AN ABSTRACT OF THE THESIS OF

Danielle M. Meritet for the degree of Master of Science in Comparative Health Sciences presented on June 6, 2019.

Title: Effects of Error-Management Training on Veterinary Student Performance in Technical Skills Training

Abstract approved: ____
Duncan S. Russell

Conventional training methods in most medical fields seek to avoid errors, framing them as negative consequences instead of as learning opportunities. Error-management training (EMT) contrasts with conventional methods by explicitly encouraging trainees to make errors through low-stakes training modules in an effort to alter perspectives on errors, utilizing them as learning opportunities, and to improve trainee performance in error-prone tasks. Simultaneously, EMT can promote the use of non-technical skills – namely metacognition and emotion control – that can provide trainees with lifelong, stress-reducing tactics that can be used across disciplines. The following describes two prospective, double-blinded studies to investigate the effects of EMT, compared to traditional error-avoidant training (EAT), on veterinary students learning to perform: 1) surgical knot tying and 2) blood smear preparation and analysis. For both studies, the performances between EMT and EAT participants were not significant when completing familiar tasks, suggesting EMT is a comparable training method. When assessing novel scenarios, participants trained by EAT showed a significant decline in long-term performance for several outcomes compared to EMT trainees. Survey data from the blood smear study found that participants in both groups felt errors were a useful part of their training and that they were able to reflect on their mistakes. These findings suggest that EMT could be an effective teaching strategy that fosters both technical and non-technical skills for veterinary students.

Effects of Error-Management Training on Veterinary Student Performance in Technical Skills Training

by
Danielle M. Meritet

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented June 6, 2019
Commencement June 2019

Master of Science thesis of Danielle M. Meritet presented on June 6, 2019.

APPROVED:

Major Professor, representing Comparative Health Sciences

Dean of the College of Veterinary Medicine

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes the release of my thesis to any reader upon request.

Danielle M. Meritet, Author

ACKNOWLEDGEMENTS

The author expresses sincere appreciation for Drs. Duncan Russell (major professor), Katy Townsend (committee member), Elena Gorman (committee member), Patrick Chappell (committee member), Robin Pappas (committee member), Erica Woekel (graduate council representative), the Department of Biomedical Sciences at Oregon State University, and the participants from the Carlson College of Veterinary Medicine Class of 2020 and Class of 2021.

CONTRIBUTION OF AUTHORS

Dr. Duncan Russell aided in the experimental design and execution as well as revisions for both manuscripts related to this thesis. Dr. Katy Townsend contributed her surgical expertise in the surgical knot tying study and assisted in data acquisition and evaluation for said study. Dr. Elena Gorman offered her clinical expertise in the blood smear study. Drs. Katy Townsend, Elena Gorman and Patrick Chappell facilitated in training sessions for both studies and provided revisions for both manuscripts. Dr. Laura Kelly guided proper statistical analyses and interpretation for both manuscripts.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

Effects of Error-Management Training on Veterinary Student Performance in Technical Skills Training

Introduction

Errors are an inevitable part of the human condition. With that in mind, it would be unreasonable to expect someone to completely prevent errors and, yet, many professional training strategies emphasize error avoidance. In professional settings, the adverse perception of errors stems from their cost with regards to finances, time, safety and individual intellectual resources (Edmondson 1996; Sitkin 1996; Van Dyck 2010). In this perspective, errors can become significant sources of emotional distress, particularly in vulnerable populations such as medical professionals (Mellanby 2004). Contrarily, in the educational setting, errors can serve as powerful impetuses for learning (Van Dyck 2010). When defined as deviations from the expected outcome, errors can also prompt innovation and progress. In 1874 the British economist Williams Stanley Jevons stated:

> "It would be an error to suppose that the great discoverer seizes at once upon the truth, or has any unerring method of divining it. In all probability the errors of the great mind exceed in number those of the less vigorous one." (Jevons Principles of Science: A treatise on logic and scientific method 1913 McMillan and Co, LTD London)

This notion is perpetuated in today's society, through nonfiction writing and peer-reviewed scientific publication (Buljac- Samardžić 2012; Kuhn 2012; Johnson 2011). In spite of this, many modern educational programs – including veterinary curricula – frame errors as unfavorable outcomes to be avoided as much as possible, entrenching the sentiment into their students.

Real-life scenarios often do not mimic those encountered during training, opening up opportunity for errors to occur. In the general sense, errors result from cognitive limitations or system failures, either singly or in combination (Oxtoby 2015). Cognitive limitations, if identified, can be addressed and ameliorated. System failures may be out of an individual's control, but how one reacts towards such failures is not. Regardless of cause, how one manages errors can significantly influence the behaviors and outcomes of individuals and organizations alike (Buljac-Samardžić 2012). To reduce the frequency and the negative consequences of errors, the medical profession has dedicated substantial research efforts into the investigation of the source of errors and the development of methods to mitigate risks (Lawton 2012; Taylor-Adams 2004; Van der Schaaf 2005; van Vuuren 1997). The result of this research has led to the implementation of several programs – including the London protocol, the Eindhoven PRISMA model and the Yorkshire framework – to instill systematic approaches to assess the cause of errors, and help to outline ways to recognize and manage them in a productive fashion (Oxtoby 2015). While research into the causes of errors in veterinary practice is starting to surface, there are no systematic methods outlining how to handle errors in this field at this time (Oxtoby 2015). Error-management training (EMT) could potentially help fill this gap and can be implemented early on in the veterinary curriculum.

EMT is an alternative teaching strategy that defines errors as learning opportunities, reframing errors in a positive light (Keith 2008). The defining characteristics of EMT are as follows: minimal instructor intervention to allow for self-directed learning, use of challenging (i.e. error-prone) tasks and explicit encouragement to make errors (Keith 2008). Prior studies have found that EMT leads to improved performance in novel tasks (adaptive tasks) and in long-term task completion when compared to traditional error-avoidance training (EAT) methods (Keith 2008). These same studies found that EAT participants significantly perform better in tasks that directly correlate to skills taught during training sessions (analogue tasks). One could argue, however, that improved performance in analogue tasks does not necessarily translate to success in the workplace where unfamiliar scenarios will inevitably be encountered. This is particularly true for most veterinary professionals who can expect to encounter unfamiliar scenarios on a daily basis depending on their type of practice.

In addition to task performance in technical skills, EMT can benefit trainees by promoting non-technical skills that aid in error recognition and management, namely those of metacognition and emotion control (Keith 2005). Metacognition is a self-regulatory practice that modulates understanding during task performance, requiring both self-awareness and self-directed learning. By relying on automaton-like practice, EAT methods – where individuals are instructed to follow instructions as closely as possible to arrive at expected results – leave little room for open-exploration and enhanced understanding. In contrast, EMT offers ample opportunity for individuals to meditate on their understanding, thus allowing for adjustments and adaptations when their understanding is insufficient or inaccurate to permit task completion. Some learning theories postulate that, for learning to occur, individuals need to recognize when and where their current knowledge is insufficient or inaccurate and then engage in self-directed methods to address that gap in knowledge (Ambrose 2010a; Merriam 2013b). Inherently, errors can serve as indicators of knowledge insufficiency/inaccuracy and error-management could serve as self-directed method to bridge that knowledge gap, mainly through the process of metacognition.

