The Effects of Expert and Augmented Feedback on Learning a Complex Medical Skill

Cecilio-Fernandes, Dario; Cnossen, Fokie; Coster, Jenifer; Jaarsma, Ad Debbie C; Tio, René A

Link to publication in University of Groningen/UMCG research database

The Effects of Expert and Augmented Feedback on Learning a Complex Medical Skill

Dario Cecilio-Fernandes[1,2] (iD),
Fokie Cnossen[3,4] (iD), Jenifer Coster[5],
AD (Debbie) C. Jaarsma[1], and
René A. Tio[6]

Abstract

Many medical skills are complex due to their requirements for integration of declarative (biomedical) knowledge with perceptual–motor and perceptual–cognitive proficiency. While feedback generally helps learners guide their actions, it is unclear how feedback supports the integration of declarative knowledge with skills. Thus, we investigated the effect of expert and augmented feedback on acquisition and retention of a complex medical skill (acquiring a transthoracic echocardiogram) in a simulation study. We randomly assigned 36 medical undergraduate students to one of three feedback sources: Expert (EF), Augmented visual (HS), and Expert plus Help Screen (EF + HS). Participants practiced until reaching proficiency. Outcome measures (knowledge test and practical skill ratings on a 5-point scale), were gathered at initial acquisition and at retention after 11 days, the time needed to obtain the images and the quality of images obtained. We divided the knowledge test into three topics:

[1]Center for Education Development and Research in Health Professions, research group LEARN, University of Groningen and University Medical Center Groningen
[2]Department of Medical Psychology and Psychiatry, School of Medical Sciences, University of Campinas
[3]Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen
[4]Research group LEARN, University Medical Center Groningen
[5]Department of Cardiology, University of Groningen and University Medical Center Groningen
[6]Department of Cardiology, Catharina Hospital, Eindhoven, the Netherlands

Corresponding Author:
Dario Cecilio-Fernandes, Cidade Universitária, Rua Tessália Vieira de Camargo, 126, Campinas SP, CEP: 13083-887, Brazil.
Email: dario.fernandes@gmail.com

names of the images, manipulation of the probe, and anatomy of the heart. At acquisition, immediately after training, EF group participants were faster at obtaining images than participants in the two other groups. On the retention test, there were no group differences for speed of obtaining images, but the EF + HS group scored significantly higher than the other two groups on image quality. Thus, expert feedback best assisted initial acquisition and combined augmented and expert feedback best assisted retention of this complex medical task. Expert assistance reduced learners' cognitive load during initial practice, helping learners integrate declarative knowledge with physical skills.

Mastering complex medical skills requires the learner to acquire both necessary medical and procedural knowledge and the requisite perceptual–motor skills and then to integrate these into a coordinated, coherent task performance. Both instruction and feedback are essential for learning such a complex skill. Instruction from an expert helps the learner integrate the necessary knowledge and skills, while feedback helps learners by guiding and correcting their actions. Although much is already known about feedback processes in learning relatively simple perceptual–motor skills, less is known about feedback in learning more complex skills that require vast background and procedural knowledge. For example, little is known of what type of feedback best supports the integration of knowledge and motor skills.

Medical training has long included both classroom instruction and clinical skill training, by observation or imitation (Torre et al., 2006). However, the clinical environment is saturated with both relevant and irrelevant information, challenging the medical student to discern relevant information. The teacher's role is to guide students to recognize relevant information in acquiring the necessary motor skills, in formulating a diagnosis and in intervening. In this review of relevant prior literature, we will first overview what is known about the feedback effectiveness in motor skill acquisition and then apply this knowledge to learning a complex medical skill selected for study here, namely, obtaining transthoracic echocardiogram (TTE) images. TTE is a noninvasive technique with an ultrasound probe placed on the patient's chest to create moving images of the heart that are displayed on a video monitor.

Feedback in Skill Acquisition

Mastering a skill involves overlearning required actions to the point of automaticity. Mastery requires more than practice, as learners must also know and

recall task instructions in order to determine whether they are practicing correctly. However, recalling complex instructions can be difficult and may lead to forgetfulness errors (Taatgen et al., 2008). As learners progress to the point of automizing the skill, the cognitive demand and the number of errors are diminished (Anderson, 1982). Thus, skill acquisition is supported by both declarative and procedural knowledge (Anderson, 1982; Anderson et al., 2004). Declarative knowledge refers to knowledge of facts or events (*knowing what*) and can be consciously inspected. Procedural knowledge refers to *knowing how* and cannot be consciously inspected. Declarative knowledge decays overtime, meaning that this knowledge can be forgotten when it is not used often enough. Procedural knowledge is not subject to decay (Anderson, 1982; Anderson & Fincham, 1994). Learning a new skill starts with receiving instructions in the declarative format (Anderson & Fincham, 1994). Then, while practicing a skill, declarative knowledge is slowly transformed into procedural knowledge. Through this *proceduralization*, the skill can be said to have become automatized (Anderson, 1982; Taatgen & Lee, 2003).

