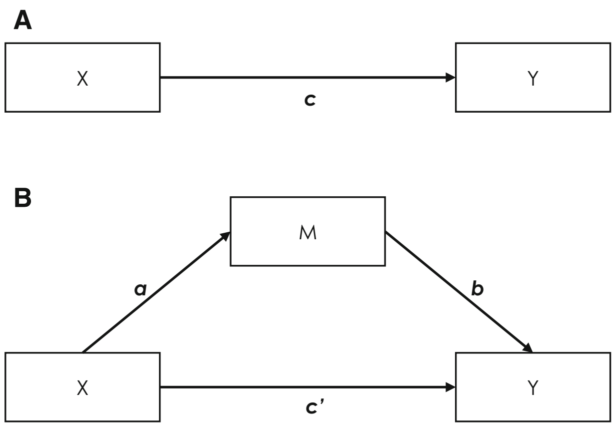

Fig. 3 Regression analyses for total effect of intervention X on outcome variable Y (c) (**A**) and regression analysis for indirect effect of intervention X via mediator variable M ($a \times b$) (**B**), i.e., simple mediation analysis. Respective path coefficients are represented by c, a, b, and c'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

Table 1 Mean group scores ± SE 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

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 ($c' = -.33$).

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

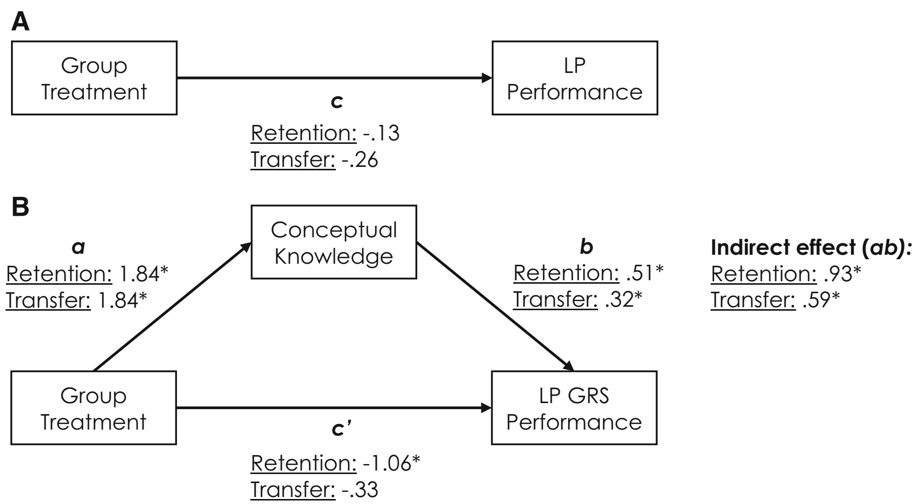

Fig. 4 Unstandardized coefficients of paths in the mediation analysis are represent by *a*, *b*, c' for the outcome variable of GRS of retention (**A**) and transfer performances (**B**). Group treatment effect on conceptual knowledge by *a*, conceptual knowledge effect on GRS retention and transfer 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*. $^*p < .05$

benefit diagnostic ability through the resulting *conceptual coherence* developed by learners (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.

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), rather than at the session level where cognitive integration is best supported (e.g., scheduling exactly how simulation will be used during each 4-h half-day) (Kulasegaram et al. 2013). Findings from education (Bransford and Schwartz 1999; DeCaro and Rittle-Johnson 2012; Kapur 2014; Rittle-Johnson et al. 2012; Schwartz and Bransford 1998) psychology (Needham and Begg 1991), and in clinical reasoning (Mylopoulos et al. 2016; Mylopoulos and 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.

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 et al. 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.

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.

Acknowledgements This study was supported through operational funds from the BMO Chair in Health Professions Education Research and graduate student funding from the NSERC Alexander Graham Bell Canada Graduate Scholarship. The authors would like to thank the Mount Sinai Surgical Skills Centre for their logistical support and Dr. Faizal Haji, Dr. Gianni Lorello, Dr. Malika Sharma, and Dr. Brandon Girardi for sharing their clinical expertise in the development of the educational materials. We also thank Dr. Jimmie Leppink for his assistance with statistical methodology.

References

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

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., 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. doi:10.1111/medu.12302.

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.

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). Mahwah, NJ: Lawrence Erlbaum Associates. http://site.ebrary.com/id/10084525. Accessed 29 July, 2015.

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. doi:10.2307/30034952.

Boet, S., Bould, M. D., Fung, L., Qosa, H., Perrier, L., Tavares, W., et al. (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. doi:10.1007/s12630-014-0143-8.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of research in education, 24,* 61–100.

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., 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. doi:10.1111/j.1365-2923.2012.04268.x.

Chamberland, M., Mamede, S., St-Onge, C., Rivard, M.-A., Setrakian, J., Lévesque, A., et al. (2013). Students' self-explanations while solving unfamiliar cases: the role of biomedical knowledge. *Medical Education, 47*(11), 1109–1116. doi:10.1111/medu.12253.

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*, *11*(5), 316–322. doi:10.1097/SIH.0000000000000171.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist, 47*(3), 177–188. doi:10.1080/00461520.2012.695709.

Cook, D. A., Brydges, R., Hamstra, S. J., Zendejas, B., Szostek, J. H., Wang, A. T., et al. (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. doi:10.1097/SIH.0b013e3182614f95.

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

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., et al. (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. doi:10.1001/jama.2011.1234.

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 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. doi:10.1016/j.jecp.2012.06.009.

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. doi: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.

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. doi:10.3200/JMBR.36.2.212-224.

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. doi:10.1111/medu.13086.

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. doi:10.1097/ACM.0000000000000130.

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

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. doi:10.1111/medu.12621.

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. doi:10.1080/01421590500046924.

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science, 38*(5), 1008–1022. doi:10.1111/cogs.12107.

Kulasegaram, K. M., 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. doi:10.1097/ACM.0000000000000896.

Kulasegaram, K. M., 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. doi:10.1097/ACM.0b013e3182a45def.

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. doi:10.1197/j.aem.2005.01.014.

Larsen, D. P., Butler, A. C., & Roediger, H. L., III. (2008). Test-enhanced learning in medical education. *Medical Education, 42*(10), 959–966. doi: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. doi:10.1007/s40037-015-0174-z.

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

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.

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. doi:10.1111/j.1365-2923.2009.03547.x.

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. doi:10.1111/medu.12893.

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

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.

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

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). Oxford: Oxford University Press. 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. doi:10.1007/s10648-015-9302-x.

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. doi:10.1111/j.2044-8279.2011.02037.x.

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. doi:10.1111/j.1751-9004.2011.00355.x.

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

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three. http://www2.psychology.uiowa.edu/Classes/31330/MotorControl/SchmidtBjork1992.pdf. Accessed 24 March 2014.

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

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–5223. doi:10.1207/s1532690xci1604_4.

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. *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.

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

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. doi: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*. http://www.royalcollege.ca/cs/groups/public/documents/document/y2vk/mdaw/~edisp/tztest3rcpsced000901.pdf

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. doi: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. doi:10.1111/j.1365-2929.2004.02036.x.

Woods, N. N., Brooks, L. R., & Norman, G. R. (2007a). It all make sense: Biomedical knowledge, causal connections and memory in the novice diagnostician. *Advances in Health Sciences Education, 12*(4), 405–415. doi: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. doi:10.1007/s10459-006-9054-y.

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. doi:10.1016/j.physio.2006.08.005.