Another self-regulatory practice enhanced by EMT is emotion control which can help reduce anxieties and other negative emotions during task completion (Kanfer 1996). Errors often lead to negative emotions, which can cascade and become dysfunctional to the individual (Van Dyck 2010). Lack of emotion control thwarts learning and performance by diverting intellectual resources towards the self and away from the task (Kanfer 1996; Van Dyck 2010). However, it is not sufficient to simply suppress negative emotions as this also drains intellectual resources, leading to cognitive decline (Muraven 2000). To prevent cognitive decline during an emotional event, it is necessary to re-evaluate the event in order to modify emotions before they are expressed; this is done through the self-regulatory process of emotion control (Richards 2000). By framing errors in a positive light, EMT can reduce the anxiety and negative emotions that may develop when facing errors, helping participants to develop and practice emotion control (Keith 2005). In contrast, EAT participants are not as prepared to regard the negative emotions resulting from errors, potentially allowing these adverse emotional reactions to detract from their cognitive ability and, so, performance (Keith 2005). Given that veterinarians frequently face challenging tasks, potentiating the possibility of error, the practice of emotion control – as is fostered through EMT – could prove a vital resource to redirect attention to the task and away from one's own self.

The effect of EMT varies based on certain dispositional features including that of goal orientation. Given that errors result in non-completion or erroneous completion of a task, it can be said that errors prevent goal attainment. Since errors occur during the pursuit of a goal, it follows that goal orientation may influence how an individual approaches errors (Heimbeck 2003). As part of a theory proposed by Dweck et al, there are two different orientations when pursing goals: learning and performance (Dweck 1986). Learning goal orientation regards goal attainment as a means to increasing knowledge and, therefore, views errors as a chance to learn (Dweck 1988). Contrarily, performance goal orientation seeks to demonstrate competence, thereby perceiving errors as an indicator of incompetence (Dweck 1988). When assessing the efficacy of EMT, goal orientation should be considered. Heimbeck et al. found those with learning goal orientation thrive in EMT settings while those with performance goal orientation profit most under EAT conditions (Heimbeck 2003).

Both medical and non-medical training programs have successfully implemented EMT; these include ultrasonographer training in fetal measurements, driving simulation, and computer software training (Dyre 2017; Ivancic 2000; Gardner 2014). Its effects in veterinary medicine have yet to be investigated. The following material describes two prospective, double-blinded studies that evaluate the efficacy of EMT on veterinary students learning two different technical skills: 1) surgical knot tying and 2) blood smear preparation and analysis. A secondary objective of the latter study was to assess the role of the aforementioned self-regulatory skills (i.e. metacognition and emotion control) and goal orientation for veterinary students training in technical skill acquisition.

Ch. 2: Investigating the Effects of Error Management Training on Veterinary Students Learning to Tie Surgical Knots

2.1 Introduction

The aim of this study is to determine the impact of EMT in veterinary students learning to tie surgical knots. This is a skill that requires advanced fine motor skills and muscle memory. Novice surgeons find this skill challenging and are likely to generate a broad spectrum of errors, which might manifest as knot instability, inadequate hemostasis, and dehiscence. Furthermore, simulated training performed on artificial pedicles does not necessarily transfer to clinical settings, where knots may need to be placed in confined spaces, possibly with limited visualization, and unfamiliar suture materials. Difficulty may be further compounded by time restrictions (i.e. such as during bleeding or surgery in a critically ill patient). Successful training equips students to utilize existing knowledge to either modify a procedure or generate a new solution. This has been named ‘adaptive transfer’, differing from ‘analogous transfer’ where the task directly resembles training (Ivancic 2000). We hypothesized that EMT results in improved performance in unfamiliar scenarios (adaptive transfer), as compared to an error-avoidance method. Our secondary hypothesis was that EMT is associated with improved adaptive transfer performance after a sustained period of training quiescence. These results may have direct implications on teaching strategies, including how students and instructors chose to manage errors.

2.2. Materials and Methods

2.2.1 Recruitment and Ethical Approval

An overview of the study timeline is illustrated in **Figure 2.1**. Second year veterinary students at the Oregon State University Carlson College of Veterinary Medicine (OSU-CCVM) were eligible for inclusion. To ensure a cohort of novice participants, the experiment was conducted prior to beginning the OSU-CCVM introductory general surgical skills course. Students who self-identified as having extensive surgical experience or formal training in tying surgical knots were ineligible.

Ethical approval was granted by the Oregon State University Institutional Review Board (approval number 8194). Participation in the study was voluntary and all students received an explanation of research prior to enrollment. Recruitment was led by a graduate researcher (DM) to minimize perceived coercion. Participants were unaware of the specific research aims and hypotheses; incomplete disclosure was deemed necessary to determine the educational value of the intervention. All participants were fully debriefed upon conclusion of the study and given the opportunity to have their results withdrawn. Data were anonymized with unique identification numbers. Recruitment incentives were a free catered meal and entry into a gift card raffle.

