Contexts, concepts and cognition: principles for the transfer of basic science knowledge

Kulamakan M Kulasegaram,[1] Zarah Chaudhary,[1] Nicole Woods,[1] Kelly Dore, [2] Alan Neville[3] & Geoffrey Norman[1,2]

CONTEXT Transfer of basic science aids novices in the development of clinical reasoning. The literature suggests that although transfer is often difficult for novices, it can be optimised by two complementary strategies: (i) focusing learners on conceptual knowledge of basic science or (ii) exposing learners to multiple contexts in which the basic science concepts may apply. The relative efficacy of each strategy as well as the mechanisms that facilitate transfer are unknown. In two sequential experiments, we compared both strategies and explored mechanistic changes in how learners address new transfer problems.

METHODS Experiment 1 was a 2 × 3 design in which participants were randomised to learn three physiology concepts with or without emphasis on the conceptual structure of basic science via illustrative analogies and by means of one, two or three contexts during practice (operationalised as organ systems). Transfer of these concepts to explain pathologies in familiar organ systems (near transfer) and unfamiliar organ systems (far transfer) was evaluated during immediate and delayed testing. Experiment 2 examined whether exposure to conceptual analogies and multiple contexts changed how learners classified new problems.

RESULTS Experiment 1 showed that increasing context variation significantly improved far transfer performance but there was no difference between two and three contexts during practice. Similarly, the increased conceptual analogies led to higher performance for far transfer. Both interventions had independent but additive effects on overall performance. Experiment 2 showed that such analogies and context variation caused learners to shift to using structural characteristics to classify new problems even when there was superficial similarity to previous examples.

CONCLUSIONS Understanding problems based on conceptual structural characteristics is necessary for successful transfer. Transfer of basic science can be optimised by using multiple strategies that collectively emphasise conceptual structure. This means teaching must focus on conserved basic science knowledge and de-emphasise superficial features.

Medical Education 2017: 51: 184–195
doi: 10.1111/medu.13145

[1]The Wilson Centre for Health Professions Education, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada
[2]Programme for Education Research and Development, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada
[3]Department of Oncology, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

Correspondence: Kulamakan (Mahan) Kulasegaram, 200 Elizabeth Street, 1ES-603 Toronto, Ontario, Canada M5G 2C4.
Tel: 416 340 3216; E-mail: mahan.kulasegaram@utoronto.ca

INTRODUCTION

The utility of basic science for teaching clinical reasoning to novices is increasingly recognised. Appropriate application and extension of basic science knowledge[1] can help learners master novel clinical concepts.[2–4] The process of applying and extending knowledge is known as transfer.[1] Despite a long-standing recognition that transfer is desirable, it remains a challenge for students[5] and is not a spontaneous result of teaching.[6,7] For example, teaching a basic science concept such as flow dynamics in the context of one pathology and organ system (e.g. asthma) does not guarantee learners will see the applicability of the concept to new relevant pathologies (e.g. heart murmurs). This form of context-specificity is a consistent finding in several studies addressing the transfer of basic science.[7–10] This is a critical issue as recent studies in clinical expertise show that although experts may not use basic science for all tasks, it remains a critical part of their knowledge structures.[11] This 'encapsulation' enables basic science to be readily available in clinical contexts and for addressing atypical problems.[12,13] Thus, teaching basic science in a manner that makes it available and relevant to future tasks is a critical challenge for undergraduate medical and health professions training.

Two complementary models from instructional science suggest ways to organise basic science instruction for transfer. The first model, schema theory,[14] suggests that learners must abstract and decontextualise conceptual knowledge to transfer it. Studies of transfer in applied settings such as health professions education,[15] mathematics[16] and skills training[17] as well as other domains[18] have tended to focus on emphasising this conceptual knowledge. For example, simple everyday analogies have been paired with teaching explanations in order to allow learners to access the essential conceptual structures of knowledge or the so-called 'deep structure' of concepts.[14,17,18] Similar interventions during practice and other phases of learning[19–23] have supported the efficacy of emphasising this conceptual structural knowledge. This implies that there is extraneous or superficial knowledge that may be necessary for immediate problem solving but does not need to be transferred to new problems. On the other hand, conceptual deep structure is conserved across problem contexts and must be understood for transfer.

An alternative theory derived from exemplar-based expertise research,[10] argues that it is important to emphasise contextual information and learners should be exposed to a variety of contexts during practice. Learners often have a difficult time understanding the complexity of abstract conceptual knowledge (i.e. the meaningfulness of the conceptual deep structure). Providing multiple contexts for the concept can reduce this cognitive load[24] by providing clear concrete examples and can increase the authenticity of learning.[21] This approach may teach learners that contextual surface details are not always consistent across problems. Learners may also be cued to which features of the context are relevant. In medical problem solving, the features of diagnostic tasks or a physiology problem are often relevant clues to the underlying conceptual structure of a problem.[7,25–27] Expert performance in domains such as diagnosis and management relies heavily on recognition of important details in the context of the problem. Thus, learning and practising for transfer in multiple contexts provide the application of experience and adaptation of knowledge necessary for transfer.[24]

Previously, we have established the efficacy of emphasising conceptual structural knowledge using teaching analogies[7] and the benefits of contextual variation by exposing learners to multiple examples of surface features during practice.[28] These studies also examined the effects of each set of interventions on near transfer (application to familiar contexts) and far transfer (application to novel contexts) of physiology concepts. There is ample evidence for the effectiveness of both strategies[29–33] but to our knowledge there are no head to head comparisons of these strategies.