Feedback during skill acquisition has been found useful for several reasons (Salmoni et al., 1984). First, feedback prevents proceduralizing incorrect knowledge by providing learners guidance on *how* to perform the actions of a task, so that learners can adapt their own actions accordingly (Salmoni et al., 1984). Second, feedback diminishes learners' cognitive demand associated with retrieving declarative knowledge by providing learners' with correct knowledge during practice, especially in complex skills (Wulf & Shea, 2002). Feedback also guides learners' attention toward relevant (vs. irrelevant) task detail (Wulf & Shea, 2002), and it keeps learners motivated by rewarding correct actions and/or punishing incorrect actions (Salmoni et al., 1984). However, for complex skills such as TTE that require both declarative (biomedical) and procedural knowledge (skills), there are still some unclear aspects of feedback, most notably those feedback aspects that are associated with how to best support integrating declarative knowledge with motor skills so as to promote skill retention.

The optimal amount of feedback needed is related to the complexity of the skill (Wulf & Shea, 2002). Since complex skills require more declarative knowledge than simple skills, tasks that heavily rely on declarative knowledge are more cognitively demanding for the learner. Retrieving declarative knowledge demands time and energy from the cognitive system. The more complex the skill is, the more time, practice, and feedback will be required to master it. However, more feedback is not always better; studies have shown that too much feedback *during* training may actually hamper later performance, after a period of nonuse (Hatala et al., 2014; Schmidt, 1991; Wulf & Shea, 2004). Students who received constant feedback during training (concurrent feedback) performed worse on a retention test than both students who received feedback only on the outcome (terminal feedback; Ranganathan & Newell, 2009; Sulzenbruck & Heuer, 2011) and students who received gradually diminishing

concurrent feedback (Kovacs & Shea, 2011). The risk of too much feedback is that feedback may become incorporated into memories of the movements so that when the feedback is removed, performance suffers (Schmidt, 1991). The explanation for this feedback dependency may lie in how declarative and procedural knowledge interact. When initially acquiring a skill, learners' procedural knowledge is general; as learners progress, this knowledge becomes highly specified (Taatgen et al., 2008). When learners receive constant concurrent feedback and their production memories include the feedback itself, their procedural knowledge is less specified so that when the feedback is later removed, production memories are harder to retrieve. To avoid learners becoming dependent on feedback, some researchers have proposed that the amount of feedback should decrease as the learners' skill mastery level increase (Guadagnoli & Lee, 2004; Wulf & Shea, 2002). This idea is consistent with the notion that learning complex skills requires understanding which task elements are most relevant to skill mastery.

TTE Feedback

As noted earlier, we investigated learning a complex medical skill that requires integrating knowledge (declarative and procedural) with motor skill proficiency—TTE. Advances in TTE simulators permit the use of this medical skill in research of this kind. First, TTE learners must acquire the necessary biomedical knowledge regarding heart anatomy, physiology, and pathology. Then they must acquire probe manipulation perceptual–motor skills while understanding the relationship between probe movements and monitor images, and they must acquire the perceptual–cognitive skills for mentally transforming two-dimensional TTE images into a three-dimensional representation of the heart.

Perceptual–Motor Skills Feedback

Perceptual–motor skills refer to the ability to integrate motor actions with observed changes in the environment. For a TTE, this refers to being able to move the probe in such a way that the required image is shown on the monitor. Many studies have examined the effect of attentional focus in motor skills, generally finding that focusing feedback on external information rather than on the actual movement is most effective. For example, in dart throwing, instructing learners to focus on the target (external) rather than on arm movement (internal) increased throwing accuracy (Marchant et al., 2007). This external attention advantage has been found in learning many other motor skills including golf (Wulf & Su, 2007), basketball (Zachry et al., 2005), tennis serves and shooting lofted passes at a target (Wulf et al., 2002).

Prior research has also found that a visual external feedback display further benefits motor learning as this so-called *augmented feedback* permits the learner

to view real-time feedback of personal movement (Kovacs & Shea, 2011; Snodgrass et al., 2010). For example, Kovacs et al., (2009) demonstrated that augmented visual feedback decreased the number of errors and time necessary to acquire a bimanual coordination task, although there was a drop in performance later, when the feedback was removed. In a later study, Kovacs and Shea (2011) found that gradually decreasing the frequency of augmented feedback overtime prevented this deterioration in performance. Snodgrass et al. (2010) demonstrated that augmented visual feedback enabled students to manually apply forces during a cervical spine mobilization that were similar to the forces applied by experts. In both the acquisition and retention phases, students who received augmented visual feedback applied forces that were similar to those of experts, while students in the control group did not. Sigrist et al., (2013) hypothesized that the beneficial effect of concurrent feedback during skill acquisition is because this feedback reduces the learner's cognitive demands and thereby avoids detrimental cognitive overload, especially in complex tasks.