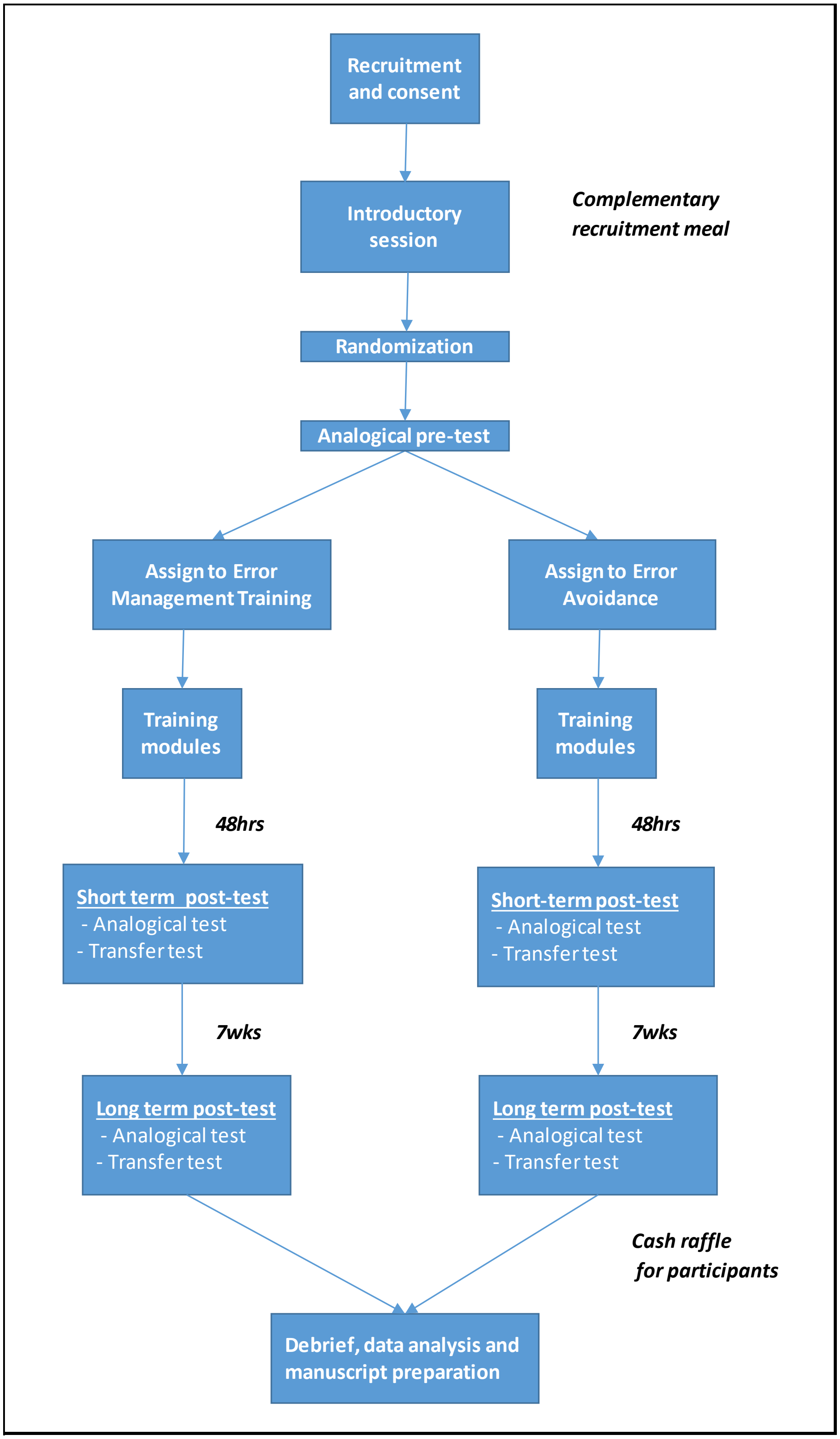

Figure 2.1. **Flowchart demonstrating progression through the study.**

2.2.2 *Introductory training and analogue pre-test*

At the beginning of the introduction session, each participant was given a survey assessing their surgical expertise (1 = no prior experience performing hand ties; 2 = student has practiced hand ties, but has not performed a hand tie in a real surgical setting; 3 = student has practiced hand ties, and has performed a hand tie in ≤5 real surgeries; 4 = student has practiced hand ties, and has performed a hand tie in >5 real surgeries). All students were randomly assigned into error avoidance training (EAT) or error management training (EAT) using a freely accessible random number generator (https://www.google.com/search?q=random+number). Randomization did not take into account responses to the introductory survey.

Participants received a 30 minute, didactic, introductory session intended to briefly summarize the purpose of surgical knot ties (Boothe 2018). The introductory session was delivered in the form of a PowerPoint presentation by a board-certified veterinary surgeon (KT). Introductory training also included a real-time demonstration of an instrument tie and a one-handed square knot. Baseline performance was determined by an analogical pre-test (i.e. directly analogous to training method), taken within 60 minutes of completing the introductory session. During the pre-test participants were asked to repeat skills covered in the introduction using a surgical knot tying board and color-coordinated rope (Ethicon US, Cincinnati, OH, USA; **Figure 2.2A**). Participants were asked to tie two consecutive square knots using two different techniques: firstly an instrument tie using needleholders, and secondly a one-handed tie. Participants were given up to two minutes to complete each task. Hand movements were recorded with participant number for identification; videos were used for evaluation of outcome measures only.

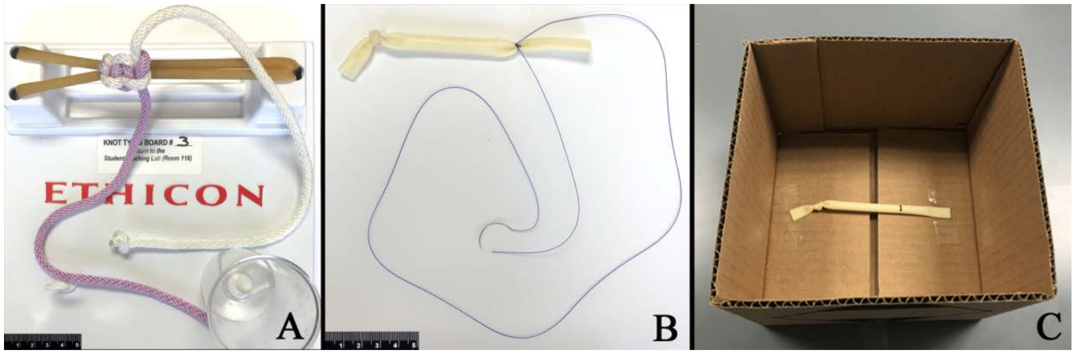

Figure 2.2. Images demonstrating materials used for training and testing. A. Suture board and rope used for the pre-test, training and analogue short-term test. The rope has been tied in two consecutive square knots. B. Short- and long-term adaptive transfer tests utilized a penrose drain and 3-0 PDS suture. The image shows two consecutive square knot around the drain, simulating a ligature. By using suture as opposed to rope, the adaptive transfer test required participants to apply their expertise to an unfamiliar set of conditions. C. The long-term adaptive transfer test introduced an additional layer of unfamiliarity by asking participants to tie knots around a penrose drain secured at the bottom of a box that restricted hand movements. Participants were asked to tie the knot around the marked line.

Outcome measures for both skills were time for completion (seconds) and number of attempts. For instrument ties, participants were also given a technical score out of four with one point assigned for each of the following categories: tying two square knots (yes/no), appropriate technique (correct/incorrect), instrument handling (correct/incorrect), and tying two knots (yes/no). The latter category was created to give credit for those participants who created two knots that were not necessarily square. For hand ties, participants were given a technical score out of three (two square knots, two knots tied, appropriate technique). Each video recording was evaluated by two trained personnel (DM and KT), blinded to the participant and method of training received. Final scoring for each video was decided by consensus. Participants who started the test, but were unable to complete the tasks and declined to use the whole time, were assigned 120s of time. Number of attempts and number of errors were scored up to the point at which they elected to end the test.

2.2.3 Training and experimental interventions

Groups separately met for one, 90 minute training session delivered seven (EAT) or eight (EMT) days after the introductory session. Training sessions consisted of three 30 minute modules to be completed sequentially: instrument tie, hand tie (one-handed technique) and open exploration (time to practice either instrument or hand tie, at discretion of participant). All participants received an individual printed color hand out containing static images and written instructions of the two skills. Images and instructions were taken from a free access online resource (http://surgsoc.org.au/wp-content/uploads/2014/03/Ethicon-Knot-Tying-Manual.pdf). Continuously looping videos of each skill were also displayed on an overhead projector. The instrument tie video included demonstration of correct instrument handling. Skills were practiced on a surgical knot tying board with color-coordinated rope, as previously.

Scripted heuristics were designed according to existing literature and are displayed in **Table 2.1** (Frese 1991). Respective EMT and EAT heuristics were delivered in the printed handout, and stated to participants at the beginning of the session. Thereafter instructions were verbally repeated at 15 minute intervals for the remainder of the session. As per previous descriptions of EMT, training sessions were designed to give participants only minimal instructor guidance, with a view to creating an environment that encourages active exploration (Keith 2008). As such, instructor interventions were only offered upon request. For both groups, verbal feedback was limited to “correct” or “incorrect” and a brief explanation of the error if applicable. Each instance of instructor feedback, regardless of participant performance, was accompanied by an appropriate intervention heuristic that was individually re-stated to the participant (i.e. “you got it right – well done! Keep practicing the same way” for EAT; “you got it right, remember you can learn from getting it wrong too” for EMT).