It would be beneficial to understand the relative importance and efficacy of each class of intervention for both practical and theoretical reasons. Curricular time is increasingly contested. If both strategies are equivalent, teachers can simply choose the strategy that makes the most logistical sense. Furthermore, we do not know if these types of intervention lead to changes in how students process new transfer problems. As Cook and colleagues have argued,[34] medical education research needs clarification studies to understand the generalisable mechanisms of learning. Transfer is optimised when students develop the appropriate cognitive processes to recognise and solve future problems. If the mechanisms or processes supported by each intervention can be clarified, then existing curricular teaching methods can be modified to facilitate these processes.

The purpose of this research is to understand how emphasising conceptual or deep structure knowledge versus contextual variation during learning mediates transfer of basic science concepts. In experiment 1, we compare interventions for their efficacy in near and far transfer. In experiment 2, we infer changes in learners' cognitive processing as a result of exposure to these interventions. Thus whereas the first experiment demonstrates effects, the second experiment illustrates potential mechanistic explanations for the results.

EXPERIMENT 1

In a factorial design, we conducted a head to head comparison of the effectiveness of deep structure emphasis versus contextual variation for the transfer of basic physiology concepts. Participants were randomised to receive increased conceptual deep structure emphasis or not and then randomised again into one of three context variation conditions. Conceptual deep structure was increased for one group of students by concept-illustrating teaching analogies, whereas context variation consisted of manipulating the number of organ systems with which concepts were practised.

Methods

Procedures and methods are similar to those of previous studies of transfer in medical education.[7,15,28] McMaster University undergraduate students taking a first-year psychology course were recruited into this study for a mandatory participation credit (all participants received a credit for completion of the study). Recruitment was on a volunteer basis and advertised to all first-year psychology students. This population was chosen for feasibility reasons but also because they were generally new to the anatomy and physiology concepts used in the experiment. Health professions students at our institution receive early instruction on these concepts and thus were considered already exposed to the content. A naïve population allows us to study instructional design issues without compromising the existing curriculum or contaminating effects as a result of prior knowledge of the teaching material. The study received ethics approval from the Faculty of Health Sciences Research Ethics Board at McMaster.

Design

Participants were randomised in a 2 × 3 design to two instruction and three practice conditions. Randomisation was carried out using a pseudo-random number generator and sequential allocation. Investigators were blinded to the allocation status. Each participant in this study learned three physiology concepts using a standard clinical explanation and diagram provided by an expert clinician (AN) or the standard explanation and an analogy illustrating deep structure. The teaching analogy was an accessible, common-sense illustration of the deep structure of each physiology concept. Teaching analogies have been used previously in many experimental studies of mathematics,[30] physics[19] and physiology teaching.[20] Contextual variation was manipulated during practice and learners had the opportunity to practise applying concepts to three clinical cases that involved the same organ system, two organ systems or three different organ systems. Following learning and practice, participants were immediately given a transfer test, and they returned after a 1-week delay to complete a new transfer test. This timeframe was chosen for feasibility and for similarity to previous studies of this topic.[3,7] Additionally, a review by Cepeda and colleagues suggests that measurements of memory recall at 1 week are predictive of performance at 30 days.[35] Transfer cases were clinical vignettes involving organ systems that learners had been exposed to during learning (near transfer) and organ systems that they had not been exposed to (far transfer). The delay condition was included to examine the durability of the interventions over time. All materials in the study were designed by a content expert with clinical expertise (AN).

Materials

Participants in this experiment were given explanations for three physiology-related concepts (laminar and turbulent flow, Laplace's law and Starling's law) as well as practice applying these concepts to explain the signs and symptoms depicted in clinical vignettes. Laminar and turbulent flow illustrates the principles of flow dynamics, which have application in multiple organ systems, including the respiratory system, cardiovascular system, gastrointestinal system and the urinary tract. Laplace's law describes tension in the walls of cylindrical vessels as it relates to the radius of the vessel and pressure of the contents; it applies to the same organ systems as the principles of laminar and turbulent flow. Starling's law describes the elastic behaviour of the heart in response to filling of the ventricles and volume of the ejected blood. Explanations included a detailed outline of the concept, its relevance to physiology and a brief example of its application, and emphasised that the concept applied to multiple organ systems (see Table S1). Practice and testing were through the use of clinical vignettes. Vignettes described a brief clinical encounter and patient presentation involving pathology located to one organ system and explainable through one of the three concepts participants were given.

Manipulation of conceptual deep structure knowledge
Participants were randomised to instructional conditions in which they were presented with either a basic clinical explanation of the physiology concepts or with the clinical explanation and an additional deep structure teaching analogy that illustrated each concept. These analogies were previously used in studies of transfer[7,15] and were shown to increase transfer performance at testing. An example of a teaching analogy is presented in Table S1.

Context variation in practice manipulation
Individuals in the one-organ-system practice condition saw three practice cases that involved the same organ system. For laminar and turbulent flow the cases all involved respiratory disorders and for Laplace's law they involved gastrointestinal disorders. In the two-organ-systems condition, laminar and turbulent flow involved two cardiovascular cases and one respiratory or gastrointestinal case. In the three-organ-systems condition, cases were taken from all three organ systems. Starling's law only applies to the cardiovascular system and was practised only with cardiovascular cases in each condition. It was added as an additional near-transfer distractor.