Perceptual–Cognitive Skills Feedback

Perceptual–cognitive skills are skills that integrate relevant information from the environment with the necessary knowledge to decide how to act next (Mann et al., 2007). For the TTE, this refers to understanding the dynamic image on the screen, relating that to declarative knowledge about heart anatomy and pathology, and then deciding whether there is yet enough evidence to formulate a diagnosis. In performing a TTE, spatial cognition is also important. Spatial skills refer to the ability to mentally understand and manipulate an object in three-dimensional space (Uttal et al., 2013). For the TTE, spatial cognition involves understanding how a two-dimensional image of the heart on the monitor relates to the three-dimensional structure of the heart; this requires mentally transforming and manipulating a two-dimensional image into a three-dimensional representation. Furthermore, medical doctors need to understand the associated three-dimensional relationship between the heart and other nearby body organs such as the lungs. Because the TTE provides a dynamic image of a beating heart, the learner should also understand how the image changes overtime.

Current Study

Despite the vast literature on feedback in learning either perceptual–cognitive or perceptual–motor skills, it remains unclear how feedback might best guide learning a complex skill that requires *integrating* declarative biomedical *knowledge* with perceptual–cognitive and perceptual–motor *skills*. For example, studies on motor skill acquisition suggest that reducing the frequency of augmented visual feedback during acquisition increases learners' performance in acquiring and

retaining perceptual–motor tasks, but, unlike TTE, these tasks typically require little declarative knowledge. Thus, for complex skills, expert rather than simple augmented feedback may be necessary to help learners integrate declarative knowledge with perceptual–cognitive and perceptual–motor skills. Also, many studies have focused on acquisition rather than the retention of complex skills, perhaps leading to oversimplified conclusions about beneficial effects of multiple sources of feedback that tend to reduce students' cognitive load during acquisition (Hatala et al., 2014; Wulf, 2007; Wulf & Shea, 2002). Meanwhile, there is a sparse prior research on the retention effect of multiple feedback sources applied to learning either complex skills (Sigrist et al., 2013) or medical skills (Hatala et al., 2014).

We sought to compare the effects of (a) augmented (visual) feedback, (b) expert feedback, and (c) the combination of both of these feedback sources on the acquisition and retention of a complex medical skill that required integration of declarative knowledge with perceptual–cognitive and perceptual–motor skills. Since expert feedback supports the integration of declarative knowledge with perceptual–cognitive and perceptual–motor skills, we hypothesized that, during skill acquisition, learners who are guided by experts would obtain the TTE images faster than those who were only guided by augmented visual feedback. We expected augmented visual feedback to help learners automate perceptual motor learning by better understanding the accuracy of their movements. Because the literature suggests that adding augmented visual feedback may reduce cognitive overload during the acquisition of the TTE, we speculated that participants who received feedback from multiple sources would obtain the images with better quality than those who were only guided by expert feedback.

Method

Participants

We recruited undergraduate medical students by an email in which they received instructions about the planning, the timing, and the advantages/disadvantages of voluntarily participating in this experiment. They were all informed that participation was voluntary and they could opt out at any moment without any consequences. All data were processed confidentially. All participants signed an informed consent form, and the experiment was performed in accordance with Dutch Law regulations. At the end of the experiment, participants received a gift voucher worth €1,000.

Participants were 39 medical students in their second, third, fourth, and sixth years of medical school (i.e., in the preclinical 1–3 years and in the clinical 4–6 year phase of undergraduate medical training). Participants were randomized to one of three equal sized groups. We followed the criteria for block

randomization by creating 15 blocks with 3 participants per group, 1 participant for each condition in each block (having created additional blocks in case participants would drop off the experiment). Using this randomization process, we assured that the number of participants were equal in all the groups and that participants had an equal probability of being allocated to one of the three conditions. However, six participants, three participants from the Expert (EF) group and three participants from Expert plus Help Screen (EF + HS) group, crossed over to the Augmented visual (HS) group, when experts for the EF and EF + HS groups were unexpectedly unavailable due to patient care needs. This decision to change these participants' group assignment occurred before the practical part of the experiment was initiated. Therefore, these participants did not acquire different knowledge by having initially been part of one group and then switching to another. Of the 39 participants, three did not attend the follow-up session and were excluded from the analysis. Of the 36 participants, 31 were females (aged 19–26 years; $M = 22.36$, $SD = 2.02$) and five were males (aged 19–26 years; $M = 22$, $SD = 2.65$). Of these remaining 36 participants, 10 were in the EF group, 16 in HS, and 10 in the EF + HS group.

The EF group received feedback from an expert in TTE with the expectation that expert feedback would help participants focus on relevant aspects of the task and learn to integrate their declarative biomedical knowledge with their perceptual–cognitive and perceptual–motor skills. The HS group received augmented visual feedback with the expectation that visual augmented feedback through a simulator would help participants acquire the necessary perceptual–motor and perceptual–cognitive skills. Visual feedback provides information regarding the accuracy of the probe movements, helping learners automate movements with feedback regarding movement accuracy. Visual feedback also reminds participants of what the target images should look like, helping them recall the correct declarative knowledge. However, visual augmented feedback was not expected to help participants integrate their cognitive and perceptual–motor skills. The EF + HS group received feedback from the expert and the visual augmented feedback from the simulator. In this group, the visual feedback provided by the help screen offers information regarding the accuracy of the participants' movements. The expert feedback guides participants' attention to relevant aspects of the tasks. It should also help participants to integrate their knowledge with the cognitive skills and perceptual–motor skills.