Table 2.1: Scripted heuristics for Error Avoidance Training (EAT) and Error Management Training (EMT)

EAT	EMT
We want you to know these tasks are really challenging! Follow the instructions exactly as demonstrated.	We want you to know these tasks are really challenging! You should expect to make lots of mistakes when learning these skills.
If you closely follow the instructions you should be able to complete them easily	Although the instructions are very specific, you do not need to stick to them exactly as demonstrated! Experiment with the technique as much as you like.
Try to make as few mistakes as possible. Do everything you can to get the correct knot as efficiently as possible!	Mistakes you make are great! The more mistakes you make, the more you will learn
You got it right – well done! Keep practicing the same way	Use your mistakes to help you learn better! Call over an instructor to check your technique when you like.
That’s incorrect. Go back and review the instructions/video closely!	You got it right. But remember you can learn from getting it wrong too! You can call me over as many times as you like
Make sure you read the instructions really closely!	That’s incorrect. Remember, don’t worry about getting the tie wrong– these are a natural part of your learning experience
Aim to do the best you possibly can	That’s incorrect. Try exploring variations in technique! These instructions are just a broad guide
	You made a mistake - now try to figure out where you might have gone wrong and how you can approach it differently.
	You have made a mistake? Great! Now you can learn something new.

2.2.4 Testing and surveys

Participants were tested 48 hours (hereafter referred to as ‘short-term’) and 7 weeks (hereafter referred to as ‘long-term’) after completion of training. The short-term test had an analogical component identical to the pre-test (i.e. instrument and hand tie using rope), and an adaptive transfer component which used an unfamiliar scenario of which they had no prior notification. The long-term testing window evaluated adaptive transfer only.

Adaptive transfer tasks were designed to test participant's ability to apply skills to an unfamiliar scenario. Tasks were intended to be structurally different from training, requiring modification of the learned procedures (Keith 2008) Participants were asked to tie two consecutive square knots using 3-0 PDS suture (product Z316H, Ethicon, Cincinnati, OH). This material contrasts from rope training due to differences in width, length, extensibility, memory, tensile strength and a needle at one end (**Figure 2.2B)**.

Knots were to be tied firstly by the instrument technique and secondly by one-handed technique. For both short and long-term adaptive transfer tests, participants were asked to tie knots around a marked ¼" penrose drain (Covidien, Mansfield MA, USA) with a knot at one end. Additional unfamiliarity was added to the long-term test by having participants tie knots within a 20cm x 20cm x 20cm box that restricted hand movements (**Figure 2.2C)**. This task was intended to simulate the ligation of tubular structures (i.e. ovarian pedicles, arteries) within a body cavity. For all adaptive transfer tests, participants were given the following additional instructions: "This drain simulates an artery. Please tie your knots over the line tightly, so as to achieve hemostasis. We will pressure test the strength of your knots once you are done". For both groups, participants were asked to complete the task exactly as demonstrated in the printed notes and videos. For the EMT participants specifically, instructions were supplemented with: "During training we asked you to experiment and make mistakes. This time please complete exactly as instructed". Participants were given up to three minutes to complete each task. Recorded hand and instrument ties were reviewed and scored as previously described.

A digital manometer (General Tools & Instruments ® DM8215, Melrose, MA) was used to determine the water pressure at which fluid leaked from the ligature (hereafter referred to as knot leaking pressure). By way of a four way stopcock (Component supply company, Sparta TN, USA, a 25-gauge butterfly catheter (Terumo, Somerset NJ) was connected to a syringe pump (Medfusion ® 3500, Smiths Medical ASD, Inc., St. Paul, MN) equipped with a 60mL syringe (Covidien, Mansfield MA, USA). The syringe was filled with blue-colored tap water and pumped at a rate of 300mL/hr into the penrose drain. Pressure (psi) was recorded when blue-colored fluid appeared past the ligature.

Upon completion of the short-term assessment, participants were asked to complete a 10 question survey that evaluated attitudes towards errors, ability to manage unfamiliar scenarios, metacognition, and whether students were performance goal orientated or learning orientated (Elliott 1988). Participants were asked the extent to which they agreed with each statement, scored on a five point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*). The survey also included opportunity for open comments.

2.2.5 Statistical analysis

All statistical analyses were performed using GraphPad Prism 7.04 for Windows (GraphPad Software, San Diego, CA). Normality of the data was tested by D'Agostino and Pearson test. Data are presented as median (interquartile range, IQR) unless otherwise stated. Continuous data (time, leaking pressure) were analyzed by student's t-test (if normally distributed) and Wilcoxon rank-sum tests (if non-normally distributed). Paired continuous data were analyzed by a paired t-test or Wilcoxon matched-pairs signed rank test. Wilcoxon matched-pairs signed rank tests were used for paired ordinal data (analogue pre-test and short-term test; short and long-term adaptive transfer tests). All *P* values were two-tailed. For all test outcomes, pre-analyses were performed to determine if data could be combined or stratified for subsequent hypothesis testing. For the survey data, it was concluded that analysis beyond descriptive observations would be inappropriate due to sample distribution.

To reduce the possibility of false positive Type I error associated with performing multiple tests, a false discovery rate (FDR) approach was used (Benjamini 1995, Chen 2010). While this procedure has been primarily applied to large scale genetic studies, the FDR procedure may be used if a small fraction of false discoveries are deemed acceptable (Chen 2010). This was selected over alternative strategies that control family-wise error rate associated with multiple testing (i.e. Bonferroni adjustment), noting that such an approach might have been too stringent (Chen 2010). The FDR analysis performed on *P*-value stacks was set for a desired FDR of Q = 5%.

2.3 Results

Forty four participants enrolled in the study (22 EMT and 22 EAT). One participant from each cohort withdrew after training and their data was excluded from analysis. The pre-training questionnaire evaluating previous knot tying experience found a median score of 1 (no prior knot tying experience; n = 32) with the 75th percentile at 2 (have practiced surgical hand ties but never in a surgical setting; n = 8). One participant indicated <5 knots in a real surgical setting, and two participants indicated ≥5 knots in a real surgical setting. One of the participants in the latter category indicated this was approximately 15 years ago. Following randomization, the median experience for EAT was 1 (1-1) and for EMT the median was 1 (1-2).

All subsequent statistical analyses that generated two-tailed *P*-values were analyzed according to the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli (Benjamini 1995). *P*-values < 0.039 were classified as ‘discoveries’.