Procedure

A summary of the design and procedure is available in Fig. S1. The entire study was conducted on a computer on a preprogramed software platform. The study was conducted in two phases: learning and testing.

Phase 1: learning
In Phase 1 participants were given the explanation of the concept in a randomised order. After viewing each explanation, participants were given three practice cases in sequence. Each concept was practised using three practice cases. An example of a practice case is given in Table S2. The participants were asked to read each case and explain how the concept they had just encountered applied to and explained the features of the patient in the case. Feedback was given in the form of the correct answer to how the previously learned concept applied to the case.

After completing the learning and practice, all participants were asked to complete a 10-question true or false quiz on the concepts they previously learned as additional practice. Participants had to achieve a score of 8/10 or greater to move onto testing or they were asked to repeat the test before moving onto transfer testing.

Phase 2: testing
Transfer testing consisted of 15 short, written clinical cases. Participants were required to identify the concept in the case and provide an explanation of how the concept applied to the case. Cases involved organ systems that were familiar and unfamiliar to all participants regardless of condition to create near and far transfer cases. Six cases involved laminar and turbulent flow, six involved Laplace's law and three involved Starling's law. The cardiovascular organ system was common to all three concepts; the gastrointestinal and respiratory systems were common to laminar and turbulent flow and Laplace's law. Far transfer cases for laminar and turbulent flow and Laplace's law were presented using the urinary tract, spinal cord and reproductive systems, which the participants had not seen during learning.

Responses were scored on a scale of 0–3, with more marks awarded for more accurate and in-depth explanations, as used in previous studies with these materials.[7,15] A subsample of responses was coded in duplicate by two raters to determine inter-rater reliability for consistency of scoring. See Table S3 for an example of a test case and scoring rubric.

Test vignettes were presented in random order. After completing the initial transfer test, participants returned after a 1-week delay to complete another transfer test with 15 new vignettes.

Outcomes and analysis

The primary outcome was the average score per case, which was calculated separately for near and far transfer cases. We also collected data on time spent during the learning phase and time spent per transfer case to determine if time on task was a significant factor for performance.

The primary analysis was a 2 × 2 × 3 repeat measures ANOVA with near versus far transfer as the repeated variable and instruction condition (analogy versus no analogy) and practice condition (one versus two versus three organ systems) as the between subjects variables. We replicated these analyses with scores on the delayed transfer test and conducted an additional repeated measures analysis with time to account for the impact of delayed testing. The least significant differences test (LSD) was used to test for post-hoc differences where appropriate. A separate sub-analysis of Starling's law cases (near transfer for all groups) was conducted as a manipulation. We examined correlations between the time variables and performance. The alpha value for all analyses was 0.05 two-tailed.

Results

Ninety undergraduate students ($n = 90$) were recruited into the study with 15 participants per group. Two students were excluded from the final analyses for not completing the transfer test. This left 44 participants allocated to the analogy and no-analogy conditions each; 29 participants remained in the one-organ-system and two-organ-system groups and 30 in the three-organ-system group. Inter-rater reliability between two blinded raters for consistency of scoring was 0.83, suggesting high consistency.

Immediate testing

A repeated measure of performance on near and far transfer showed a general main effect for near transfer ($F(1,82) = 5.92$, $p < 0.017$, $\eta_p^2 = 0.07$), with average performance on near transfer cases at 1.09 (SD = 0.51) compared with far transfer at 0.96 (SD = 0.45). Significant interactions were also found with the analogy manipulation ($F(1,82) = 7.35$, $p < 0.01$, $\eta_p^2 = 0.08$) and practice condition ($F(1,82) = 13.87$, $p < 0.0001$, $\eta_p^2 = 0.25$). Specifically, performance on near transfer was higher for the non-analogy condition but dropped for far transfer, whereas the analogy condition maintained similar performance on near and far transfer (see Fig. 1). Far transfer performance was higher for the analogy group 1.06 (0.50) compared with the no analogy group 0.86 (0.37). Post-hoc testing showed this difference to be significant ($F(1,82) = 4.6$, $p = 0.035$).

Similarly, the one-organ-system group had high near transfer scores but dropped for far transfer, whereas the two and three-organ-system groups maintained performance across both transfer tasks (see Fig. 2) ($F(1,82) = 13.87$, $p < 0.0001$). The one-organ-system group had the highest near transfer performance at 1.28 (0.57), compared with the two-organ-system group at 0.94 (0.41) and three-organ-system group at 1.04 (0.47). By contrast, the highest far transfer performance was seen for the two-organ-system group at 1.03 (0.48) and three-organ-system group at 1.07 (0.46), compared with the one-organ-system group at 0.78 (0.34). Post-hoc testing on far