TTE Simulator

In this study, we used the CAE VIMEDIX^{TM} ultrasound simulator. The simulator consists of a life-size mannequin torso with soft skin, and accurate and palpable anatomical parts, a TTE transducer, and a computer with monitor.

Procedure and Design

Data gathering for measures in this experiment took place immediately after a training session and at a retention test after 11 days. The training session was 2-3 hours long, depending on how fast the participant learned to make the images. At the start of the experiment, participants were given written information on the acoustic windows and views of the heart that they would learn. Then, participants watched a video that explained the anatomy of the heart and showed how to manipulate the probe to obtain the images. Next, we provided a second set of written material explaining how to place and manipulate the probe. They watched the same video one more time, but this time they could stop it, and fast-forward or go back until they felt confident enough to take the knowledge test. They could also study all the material provided while watching the video. To ensure that participants had acquired all the necessary knowledge to perform the TTE, they were required to take a knowledge test on the computer and answer all the questions correctly. If they made a mistake, they received feedback on their answers and had to repeat the test until they were able to answer all questions correctly on two consecutive test administrations.

The EF group practiced on the simulator with the assistance of an expert cardiologist. The HS group practiced on the simulator with the assistance of the help screens available on the simulator. The EF + HS group practiced on the simulator with the assistance of both an expert cardiologist and the help screens. All the participants were instructed to obtain the best images possible without any time restriction. After practicing on the simulator, participants took a practical test, in which they had to obtain all the images correctly twice in a row. If an image was not correct, participants could practice on that image and repeat the test for that image until they were able to obtain the image correctly twice in a row. During the test, all the simulator's help screens were disabled. At the end of the session, participants filled out a short questionnaire, to evaluate the TTE training that they had received. After 11 days, participants were invited back for the final test. They performed the same knowledge and practical tests as in the training session and, at the end, filled out questionnaires.

Two experienced cardiologists with expertise on TTE participated in this study. They provided feedback similar to real-life training for the EF and the EF + HS groups. When the help screens were on, the experts provided feedback similar to real-life training, but they also took advantage of the help screens. For example, when the probe was in the wrong position, the instructor would point to the help screens and give feedback regarding the position of the probe.

Outcome Variables

Knowledge Test. The knowledge test consisted of 20 multiple-choice questions and one match-pattern question, which were divided into three blocks: names of the images, manipulation of the probe, and anatomy of the heart. Participants took the

test on a computer. All questions were automatically scored as correct or incorrect. At the end of each question block, participants received feedback for each question.

Quality of the Images. Participants were asked to obtain the following five images on the simulator: Parasternal long- and short-axis view, Apical four-chamber view, and Subcostal four-chamber, and Inferior Caval Vein view. The quality of each image was graded independently by two experienced cardiologists who were blinded to the groups and participants. They graded the images on a scale ranging from 0 to 4 points based on the following criteria:

0. Chambers are not displayed/wrong image.
1. One or more chambers are not displayed/not fully displayed.
2. All chambers are displayed, however, one chamber is incomplete.
3. All chambers are displayed, however, the angle/cross-cut is off.
4. All chambers are displayed with the right angle/cut.

Performance Measures. In addition to the quality of the images, the time it took to complete each image was measured during the practical test, using a chronometer. The number of attempts for each block of the knowledge test and for each image was measured as well.

Questionnaire. The questionnaire was answered immediately after the final test session. A questionnaire was used to collect extra information about the participants and the experiment. The questionnaire (in Online Appendix) was divided into three parts. The first part contained questions regarding the participants' demographics and how many times they had watched or performed the TTE. The second part contained nine questions on the instructions and the experienced cognitive load of the training, which had to be scored on a Likert-type scale ranging from 1 to 5. The last part contained open questions in which participants could elaborate more on the quality of the training and give suggestions for improving the training session.

Data Analysis

All analyses were conducted using SPSS 21.0, and the alpha level for determining statistical significance was set to .05. To investigate the influence of the three types of feedback on participants' skill acquisition, we conducted the Kruskal–Wallis *H* test using *group* as independent variable and *number of participants' attempts for the knowledge test, number of participants' attempts for the practical test, and time that participants took to complete the practical test after the training session* as dependent variables. To determine which groups differed on the Kruskal–Wallis test, we performed post hoc multiple-comparison analyses.

To investigate the influence of the three types of feedback on participants' skill retention, we conducted an analysis of covariance. *Groups* were added as a fixed factor. The scores of the knowledge test, gender, medical year, and how many times participants had performed and watched a TTE were added as covariates in the model. To determine which groups differed on the corrected means of the model, we performed post hoc multiple-comparison tests. For this analysis, we used an alpha level of .05.

To investigate the relationship between the scores in the different blocks of the knowledge test and the TTE performance, we conducted a Pearson correlation analysis between the three block tests and scores on the quality of the images. To investigate participants' perceptions of the instructions and the experienced cognitive load of the training, we calculated the median of participants' answers, since those questions were in a Likert-type scale format.