2.3.1 Performance in analogue pre-test and short term test

Times to complete the analogue instrument and hand ties are shown in **Table 2.2**. For all participants, the median pre-test and short term times for the instrument tie were 45s (39-54s) and 31s (27-37s); this difference was statistically significant (two-tailed $P < 0.0001$). For the hand ties, times were 87s (57-120s) and 27s (20-36s) respectively; this difference was also significant (two-tailed $P < 0.0001$). Technical performance is shown in **Figure 2.3**. For all participants, the total number of technical errors in the pre-test was 72 (instrument tie) and 93 (hand tie). At the short term test, there were 41 and 21 errors for the two respective tests. For the instrument tie, errors in all four categories decreased in the range of 25-77%. In the hand tie, errors in all three categories decreased in the range of 70 – 84%. Total attempts for the instrument and hand ties decreased by 15 (23%) and 80 (46%) attempts respectively, between the two time points (**Figure 2.3**, right half). There were significant differences in technical errors and number of attempts for both the instrument and hand ties ($P \leq 0.013$).

Table 2.2. Performance in analogue tests, before (pre-test) and shortly after training (short-term).

Results are shown as the median (IQR) and compared by Wilcoxon rank-sum tests at one time point (*listed in rows for each outcome) and Wilcoxon matched-pairs signed-rank tests between two time points (**listed in far right column). Bolded *P*-values are those that are statistically significant after performing a false discovery rate (FDR) controlling procedure.

Table 2.2. Performance in analogue tests, before (pre-test) and shortly after training (short-term).

		Pre-test	Short-term	*P*-value**
INSTRUMENT				
Time				
	All	45 (39-54)	31 (27-37)	**<0.0001**
	EAT	45 (37-56)	29 (25-36)	**<0.0001**
	EMT	45 (40-54)	32 (27-40)	**0.0039**
	P-value*	-	0.16	
Number of attempts				
	All	1 (1-2)	1 (1-1)	**0.013**
	EAT	1 (1-2)	1 (1-1)	0.063
	EMT	1 (1-2)	1 (1-1)	0.22
	P-value	-	1	
Total errors				
	All	2 (1-3)	1 (0-2)	**0.0004**
	EAT	2 (1-3)	0 (0-2)	**0.0017**
	EMT	1 (1-2)	1 (0-2)	0.11
	P-value*	-	0.64	
HAND				
Time				
	All	87 (57-120)	27 (20-36)	**<0.0001**
	EAT	87 (60-120)	25 (19-33)	**<0.0001**
	EMT	66 (56-120)	27 (21-39)	**<0.0001**
	P-value*	-	0.52	
Number of attempts				
	All	4 (2-6)	1 (1-2)	**0.0001**
	EAT	3 (2-6)	1 (1-2)	**0.014**
	EMT	4 (2-6)	1 (1-2)	**0.0049**
	P-value*	-	0.76	
Total errors				
	All	3 (2-3)	0 (0-1)	**<0.0001**
	EAT	2 (2-3)	0 (0-1)	**<0.0001**
	EMT	3 (1-3)	0 (0-1)	**<0.0001**
	P-value*	-	0.67	

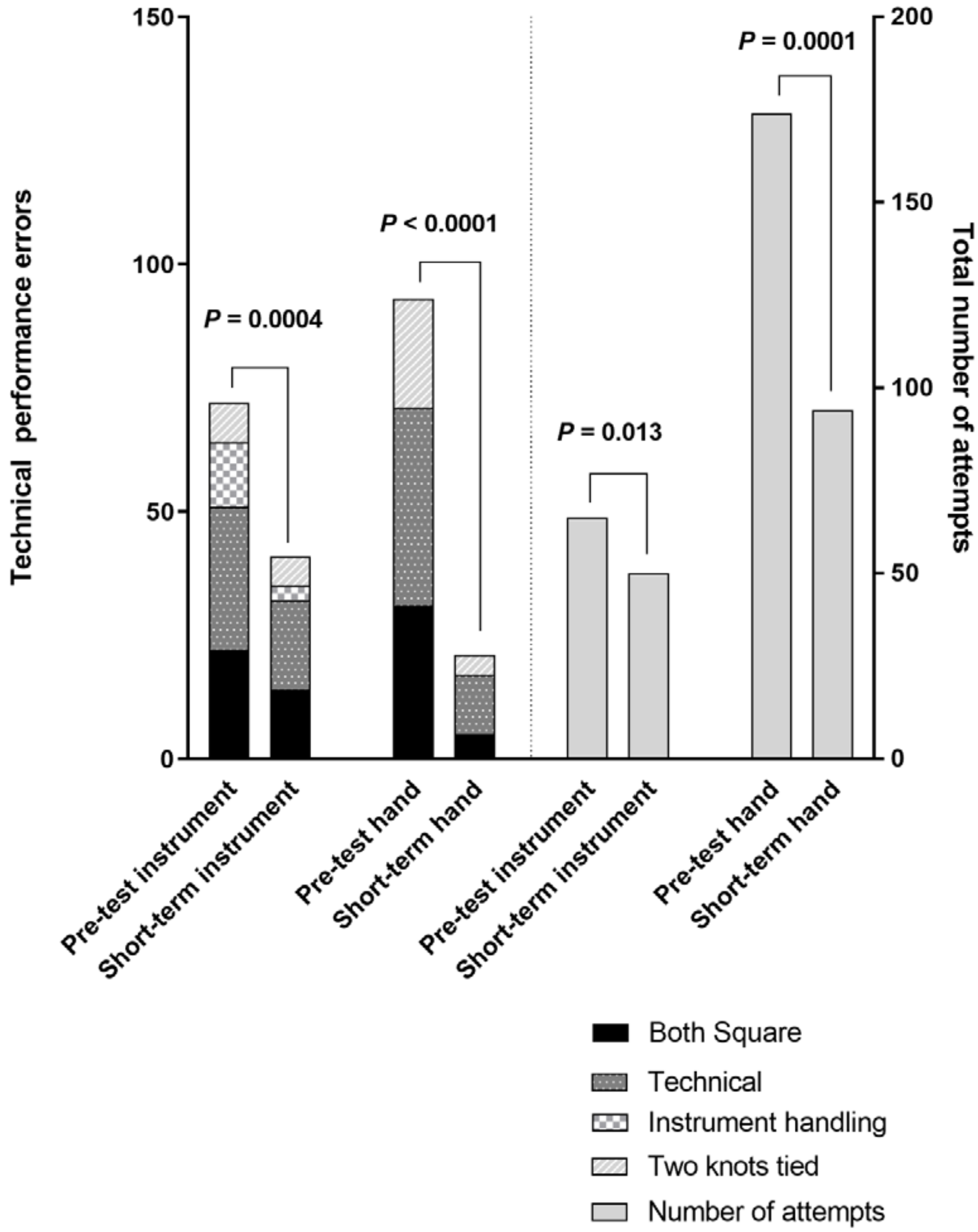

Figure 2.3. Analogue performance for all participants. A stacked bar chart showing performance in instrument- and hand-ties, before (pre-test) and shortly after training (short-term). Technical performance errors are shown in the left half and number of attempts are shown in the right half. Technical performance errors are subdivided into instrument handling (instrument tie only), other technical errors, squareness of knot, and two knots tied. Bolded *P*-values are those that are statistically significant after performing a false discovery rate (FDR) controlling procedure. All participants showed significant improvement in both technical performance ($P \leq 0.0004$) and number of attempts ($P \leq 0.013$).