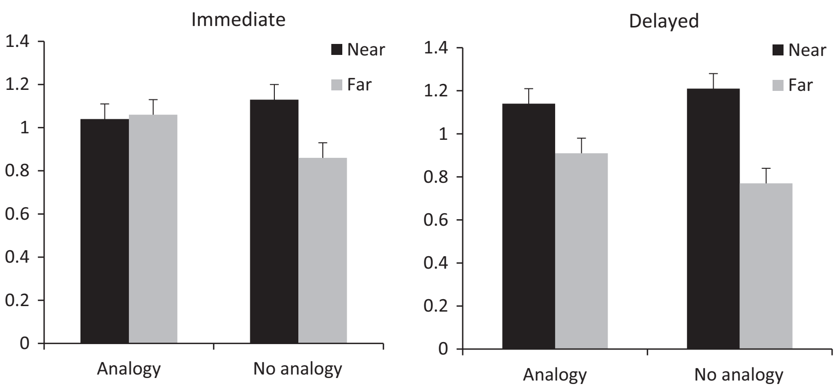

Figure 1 A significant interaction between mean near and far transfer score at immediate and delayed testing. Immediate test showed a significant for near vs. far between the Analogy and No Analogy conditions which was replicated at delayed test ($F(1,82) = 7.35$, $p < 0.01$). The analogy condition maintained similar performance across both transfer tasks while the No Analogy instruction group dropped significantly for far transfer on both tests. Far transfer performance dropped at retention testing ($F(1,69) = 8.06$, $p < 0.006$) but no difference was detected for near transfer.

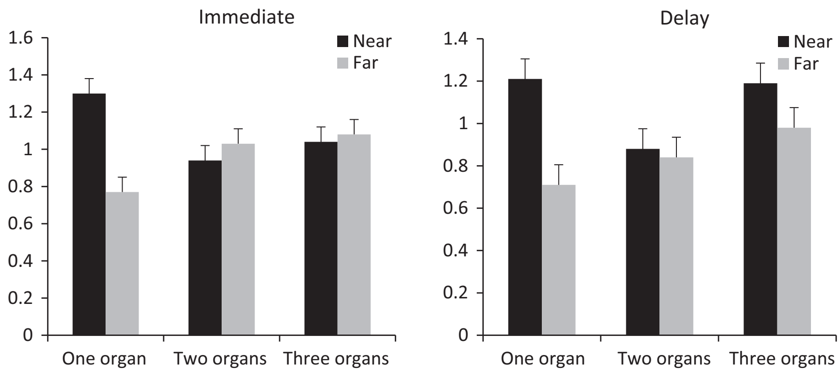

Figure 2 A significant interaction was detected for near vs. far and the number of organ systems at immediate (F (1,82) = 13.87, p < 0.0001) and delayed test (F(2,69) = 5.72, p < 0.005, η_p^2 = 0.14). The one-organ-system had the highest near transfer performance but performance dropped for far transfer. The two- and three-organ-system groups had similar near and far transfer scores; average far transfer score was higher for the two- and three-organ-system groups compared to the one-organ-system group.

transfer showed that there were no significant differences between two- and three-organ-system groups but both groups were significantly higher than the one-organ-system group. Analysis of only the Starling's law transfer case performance was not significant between conditions.

Delay

Thirteen participants (n = 13) did not return for delayed testing, limiting sample size for delayed analysis to 75. This left 35 and 40 participants in the analogy and no-analogy conditions; there were 24, 27 and 24 participants in the one-, two- and three-organ-system conditions.

A significant effect of time delay was seen for far transfer performance (F(1,69) = 8.06, p < 0.006, η_p^2 = 0.11), with immediate far transfer performance at 0.98 (0.41) and delayed performance at 0.85 (0.38). The 1-week time delay did not significantly decrease near transfer performance (F(1,69) = 0.051, p < 0.82, η_p^2 = 0.001).

The same pattern of results was seen as with the immediate test, with significant interactions between instructional condition (F(1,69) = 7.78, p < 0.007, η_p^2 = 0.10) and near versus far transfer performance as well as practice condition and near versus far transfer (F(2,69) = 5.72, p < 0.005, η_p^2 = 0.14). Once again, the analogy condition had a smaller difference between near and far transfer performance (0.99 versus 0.91) compared with the no-analogy group (1.21 versus 0.77) (see Fig. 1). As at the immediate test, the one-organ-system group had high near transfer performance but the lowest performance for far transfer, whereas the two- and three-organ-system practice groups maintained similar scores for both types of transfer tasks (see Fig. 2). The far transfer scores between the two- and three-organ-system groups were not significantly different. Once again the performance of Starling's law did not differ between the manipulations but was impacted by delay (F(1,69) = 4.2, p < 0.05).

Discussion

Experiment 1 demonstrated that emphasising deep structure and exposure to multiple contexts, independently and additively contribute to improve far transfer performance. Of note, increasing the number of contexts from two to three had marginal benefit for transfer performance. This suggests that even the limited contextual variation facilitates greater structure abstraction[32,33] and two examples of contexts are sufficient for this. This latter finding suggests that although context variation promotes transfer, it may do so by showing learners that contextual information may not always be a reliable cue for concept retrieval. Thus, the predictions of schema theory seem to be more salient for transfer of basic science.

The one-organ-system and no-analogies group had high near transfer but very low far transfer performance in comparison to the other conditions. However there was no near–far difference for the teaching analogy and two- or three-organ-system groups, and their far transfer scores outperformed the no-analogy and one-organ-system groups. This finding hints at a possible mechanism for how the learning interventions change performance during transfer problem solving. Relying on surface features as clues to conceptual structure can be an effective and efficient strategy *when there is a close association between surface and structure*. In learning conditions where this association is reinforced (i.e. practice with only a single organ system), it is possible learners over-contextualise their knowledge. Some transfer researchers have gone so far as to argue that this is a general 'default' for learners.[10,14,24] This would continue to support the theory that successful transfer results from greater abstraction instead of reliance on contextualisation.