Results

Baseline Measurement

Two baseline variables could have affected the experiment, namely, whether participants had performed a cardiac ultrasound before and whether they had watched others make a cardiac ultrasound. Since only two participants had performed a TTE previously, an analysis of variance to determine any difference between groups was not possible. The second variable concerned how many times participants had watched someone else perform a TTE. Since the second variable was nonnormally distributed, we conducted the Kruskal–Wallis *H* test as an analysis of variance for nonparametric data. We did not find a significant difference between the three experimental groups, $\chi^2(2) = 0.144$, $p = .930$, meaning that all groups had comparable previous experience with echocardiography.

Training Session

During the training session, participants practiced until they reached proficiency. Thus, all participants achieved the highest scores on the knowledge and practical tests. Figure 1 displays the various number of test administrations required for participants to reach this level. The number of participant attempts at the knowledge test, $\chi^2(2) = 8.224$, $p = .894$, and at the practical test, $\chi^2(2) = 2.457$, $p = .293$, were not significantly different between the groups. However, the time participants took to complete the practical test, $\chi^2(2) = 7.746$, $p = .021$, $\eta^2 = 0.221$, was significantly different between the groups, and post hoc multiple-comparison analyses demonstrated that the EF group was significantly faster than the HS group but not significantly faster than HS + EF group (see Figure 1). The effect size for the significant difference between EF group and HS group was 0.350.
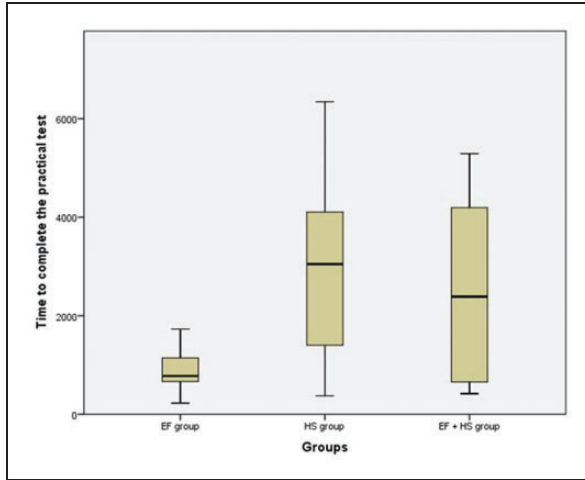
Figure 1. Box Plot Comparison of Time (in seconds) to Complete the Practical Test Across Groups EF, HS, and EF+HS, Immediately After Training. EF = Expert; HS = Augmented visual; EF+HS = Expert plus Help Screen.
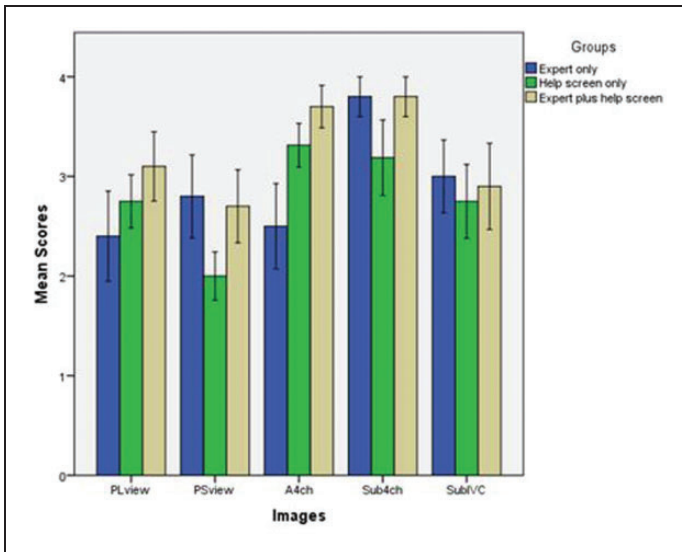

Figure 2. Means and Standard Error Scores for TTE Image Quality Across Groups EF, HS, and EF+HS, After 11 Days. PLview = Parasternal Long Axis View; PSview = Parasternal Short Axis View; A4ch = Apical 4 chamber view; Sub4ch = Subcoastal 4 chamber view; SubIVC = Subcoastal Inferior Caval Vein view.
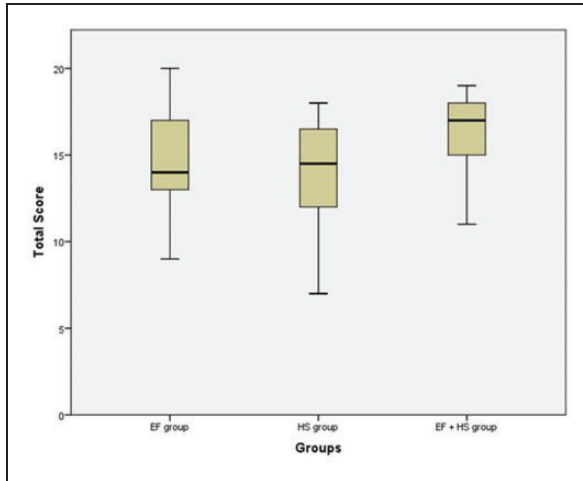
Figure 3. Box Plot Comparison of Participant Total Imagery Quality Scores on the Practical Retention Test Across Groups EF, HS, and EF+HS. Scores ranging from 0 to 20 points. EF = Expert; HS = Augmented visual; EF+HS = Expert plus Help Screen.