2.3.2 Performance in transfer short term and long term tests

Descriptive statistics for leaking pressures at the two time points are shown in **Table 2.3**. Among all participants, there was significant decline in leaking pressures for the instrument tie (P = 0.0086) but not in the hand tie (P = 0.27). The same period saw a decline in technical performance (i.e. greater errors), as shown in **Figure 2.4**. Total technical errors increased by 27 (58.7%) for instrument ties and 44 (157.1%) for hand ties (range from 50% to 500% within individual categories). Between the two time points, total technical errors for each participant were significantly different for both instrument ties (P = 0.0078) and hand ties ($P < 0.0001$). Attempts were significantly less for the instrument tie (P = 0.0061) and more for the hand tie (P = 0.0018).

Table 2.3. Performance in transfer tests, shortly after training (short-term) and 7 weeks after training (long-term). Results are shown as the median (IQR) and compared by Wilcoxon rank-sum tests at one time point (*listed in rows for each outcome) and Wilcoxon matched-pairs signed-rank tests between two time points (**listed in far right column). Bolded *P*-values are those that are statistically significant after performing a false discovery rate (FDR) controlling procedure.

Table 2.3. Performance in transfer tests, shortly after training (short-term) and 7 weeks after training (long-term).

		Short term	Long term	*P*-value**
INSTRUMENT				
Leaking pressure (psi)				
	All	11 (9-18)	10 (7-14)	**0.0086**
	EAT	10 (8-19)	10 (7-12)	0.33
	EMT	13 (11-28)	10 (5-16)	**0.0085**
	P-value*	0.053	0.75	-
Time				
	All	61 (50-77)	77 (59-94)	**0.021**
	EAT	61 (51-77))	79 (61-97)	**0.0057**
	EMT	60 (45-78)	67 (55-94)	0.49
	P-value*	0.86	0.31	-
Number of attempts				
	All	3 (2-3)	2 (1-3)	**0.0061**
	EAT	3 (2-4)	2 (1-3)	0.045
	EMT	3 (2-3)	2 (1-3)	0.086
	P-value*	0.46	0.78	-
Errors				
	All	1 (0-2)	2 (0-3)	**0.0078**
	EAT	1 (0-2)	2 (1-3)	**0.023**
	EMT	1 (0-2)	1 (0-3)	0.16
	P-value*	0.69	0.29	-
HAND				
Leaking pressure (psi)				
	All	15 (10-23)	11 (9-19)	0.27
	EAT	14 (9-22)	10 (8-15)	0.23
	EMT	15 (12-27)	13 (11-25)	0.66
	P-value*	0.36	0.041	-
Time				
	All	46 (37-59)	66 (50-117)	**<0.0001**
	EAT	45 (35-55)	95 (48-133)	**<0.0001**
	EMT	47 (38-70)	66 (53-90)	0.1
	P-value*	0.27	0.041	-

Table 2.3. Performance in transfer tests, shortly after training (short-term) and 7 weeks after training (long-term) "(Continued)"

Number of attempts				
	All	1 (1-2)	3 (2-8)	**0.0018**
	EAT	1 (1-2)	6 (3-9)	**<0.0001**
	EMT	2 (1-3)	2 (1-4)	0.75
	P-value*	0.19	**0.023**	-
Errors				
	All	0 (0-1)	2 (1-3)	**<0.0001**
	EAT	0 (0-1)	2 (1-3)	**0.0002**
	EMT	0 (0-1)	2 (1-3)	**0.0009**
	P-value*	0.93	0.93	-

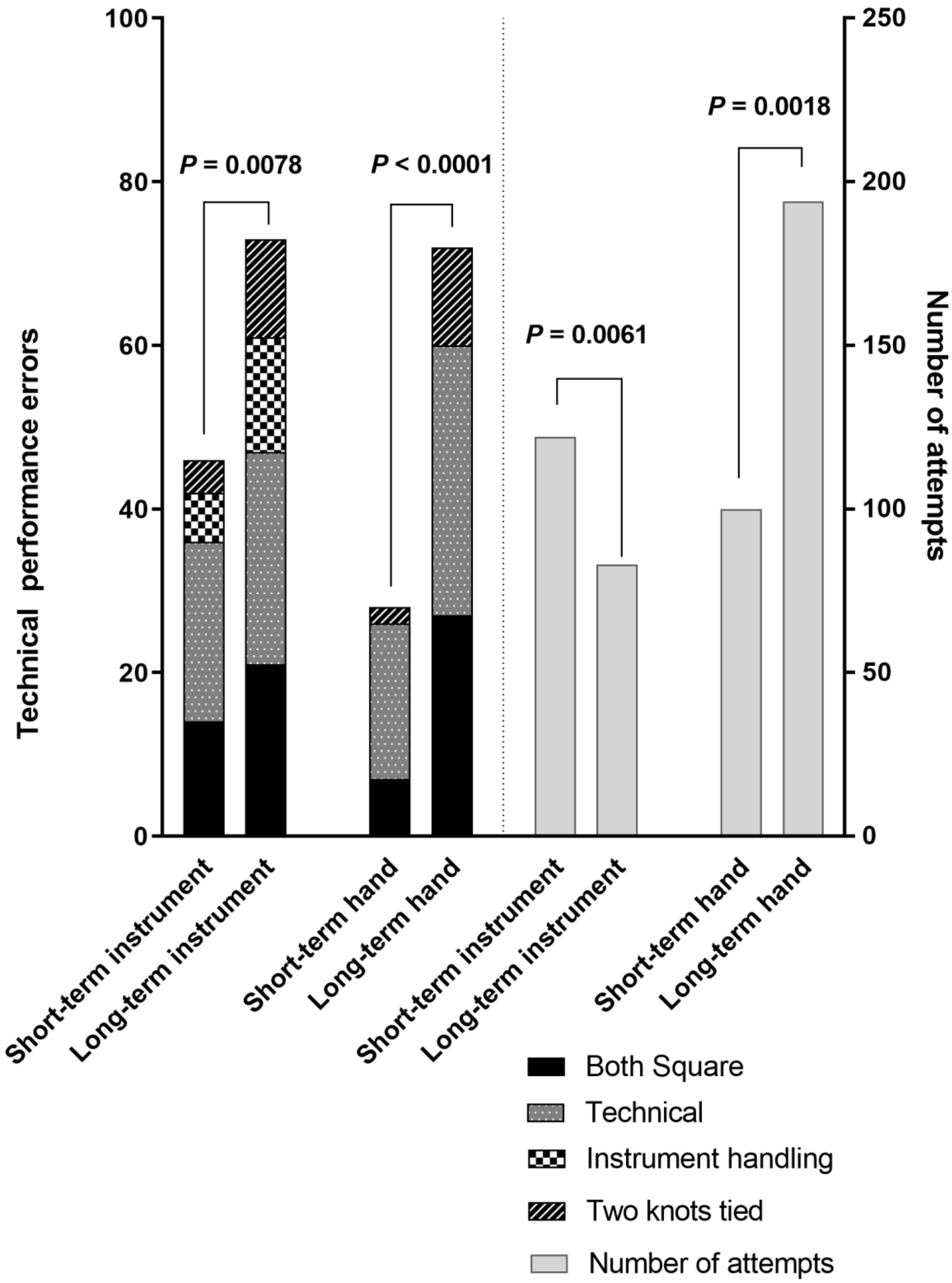

Figure 2.4. Transfer performance for all participants. Stacked bar chart showing performance in instrument- and hand-ties for all participants, shortly after training (short-term) and 7 weeks after training (long-term). Technical performance errors are shown in the left half and number of attempts are shown in the right half. Technical performance errors are subdivided into instrument handling (instrument tie only), two knots tied, other technical errors and squareness of knot. Bolded P-values are those that are statistically significant after performing a false discovery rate (FDR) controlling procedure. All participants showed significant improvement in both technical performance and number of attempts ($P \leq 0.0078$).