To test whether this was indeed the mechanism of action for the transfer interventions, experiment 2 examined how analogies and multiple contexts changed the categorisation of new problems by learners. We hypothesised that learners would be more likely to classify problems on the basis of conceptual knowledge if exposed to either teaching analogies or contextual variation. We also modified our conceptual recall test to a multiple choice format. This test was used to determine if increasing conceptual knowledge was related to changes in transfer processing.

EXPERIMENT 2

Experiment 2 was similar to experiment 1 with two notable exceptions. First, practice with three organ systems was dropped because it was not statistically different from the two-organ-system group. This reduced our costs and ease of recruitment. Second, we changed the outcome of this experiment to include a multiple choice question (MCQ) test to understand if individual differences in conceptual learning contributed to transfer performance and we changed the transfer task to a forced classification or forced choice test known as the similarity classification task.

Methods

Undergraduate students (different from experiment 1) taking a first-year psychology course were recruited into this study for participation credit. The study received ethics approval from the Faculty of Health Sciences Research Ethics Board at McMaster.

Design

We used similar design, procedures and materials in experiment 2 as in the first experiment, with three major changes (see Fig. S2). In a 2 × 2 design, participants were randomised to learn physiology concepts with a basic clinical explanation or basic clinical explanation. Participants were then randomised again to practising with one or two organ systems for laminar flow and Laplace's law. As with the previous experiment, Starling's law was practised with one organ system. After completing learning, participants took a multiple-choice test to test recall and a similarity categorisation test, which required them to make similarity judgements in order to classify written medical cases. Lastly, the delay condition was eliminated to make participation more convenient for participants and because delay did not affect the pattern of results in experiment 1.

Procedure

Phase I: learning

The learning phase used the same materials as in experiment 1, with the exception that participants only practised with two practice cases, with feedback in the form of the correct answer. Participants completed a 10-question true or false quiz prior to testing. They were required to achieve a minimum score of 8/10 to move on; if not they were asked to repeat the test.

Phase II: testing

In the testing phase, all participants were first given a 15-question multiple-choice test that tested their recall of the concept explanations. Questions were focused solely on their understanding of the explanations given of each concept. Participants were not given feedback on their performance before moving on to the similarity classification test.

Similarity classification test

The similarity classification test was framed as a forced-choice recognition task[31] and is similar to the concept-sorting[36] developed by McLaughlin and colleagues. Participants were asked to view a clinical case designated the target case. Participants were not told which concept was involved and given instructions to read and think about the vignette but were not required to explain the features or answer any questions. After completely reading the case, they proceeded to view three new cases and were asked to select the case that was most similar to the previous target case. Participants were not given any instructions on what constituted similarity. Time per target and match decision was limited to 5 minutes.

The three choices were manipulated to have (i) surface similarity to the target, (ii) structural similarity to the target or (iii) both surface and structural similarity or no similarity. For example, if the target case involved a gastrointestinal disorder explained by Laplace's law then the possible matches could be (i) a gastrointestinal disorder explained by laminar turbulent flow, (ii) a cardiovascular disorder explained by Laplace's law or (iii) a gastrointestinal disorder explained by Laplace's law or a cardiovascular disorder explained by Starling's law.

Participants using contextual surface information to identify similarity would select case (i), whereas those matching on conceptual structural features could pick case (ii). The third case was used as a distracter or control. Participants had to complete 10 classifications, with four involving laminar flow cases from various organ systems, four involving Laplace's law and two involving Starling's law. Clinical cases were similar to those in experiment 1.

The outcomes from the forced choice similarity classification test were: (i) *context matches*, the number of times the selected case had the same organ system (i.e. contextual surface detail) as the target case but a different concept (i.e. structural features) explaining the clinical presentation; (ii) *concept matches*, the number of times each participant selected a matching case that had the same concept involved but had a different organ system from the target case; (iii) *double hit*, the number of times a match had the same concept and organ system; and (iv) *double miss*, had neither the same concept or organ system.

Context matches indicate the extent to which surface features were primarily used for categorisation, whereas concept matches indicate the extent to which participants were using structural or conceptual knowledge for categorisation despite conflicting surface features. The primacy of one type of match over the other for a group would indicate the extent to which categorisation of transfer problems had been impacted by the interventions.

Analysis

Analysis of the multiple choice recall test was by a 2 × 2 ANOVA with analogy or no-analogy and one- or two-organ-system as the between subjects factor. Each type of match on the similarity categorisation test was analysed using a 2 × 2 × 2 ANCOVA with the number of context and concept matches as the within subjects factor. Double hits and double misses were considered distractors and their analysis is not reported as it would not provide useful information.

Results

Forty undergraduate students were recruited into the study, with 10 students per group. Average time between conditions did not differ significantly.

Multiple-choice knowledge test

There were no significant differences between any groups on MCQ testing. Mean score (SD) for the no-analogy and one-organ-system group was 10.22 (2.10) and for two-organ-system was 9.27 (1.61). Mean scores for the analogy with one- and two-organ-system were 9.3 (1.76) and 8.45 (2.06), respectively. The MCQ test was not significant in predicting the number of double hit or double miss matches. MCQ scores had a significantly negative correlation with the number of context matches ($r = -0.36$, $p < 0.020$) and non-significant positive correlation with concept matches ($r = 0.11$, $p < 0.48$). Thus the MCQ score was used as a covariate in the analyses of the similarity categorisation test.