Retention Session

The scores for the quality of the images at the retention session are shown in Figure 2. The analysis of covariance for image quality demonstrated a significant difference between groups, $F(2, 32) = 3.374$, $p = .049$, $\eta^2 = 0.194$. Post hoc multiple-comparison analyses demonstrated that the EF + HS group scored significantly higher on the practical test than the HS group ($M_{\text{EF}+\text{HS}} = 16.20$; $M_{\text{HS}} = 14.00$). The EF group was not significantly different from the other two groups (see Figure 3). There was no group difference in the time participants took to complete the final practical test, $\chi^2(2) = 1.045$, $p = .593$.

A Pearson correlation analysis demonstrated that image quality scores correlated positively with knowledge test scores involving names of images, manipulation of the probe, and anatomy of the heart (see Table 1). Only the scores for knowledge related to manipulation of the probe significantly correlated with the practical test scores ($r(34) = .45$, $p = .005$). The magnitude of this correlation is considered moderate (Taylor, 1990).

Questionnaire

Overall, participants found the instructions used in the training session very clear (median score 4). They perceived the theoretical part of the training as easier than the practical part (median scores 3 and 4, respectively) and indicated

Table 1. Correlations Between Rated Quality of the TTE Images and Scores on Blocks in the Knowledge Test.

		Blocks in knowledge test		
		Name of the image	Manipulation of the probe	Structure of the heart
Quality of images	R	.21	.45	.31
	P	.222	.005	.660
	Degrees of freedom	34	34	34

Note. r = Pearson correlation; p = significance.

that they put a lot of effort into both parts (median score 4, for both parts). The video and the knowledge test were considered the most helpful materials; participants stated that both forced them to remember and apply what they had studied. Nevertheless, participants found it unnecessary to be tested twice in a row to demonstrate knowledge achievement. Some participants found the number of help screens available to be overwhelming. A few participants suggested to include more information/practice on how to manipulate the probe and the effect of manipulating the probe on the image. Furthermore, two participants suggested spreading the practice over a few days instead of practicing in one session.

Discussion

In this study, we sought to determine the optimal way of guiding learners' acquisition and retention of a complex medical skill, transthoracic cardiac ultrasound (TTE), with a focus on feedback sources. We manipulated the source of feedback participants received during their practice on the simulator in three feedback conditions: (a) expert provided (EF), (b) help screen provided (HS), or (c) expert and help screen provided (EF + HS). At the acquisition test conducted immediately after the training session, all the participants were required to acquire all images correctly twice consecutively, meaning that image quality across the conditions could not be compared at that time. However, EF participants were faster at obtaining TTE images than HS participants (but not faster than HS + EF participants). This finding partially contradicts our hypothesis and prior literature. In Hatala et al. (2014)'s meta-analysis of feedback effectiveness for simulation-based medical education, there was a moderate advantage of expert feedback to simulator feedback and an further performance improvement from an extra source of feedback in acquisition learning. Our contradictory finding may be due in part to variations in the type of medical

skill tasks investigated, as primary studies focused primarily on procedural (perceptual–motor) skills, while we chose TTE for its required integration of declarative knowledge and perceptual–cognitive and perceptual–motor skills.

At the retention test after 11 days, we found no group differences in time needed to acquire the images. In terms of image quality, participants from all groups had already shown a decay in their knowledge and a decrease in their TTE skill performance levels, even though all participants had perfect scores in both areas immediately after the training. These findings are in accordance with prior literature in that even though participants trained until proficiency, they demonstrated a decay after a period of nonuse (Barsuk et al., 2010; Moazed et al., 2013; Wik et al., 2002). The explanation for this decline after only 11 days may lie with problems in acquired declarative and procedural knowledge. Although learners achieved an apparently high proficiency level, perceptual–motor and perceptual–cognitive learning may not have yet automatized. The part of this skill that was declarative knowledge or the production rules may not have been sufficiently learned to resist decay. Complex skills that require more declarative knowledge should benefit more from spacing training sessions (as our participants subsequently requested) than skills that require less declarative knowledge (Cecilio-Fernandes et al., 2018). An important implication of this finding is that teaching and learning strategies are different for declarative and procedural knowledge. When skill knowledge is in a declarative format, spacing the training sessions should be preferred, while when the knowledge is in a procedural format, massed practice should be preferred (Kim et al., 2013).