2.3.3 Performance between EAT and EMT

Performance differences between EMT and EAT groups are shown in **Table 2.2** (analogue tests) and **Table 2.3** (adaptive transfer tests). In the analogue short-term tests, performances were not significantly different in any of the outcomes (P all ≥ 0.16). For the short-term adaptive transfer outcomes, only leaking pressure for the instrument tie approached statistical significance ($P = 0.053$). At the long-term test, EMT participants took significantly fewer attempts for the hand tie ($P = 0.024$); leaking pressures and time for the hand tie approached significance (both $P = 0.041$) after the FDR correction.

By comparing performance trends within groups over the seven week interim period, paired data were used to evaluate the second hypothesis. In so doing, the paired design allowed individuals to serve as their own matched control. A total of seven paired outcomes yielded statistically significant differences among all participants (all four outcomes for instrument tie; 3/4 outcomes for the hand tie). In the instrument tie, EAT was primarily responsible for slower times and greater errors. For the hand ties, EAT was primarily responsible for the slower times and greater number of attempts.

2.3.4 Post-training surveys

Survey results are shown in **Figure 2.5**. **Question 1** was designed to evaluate participants' receptiveness to challenging situations, and specifically if participants were learning goal oriented (Heimbeck 2003). Greater than 95% of participants in both groups either agreed or strongly disagreed. **Question 2** aimed to determine if students were performance goal oriented, motivated by demonstrated good performance and favorable judgment (Heimbeck 2003). These participants might perceive errors as indicators of poor performance (Heimbeck 2003). Seventy percent or more participants in both groups agreed or strongly agreed with this statement.

Performance-avoid goal orientation (i.e. tendency to evade negative judgment) was tested in **question 3**, with 15-43% agreeing and 33-45% disagreeing. **Questions 4 and 5** evaluated perceived performance in adaptive transfer: 45% or more in both groups generally felt that training prepared them for unfamiliar scenarios. Between 19.1% - 40% of participants felt unprepared for adaptive transfer. **Question 6** evaluated participants' perception of mistakes: ≥ 90% in both groups felt that mistakes were an important part of their training. **Questions 7-10** assessed student self-awareness: across all questions and groups, ≥ 50% positively reflected on their ability to self-evaluate.

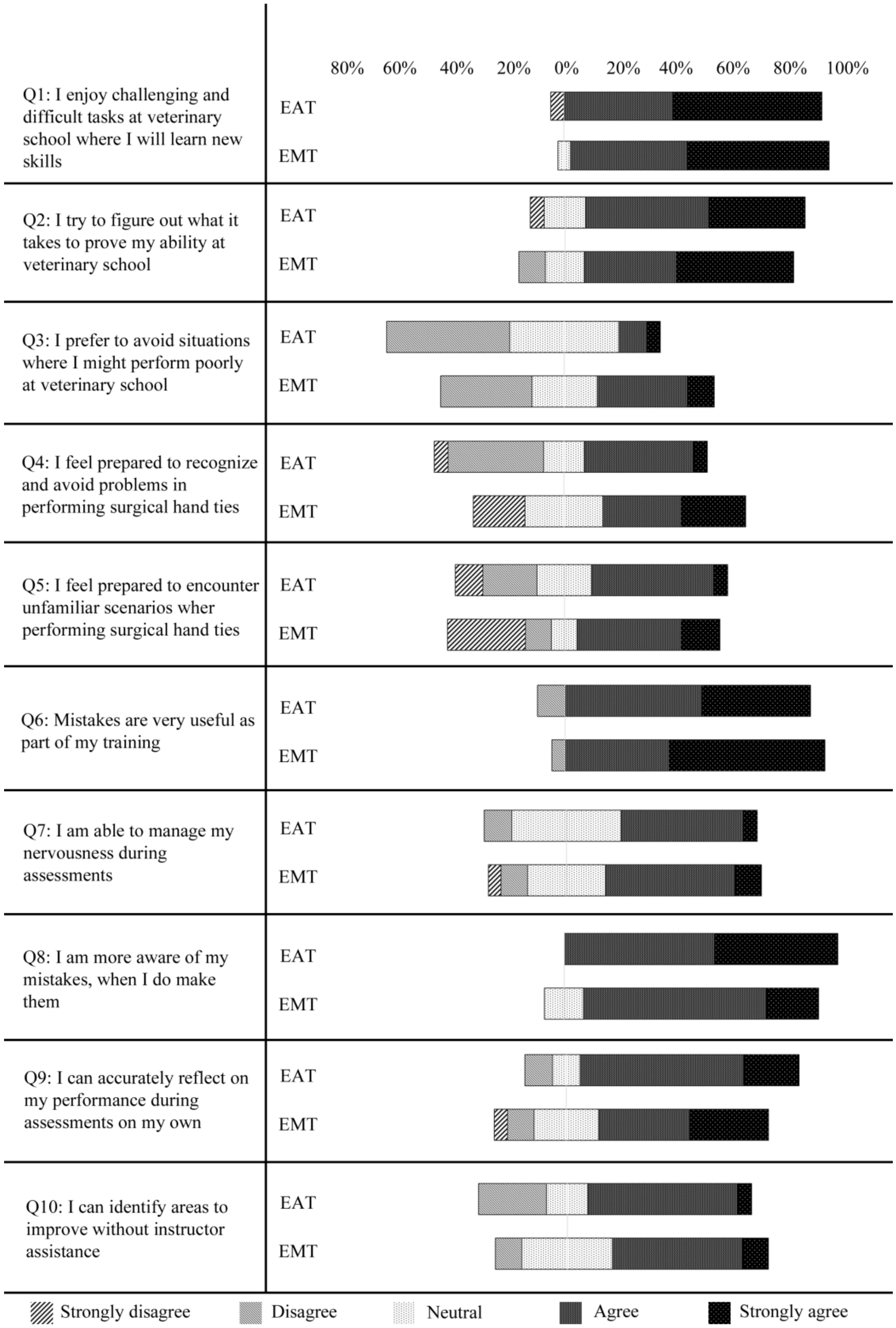

Figure 2.5. Results of 10 survey questions taken immediately after short-term testing. Data are expressed on a five point Likert scale (strongly disagree, disagree, neutral, agree, strongly agree). The Y-axis is centered on the median neutral responses, with agree/strongly agree responses on the right half and disagree/strongly disagree responses on the left half. EAT participants are shown above EMT participants.