Similarity categorisation test

Analysis of the number of context matches versus concept matches (when the matched case had a similar organ system but different concept from the target case versus matches of similar concepts despite a different organ system from the target case) showed a significant interaction with instruction (analogy) ($F(1,35) = 9.4$, $p < 0.004$, $\eta_p^2 = 0.21$). Participants in the no-analogy condition made a similar number of context and concept matches (see Fig. 3), whereas those in the analogy condition made more concept matches. The analysis of the effect of organ system showed a similar within-subjects interaction ($F(1,35) = 4.1$, $p < 0.05$, $\eta_p^2 = 0.10$) (see Fig. 4). The participants in the one-organ-system condition made a similar number of context and concept matches, whereas the two-organ-system group made a greater number of concept matches. There was no interaction between analogy and organ system condition versus type of match ($F(1,35) = 0.004$, $p < 0.9$, $\eta_p^2 = 0.0001$). The MCQ score was a significant interacting covariate ($F(1,35) = 7.9$, $p < 0.008$, $\eta_p^2 = 0.18$). As expected, the MCQ score was a significant negative covariate for contextual similarity matches and a positive covariate for conceptual matches ($F(1,35) = 4.35$, $p < 0.04$). There were no significant between-subject effects. On average, the no-analogy with one-organ-system group made 2.89 matches on contextual similarity, whereas the analogy with two-organ-system

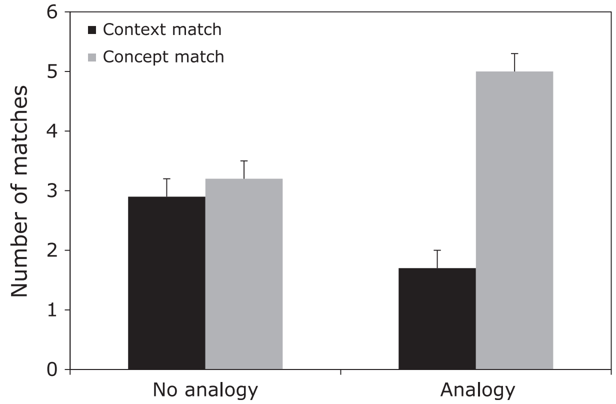

Figure 3 The mean number of Context vs. Concept Matches by Analogy Condition. A significant interaction between mean number of matches on contextual similarity vs. conceptual similarity by Analogy condition ((F,135) = 9.4, $p < 0.004$).

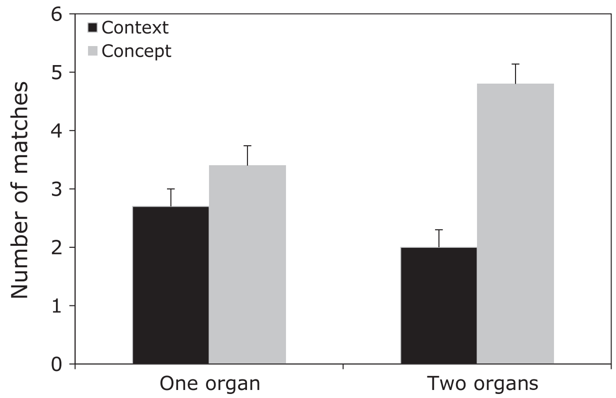

Figure 4 The mean number of Context vs. Concept Matches by Practice Condition. A significiant interaction between mean number of matches on contextual similarity vs. conceptual similarity by practice condition ($F(1,35) = 4.1$, $p < 0.05$).

groups made 1.79 matches. The no-analogy with one-organ-system group made 2.5 matches on conceptual similarity, whereas the analogy with two-organ-system groups made 5.58 matches.

Discussion

Experiment 2 examined whether provision of an emphasis on deep structure via a teaching analogy and contextual variation through increasing the number of practice contexts would lead to a representational shift in how learners classified transfer problems. We hypothesised that the mechanism facilitating transfer was the shift to deep structure and conceptual processing. Evidence for this was shown by the tendency of participants in analogy and two-organ-systems practice conditions to make significantly more similarity matches on conserved deep structure of the underlying concept despite contextual dissimilarity. Furthermore, learners without the benefit of the deep structure analogy or an additional organ system of practice tended to be more likely to make similarity judgements on contextual surface similarity. Although the MCQ score did not differ significantly between groups, it was a significant covariate and negatively predicted context matches. This suggests that better knowledge of deep structure inhibits dependence on surface details.

The results suggest that reliance on surface similarity is the default strategy for unaided learners. Learning interventions that rely on illustration of deep structure, such as teaching analogies or context variation, move learners away from the default strategy by highlighting the inaccuracy of contextual cues. These findings provide an explanation for the improved performance seen in experiment 1 and support previous work that suggests analogies and context variation shift learners to more effective transfer strategies.[37,38]

GENERAL DISCUSSION

Teaching basic science so that it has utility and relevance for learners is a real and practical challenge for medical educators. Given this challenge, what do the results of these two small experimental studies add? Our results in highly controlled experiment settings with novice learners may not immediately translate to educational interventions. Rather, they shed light on the deeper mechanistic processes by which learning can facilitate transfer. It is important that teachers do not reify the interventions *per se* but rather focus on what they represent for teaching as a whole. The purpose of this paper is to help conceptualise the general strategies and mechanisms by which basic science can be taught for transfer.