Of all groups, the EF + HS group acquired the best quality images. We think that this may be caused by a reduction of cognitive load during the practice session, as suggested by Hatala et al. (2014). Experts likely guided learners in both skill acquisition and in using and interpreting the help screen, minimizing the metacognitive load associated with using help screens by themselves. Although we hypothesized that augmented visual feedback would support learners to automate their perceptual motor learning, participants in the HS group, had to rely exclusively on previously given instructions; they were observed to explicitly recall all instructions. We believe this required cognitive load may have hampered their automation of the motor actions, because these participants needed to actively remember the relationship between the help screen and the image. This resulted in a poor retention later when compared with the EF + HS group. Furthermore, participants in the HS group may have focused their attention on the motor actions but not on the relationship between the anatomy and the images, which is one of the most difficult aspects of the TTE (Weidenbach et al., 2005). Studies on perceptual–motor skills have shown that learners who focus on their own internal movements may perform worse than those who focus on the external effects of their movements (for a review, see Wulf et al., 2010). At least for this complex medical skill, our finding suggests that visual

augmented feedback may support participants' skill retention only when it is associated with another type of feedback.

The optimal feedback source depends partly on skill complexity (Wulf & Shea, 2002). Although suturing and performing a TTE, for example, both require biomedical knowledge and perceptual–motor skill, the relative amounts of biomedical knowledge and perceptual–motor skill necessary for each task is different. Whereas simple motor skills benefit from feedback at the end of the session (terminal feedback), complex motor skills benefit from feedback during the training session (augmented feedback; for a review see Sigrist et al., 2013 and see Wulf & Shea, 2002). Thus, dividing skill components into declarative and procedural knowledge may help design optimal feedback for skill training (Kim et al., 2013) by matching best teaching strategies to the type of knowledge a skill demands. In summary, it appears that although expert feedback may reduce the training time for acquiring a complex medical skill, expert feedback alone may not be optimal for skill retention. Combined augmented visual and expert feedback seems to best promote learners' skill retention for complex medical skills, though whether this finding also holds after more practice sessions or over longer retention intervals, remains to be seen.

We studied TTE in this experiment because it is a complex medical skill requiring learners to integrate declarative (biomedical) knowledge, perceptual–cognitive, and perceptual–motor skills. This integration requirement for TTE was supported by our finding of a positive correlation between participant scores on the declarative test on probe manipulation and the quality of their produced images, suggesting that this declarative knowledge may have assisted probe manipulation for acquiring better quality images. Practically, training TTE skills has often emphasized acquisition rather than retention. But, since the acquisition and retention of TTE skills appear to be affected by different feedback sources, effective TTE training for persistent skill proficiency should combine expert and augmented visual feedback.

A limitation of this study lies in our failure to control for practice time on the simulator. However, we standardizing simulator time for competency-based proficiency also enabled a more reliable comparison of skill decay overtime associated with different training conditions. We reconfirmed the established fact that individual learners vary considerably in the time needed to acquire a skill through participant differences in time needed to master the TTE skills, with training sessions that ranged from 60 to 120 minutes. Another limitation lies in the short time between our training and retention session, since in clinical practice, medical skills must be retained over many months. However, 11 days proved to be enough time for participants to show a drop in their knowledge and skills, possibly because our participants were not able to automatize their skills in a single training session. Our data suggest that repeat training sessions are necessary (even within 11 days for TTE) for participants to acquire and retain their skills overtime. Our small sample size and our failure to conduct

an *a priori* power analysis may have led us to underpower this study statistically, perhaps contributing to some nonsignificant findings, though we detected some significant findings. In addition, we did not calculate interrater reliability for scoring image quality, relying heavily on the fact that raters were experts in cardiology and TTE.

Conclusion

In accordance with prior suggestions from the literature, our results demonstrate a superior combined effect of augmented feedback with expert guidance and expert feedback during skill acquisition on retention of a complex medical task. Experts seemed to be especially useful for reducing the cognitive load during practice and thereby helping learners integrate declarative knowledge with perceptual–motor skills.

ORCID iDs

Dario Cecilio-Fernandes https://orcid.org/0000-0002-8746-1680
Fokie Cnossen https://orcid.org/0000-0001-7435-4889

Supplemental Material

Supplemental material for this article is available online.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–406.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060.

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of Experimental Psychology Learning, Memory, and Cognition*, 20(6), 1322–1340.

Barsuk, J. H., Cohen, E. R., McGaghie, W. C., & Wayne, D. B. (2010). Long-term retention of central venous catheter insertion skills after simulation-based mastery learning. *Academic Medicine*, 85(10), S9–S12.

Cecilio-Fernandes, D., Cnossen, F., Jaarsma, A. D. C., & Tio, R. A. (2018). Avoiding surgical skill decay: A systematic review on the spacing of training sessions. *Journal of Surgical Education*, 75(2), 471–480.

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.

Hatala, R., Cook, D. A., Zendejas, B., Hamstra, S. J., & Brydges, R. (2014). Feedback for simulation-based procedural skills training: A meta-analysis and critical narrative synthesis. *Advances in Health Sciences Education*, 19(2), 251–272.

Kim, J. W., Ritter, F. E., & Koubek, R. J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, 14(1), 22–37.

Kovacs, A. J., Buchanan, J. J., & Shea, C. H. (2009). Bimanual 1:1 with 90° continuous phase: Difficult or easy? *Experimental Brain Research*, 193, 129−136.

Kovacs, A. J., & Shea, C. H. (2011). The learning of 90 continuous relative phase with and without lissajous feedback: External and internally generated bimanual coordination. *Acta Psychologica*, 136(3), 311–320.

Mann, D. T., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology*, 29(4), 457–478.

Marchant, D. C., Clough, P. J., & Crawshaw, M. (2007). The effects of attentional focusing strategies on novice dart throwing performance and their task experiences. *International Journal of Sport and Exercise Psychology*, 5(3), 291–303.

Moazed, F., Cohen, E. R., Furiasse, N., Singer, B., Corbridge, T. C., McGaghie, W. C., & Wayne, D. B. (2013). Retention of critical care skills after simulation-based mastery learning. *Journal of Graduate Medical Education*, 5(3), 458–463.

Ranganathan, R., & Newell, K. M. (2009). Influence of augmented feedback on coordination strategies. *Journal of Motor Behavior*, 41(4), 317–330.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, 95(3), 355–386.

Schmidt, R. A. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. In J. Requin & G. E. Stelmach (Eds.), *Tutorials in motor neuroscience* (pp. 59–75). Springer.

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin & Review*, 20(1), 21–53.

Snodgrass, S. J., Rivett, D. A., Robertson, V. J., & Stojanovski, E. (2010). Real-time feedback improves accuracy of manually applied forces during cervical spine mobilisation. *Manual Therapy*, 15(1), 19–25.

Sulzenbruck, S., & Heuer, H. (2011). Type of visual feedback during practice influences the precision of the acquired internal model of a complex visuo-motor transformation. *Ergonomics*, *54*(1), 34–46.

Taatgen, N. A., Huss, D., Dickison, D., & Anderson, J. R. (2008). The acquisition of robust and flexible cognitive skills. *Journal of Experimental Psychology: General*, *137*(3), 548–565.

Taatgen, N. A., & Lee, F. J. (2003). Production compilation: A simple mechanism to model complex skill acquisition. *Human Factors*, *45*(1), 61–76.

Taylor, R. (1990). Interpretation of the correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography*, *6*(1), 35–39.

Torre, D. M., Daley, B. J., Sebastian, J. L., & Elnicki, D. M. (2006). Overview of current learning theories for medical educators. *The American Journal of Medicine*, *119*(10), 903–907.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402.

Weidenbach, M., Wild, F., Scheer, K., Muth, G., Kreutter, S., Grunst, G., Berlage, T., & Schneider, P. (2005). Computer-based training in two-dimensional echocardiography using an echocardiography simulator. *Journal of the American Society of Echocardiography*, *18*(4), 362–366.

Wik, L., Myklebust, H., Auestad, B. H., & Steen, P. A. (2002). Retention of basic life support skills 6 months after training with an automated voice advisory manikin system without instructor involvement. *Resuscitation*, *52*(3), 273–279.

Wulf, G. (2007). Attentional focus and motor learning: A review of 10 years of research. *E-Journal Bewegung Und Training*, *1*(2–3), 1–11.

Wulf, G., McConnel, N., Gärtner, M., & Schwarz, A. (2002). Enhancing the learning of sport skills through external-focus feedback. *Journal of Motor Behavior*, *34*(2), 171–182.

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, *9*(2), 185–211.

Wulf, G., & Shea, C. H. (2004). Understanding the role of augmented feedback. In A. M. Williams & N. J. Hodges (Eds.), *Skill acquisition in sport: Research, theory and practice* (pp. 121–144). Routledge.

Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: A review of influential factors. *Medical Education*, *44*(1), 75–84.

Wulf, G., & Su, J. (2007). An external focus of attention enhances golf shot accuracy in beginners and experts. *Research Quarterly for Exercise and Sport*, *78*(4), 384–389.

Zachry, T., Wulf, G., Mercer, J., & Bezodis, N. (2005). Increased movement accuracy and reduced EMG activity as the result of adopting an external focus of attention. *Brain Research Bulletin*, *67*(4), 304–309.

Author Biographies

Dario Cecilio-Fernandes is a researcher in the Department of Medical Psychology and Psychiatry, School of Medical Sciences, University of Campinas and has research interest on assessment as well as skill acquisition and retention. He has also working with faculty development in the University of Campinas.

Fokie Cnossen is a senior lecturer in Medical Cognition, and director of Education at the department of Artificial Intelligence of the University of Groningen, the Netherlands. She also coaches national-level athletes.

Jenifer Coster is a cardiologist in UMCG, member of acute cardiology team (responsible for cardiac care unit and emergency department regarding cardiology). Helped to develop e-learning cardiac ultrasound for the medical students of Rijksuniversiteit Groningen. Team member of supervisors of the cardiac ultrasound laboratory UMCG.

AD (Debbie) C. Jaarsma, DVM, PhD is professor of health professions education and program director of LEARN and director of the Center for Education Development and Research in Health Professions, University Medical Center Groningen, Groningen, the Netherlands.

René A. Tio is cardiologist with a broad experience in Medical Education amongst others he is involved in how testing drives learning and how students learn clinical reasoning; he is former chair of the Dutch progress test consortium and former chair of the Examination board of the University Medical Center Groningen.