In experiment 1, we found that learners exposed to analogies illustrating deep structure and more than one organ system during practice had superior far-transfer performance. The results strongly suggest that to facilitate transfer, learners must have the conceptual deep structure of basic science emphasised, along with exposure to multiple clinical examples. We hypothesised that doing this changes the way that learners process or 'see' clinical cases and pushes them to focus on conserved conceptual structure instead of solely relying on contextual information (i.e. the organ system). As evidenced by experiment 2, learners with the successful interventions from experiment 1 were more likely to identify similarity of clinical cases on conceptual basic science despite superficial dissimilarity. This representational shift is necessary for successful far transfer and if unaided learners will default to reliance on surface features. Thus, our results yield two principles of note: (i) novice learners rely heavily on contextual information without understanding the relationship between contextual information and the concept they are supposed to transfer; and (ii) teachers must emphasise structure and conceptual representation of knowledge. The former principle requires that transfer cannot be taken for granted, whereas the latter requires that teaching for transfer may require multiple strategies.

Not every concept in the curriculum can be framed with an analogy or represented in the manner suggested by our interventions. Our results should not be interpreted as providing prescriptive evidence for analogies in teaching or a stable estimate of their relative effect size. Rather, the message of our work is that abstract relationships are difficult to grasp and must be framed appropriately for learners in order for them to grasp the deep structure and create the necessary representational shift. In our study materials, we study transfer of physiology principles derived from physics, where the deep structure was relatively conserved. Other types of concepts may have a different deep structure or relationship to context, although learners will probably still face difficulty with transfer. Teaching that focuses on exposing as much of the appropriate deep structure as possible to students will probably be more successful at helping foster transfer.

Some teaching interventions have been shown to help facilitate this conceptual understanding when employed appropriately.[5] Use of mixed practice has been shown repeatedly to facilitate transfer of learning.[28,39] Mixed or interleaved practice involves learners practising multiple concepts together or cumulatively as opposed to each concept separately. This promotes active comparison of examples and forces learners to identify the distinguishing structural features of concepts. It aligns with the principles emerging from this study showing the importance of emphasising conceptual knowledge. Similarly, showing the structural relationships between basic science knowledge and clinical knowledge promotes the conceptual understanding necessary for clinical reasoning. In other words, effective integration promotes transfer. A series of experimental studies have demonstrated this for near and far-transfer outcomes relevant to practical clinical reasoning.[3,4,40,41]

This study has some limitations common to experimental research. Firstly, it was carried out in a tightly controlled setting and with novice participants. Changing some of these factors, such as allowing student collaboration or including students with some prior knowledge, may reduce the effects detected. It may be that the analogies or examples need to be more complex or sophisticated for a more expert group. Secondly, we had a relatively short time delay before our second assessment of transfer. Although several medical education studies have used this experimental delay,[7,15,40,41] longer-term effects need to be studied in future work.

CONCLUSION

Transfer of basic science concepts can be aided by interventions that make abstract conceptual deep structure explicit and by context variation using examples with dissimilar surface features. Both strategies encourage categorisation and recognition using conceptual knowledge while assisting learners to recognise the limitations of surface features. This shift is necessary for successful far transfer and can be used as an organising framework for teaching basic science for transfer.

Contributors: KK conceived the study, assisted in data collection, conducted the final analyses, completed the first draft and approved the final draft. ZC assisted in data collection and analyses, revised the work critically for important intellectual content and approved the final draft. AN helped conceptualise the study, developed material, revised the work critically for important intellectual content and approved the final draft. KD and NW helped conceptualise the study, assisted in the manuscript preparation and approved the final draft. GN helped conceptualise the study, assisted in the analysis of data, revised the work critically for important intellectual content and approved the first and final drafts.
Acknowledgements: Our thanks to Elizabeth Howey for creation of the data collection software as well as to the Department of Psychology and Program for Education Research and Development at the Faculty of Health Sciences, McMaster University.
Funding: None.
Conflicts of interest: None.
Ethical approval: Given by the McMaster Faculty of Health Sciences Ethics Board.

REFERENCES

1 Eva KW, Neville AJ, Norman GR. Exploring the Etiology of content specificity: factors influencing analogical transfer and problem solving. *Acad Med*, 1998;**73** (10 Suppl):S1–5.
2 Woods NN. Science is fundamental: the role of biomedical knowledge in clinical reasoning. *Med Educ* 2007;**41** (12):1173–7.
3 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 Theory Pract* 2012;**17** (4):547–55.
4 Mylopoulos M, Woods N. Preparing medical students for future learning using basic science instruction. *Med Educ* 2014;**48** (7):667–73.
5 Norman GR. Teaching basic science to optimize transfer. *Med Teach* 2009;**31** (5):807–11.
6 Perkins DN, Salmolon G. Are cognitive skills context-bound? *Educ Res* 1989;**18** (1):16–25.
7 Kulasegaram K, Min C, Ames K, Howey EH, Neville AJ, Norman GR. The effect of conceptual and contextual familiarity on transfer performance. *Adv Health Sci Educ Theory Pract* 2012;**17** (4):489–99.
8 Laksov KB, Lonka K, Josephson A. How do medical teachers address the problem of transfer? *Adv Health Sci Educ Theory Pract* 2008;**13** (3):345–60.
9 Day SB, Goldstone RL. Analogical transfer from a simulated physical system. *J Exp Psychol Learn Mem Cogn* 2011;**37** (3):551–67.
10 Goldstone RL, Son JY. The transfer of scientific principles using concrete and idealized simulations. *J Learn Sci* 2005;**14**:69–110.
11 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):765–73.
12 Rikers RM, Loyens S, te Winkel W, Schmidt HG, Sins PH. The role of biomedical knowledge in clinical reasoning: a lexical decision study. *Acad Med* 2005;**80** (10):945–9.
13 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):1035–43.
14 Reeves L, Weisberg RW. The role of content and abstract information in analogical transfer. *Psychol Bull* 1994;**115** (3):381–400.
15 Norman GR, Dore K, Krebs J, Neville AJ. The power of the plural: effect of conceptual analogies on successful transfer. *Acad Med* 2007;**82** (10 Suppl):S16–18.
16 Kaminski JA, Sloutsky VM, Heckler AF. The cost of concreteness: the effect of nonessential information on analogical transfer. *J Exp Psychol Appl* 2013;**19** (1):14–29.
17 Lowenstein J, Thompson L, Gentner D. Analogical encoding facilitates transfer in negotiation. *Psychol Bull Rev*, 2003;**6**:586–97.
18 Barnett SM, Ceci SJ. When and where do we apply what we learn? A taxonomy for far transfer. *Psychol Bull* 2002;**128** (4):612–37.
19 Paris NA, Glynn SM. Elaborate analogies in science text tools for enhancing preservice teachers' knowledge and attitudes. *Contemp Educ Psychol* 2007;**29**:230–47.
20 Newby TJ, Ertmer PA, Stepich DA. Instructional analogies and the learning of concepts. *Education Tech Research Dev* 1995;**43**:5–18.
21 Van Merrienboer JJG, Kester L, Paas F. Teaching complex rather than simple tasks: balancing intrinsic and germaine load to enhance transfer of learning. *Appl Cogn Psychol* 2006;**20**:343–52.
22 Rohrer D, Taylor K. The shuffling of mathematics practice problems improves learning. *Instr Sci* 2007;**35**:481–98.
23 Needham DR, Begg IM. Problem-oriented training promotes spontaneous analogical transfer: memory-

oriented training promotes memory for training. *Mem Cognit* 1991;**19** (6):543–57.
24 Ross BH, Kilbane MC. Analogical mapping views and the effects of principle explanations. *J Exp Psychol Learn Mem Cogn*, 1997;**23**:427–40.
25 Brooks LR, Norman G, Allen SW. Role of Specific Similarity in a Medical Diagnosis. *J Exp Psychol Gen* 1991;**120**:278–87.
26 Young ME, Brooks LR, Norman GR. The influence of familiar non-diagnostic information on the diagnostic decisions of novices. *Med Educ* 2011;**45** (4):407–14.
27 Dore KL, Brooks LR, Weaver B, Norman GR. Influence of familiar features on diagnosis: instantiated features in an applied setting. *J Exp Psychol Appl* 2012;**18** (1):109–25.
28 Kulasegaram K, Min C, Howey E, Neville A, Woods N, Dore K, Norman G. The mediating effect of context variation in mixed practice for transfer of basic science. *Adv Health Sci Educ Theory Pract* 2015;**20** (4):953–68.
29 Kaminski JA, Sloutsky VM, Heckler AF. The advantage of abstract examples in learning math. *Science* 2008;**25**:454–5.
30 Richland LE, Zur O, Holyoak KJ. Mathematics. Cognitive supports for analogies in the mathematics classroom. *Science* 2007;**316** (5828):1128–9.
31 Gentner D, Loewenstein J, Thompson L. Learning and transfer: a general role for analogical encoding. *J Educ Psychol* 2003;**95**:393–408.
32 Markman AB, Gentner D. Structural alignment during similarity comparisons. *Cogn Psychol* 1993;**25**:431–67.
33 Catrambone R, Holyoak KJ. Overcoming contextual limitations on problem-solving transfer. *J Exp Psychol Learn Mem Cogn*, 1989;**15** (6):1147–56.
34 Cook DA, Bordage G, Schmidt HG. Description, justification and clarification: a framework for classifying the purposes of research in medical education. *Med Educ* 2008;**42** (2):128–33.
35 Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D. Distributed practice in verbal recall tasks: a review and quantitative synthesis. *Psychol Bull* 2006;**132** (3):354–80.
36 McLaughlin K, Mandin H. Using, “concept sorting” to study learning processes and outcomes. *Acad Med* 2002;**77** (8):831–6.
37 Gentner D, Medina J. Similarity and the development of rules. *Cognition* 1998;**65**:263–97.
38 Markman AB, Ross BH. Category use and category learning. *Psychol Bull* 2004;**129** (4):592–613.
39 Hatala RM, Brooks LR, Norman GR. Practice makes perfect: the critical role of mixed practice in the acquisition of ECG interpretation skills. *Adv Health Sci Educ Theory Pract* 2003;**8** (1):17–26.
40 Woods NN, Brooks LR, Norman GR. It all make sense: biomedical knowledge, causal connections and memory in the novice diagnostician. *Adv Health Sci Educ Theory Pract* 2007;**12** (4):405–15.
41 Baghdady MT, Pharoah MJ, Regehr G, Lam EW, Woods NN. The role of basic sciences in diagnostic oral radiology. *J Dent Educ* 2009;**73** (10):1187–93.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. General Design of Experiment 1.

Figure S2. General Design of Experiment 2.

Table S1. Examples of Standard Clinical Explanation and Teaching Analogy.

Table S2. Example Practice Vignette: Goethe’s Law.

Table S3. Sample test vignette and scoring guide.

Received 28 December 2015; editorial comments to author 5 February 2016, accepted for publication 1 June 2016