A total of 19/42 participants provided written feedback in the open comments section (11 EAT, 8 EMT). Many comments expressed broad appreciation for learning a new, clinically applicable skill (i.e. “I really enjoyed this”; “will help me a lot next term and after”). Some participants indicated a preference for learning by video as opposed to printed hand outs (4/19). One comment in the EMT group specifically pertained to the experimental intervention: “I enjoyed the feedback – mostly to show off the off hand, eyes closed method! This was fun”. Seven out of nineteen comments contained feedback directed at the training format necessitated by experimental design. Examples included: “I thought it was great that there was no communication that we would be expected to perform the tasks with suture and to challenge us in that way”; or “I did not call the instructor over more since they couldn’t actually help, only say if I was doing something right or wrong”.

2.4 Conclusions

This study sought to determine how veterinary students utilize errors, and if errors could be leveraged to improve surgical knots created by novice surgeons. Specifically, this study aimed to evaluate the impact of EMT in veterinary students learning to apply surgical knots to unfamiliar scenarios. EMT modifies attitudes towards errors, and in doing so, might better prepare students for real surgery and routinely arising unanticipated scenarios. Our prospective, blinded and partially double blinded study tested the hypotheses that participants trained by EMT would out-perform those trained by an error avoidance method (EAT) when asked to perform unfamiliar (adaptive transfer) skills. We also hypothesized that EMT students would demonstrate improved adaptive transfer performance after a period of training quiescence. While there were no significant differences between EMT and EAT at the short-term and long-term tests, these data suggest that at a minimum, EMT is comparable to error-avoidance surgical knot training.

This finding itself is notable, given the dramatically contrasting heuristics delivered to the two groups (i.e. “you do not need to stick to [the instructions] exactly as demonstrated”, “try exploring variations in technique” versus “follow the instructions exactly as demonstrated”). These findings may have important implications for educators within both didactic and laboratory training environments.

Training produced demonstrable improvements in analogous hand and instrument ties, showing a positive effect of the study training procedure generally (significant *P*-values detected in all 6 outcomes after FDR correction). Faster times were especially notable for both the instrument and hand ties, likely due to improved proficiency and confidence associated with the two skills. There were broadly similar performances between the two groups: of these five, only one was disproportionately affected by the EAT (total errors – instrument). Previous studies have suggested that students trained by proceduralized methods might perform comparably to EMT participants in tasks that directly recapitulate training (Heimbeck 2003; Keith 2005).

These data fail to confirm the hypothesis that EMT results in superior performance in short-term adaptive transfer as it relates to surgical knot tying. Results were not significantly different between the two groups, but approached statistical significance for one of the eight outcome measures (leaking pressure – instrument tie; $P = 0.053$). EMT has been successfully applied to a variety of other training settings, most notably word processing and computing, (Nordstrom 1998; Chillarege 2003) and also driver training, (Ivancic 2000) radiology technology (Gardner 2014) and medical ultrasound (Dyre 2017). Not all studies have consistently demonstrated that EMT is superior to proceduralized or error avoidance training (Debowski 2001; Gully 2002). A 2008 meta-analysis looked to explain inconsistent reported effects of EMT (Keith 2008). One explanation for varying effect sizes might be the degree to which the adaptive transfer assessment differed from training. In those instances where the tasks resemble training (as partially applicable to this experiment), the positive effects of EMT might be less apparent. Perhaps one additional explanation for discrepancy might be that while our adaptive transfer tests, by definition, added unfamiliarly - they also added a notable degree of difficulty for which participants might have been unprepared for.

Performance for the entire cohort generally deteriorated for the long-term adaptive testing. This was expected given the prolonged interval between training and testing, during which period participants were given no explicit instructions to practice or seek out additional resources. While analogue performance was not tested at 7 weeks, a decline in this might have been expected too. When compared to the short-term test performance, 6/8 outcomes were associated with significant reductions in performance (number of attempts for the instrument tie significantly improved). When comparing changes that occurred within groups, declining adaptive task performance was disproportionately associated with EAT in 4/6 of these outcomes (as opposed to one, leaking pressure – instrument, for the EMT cohort). Even though only 1/8 outcomes had significant differences when directly comparing EAT and EMT at 7 weeks (number of attempts, hand tie), different trends in the paired data lend some support to the second hypothesis that EMT might help students retain adaptive transfer skill after a prolonged period of training quiescence. Perhaps this could be attributed to an improved tolerance for errors, which might inspire students towards a learning goal orientation with improved overall learning (Heimbeck 2003; Bell 2008). Future studies might seek to determine if possible differences in long-term performance are associated with inclinations to practice during the interim period.

Positive effects of EMT have been linked to the way in which it influences trainee attitudes towards errors (Keith 2008). Specifically, EMT works to control the negative emotions associated with errors (Keith 2005; Bell 2008). The post-training survey offers some insight into junior veterinary student perceptions towards errors, and notably their comfort in managing unfamiliar scenarios. The majority of participants recognized the importance of errors in training (**Figure 5**, question 6), including those who received EAT. Yet participants in both groups indicated that training left them feeling unprepared for adaptive transfer (questions 4 and 5). This might reflect the relatively brief period of training and/or the nature of training, which by experimental design, did not directly recapitulate test scenarios. Alternatively, difficulty associated with the adaptive transfer tests might have left participants with a greater awareness of the improvements necessary to be competent in these skills.

Metacognition also mediates the effectiveness of EMT (Keith 2005; Bell 2008). Errors can serve to break a routine, requiring deeper mental engagement and higher order thinking (Ellis 2005). An inevitable by-product of errors is the planning, monitoring, evaluation and revision of subsequent tasks. Survey questions 7-10 served as a tool for self-evaluating metacognition. While responses to these questions were largely positive, the training itself could be completed with relatively limited self-reflection. Commonly EMT participants liberally experimented with technique: such examples included reversing hands for ties, holding instruments in different ways, changing directions of throws, tying with eyes closed and attempting different knots entirely. However, instructors also noted that EMT participants did not consistently reflect on these deviations – especially the impact they might have had on the overarching objectives of surgical knot tying (i.e. generating tight, strong and stable knots while maintaining adequate control of the suture material). Motivation to learn from errors may be especially potent for those errors with severe consequences (Zakay 2004). Unsystematic ‘trial and error’, not guided by a clear hypothesis or associated with reflection, is not necessarily an effective learning strategy (van der Linden 2001). The impact of EMT might have been more robust if training made a more explicit effort to emphasize the relationship between errors and impact, thus integrating metacognition to a greater extent.

In order for EMT to be effective, participants must be able to recognize errors shortly after they occur. In question 8 participants were asked to reflect on their ability to detect errors, with the majority indicating they were aware of their own mistakes. The ability to detect errors has been associated with the complexity of a task, with detection lower in more complex tasks (Zapf 1994). For the purposes of this experiment, our investigative team interpreted these tasks to have a high degree of complexity (i.e. complex series of rope/suture manipulations, combined with handling of a needle, suture memory and working within a confined space). Perhaps for complex skills such as knot tying, participants might need an improved feedback mechanism for detecting errors, and/or longer periods of EMT than that offered.

The survey instrument also examined personality traits capable of interacting with EMT. Learning goal orientation describes those individuals focused on improving their ability and mastering new tasks over time (Elliott 1988). ‘Performance-prove’ goal orientation defines those who look for favorable judgment of competence; ‘performance-avoid’ orientation is those who look to avoid negative judgement of competence (Mangels 2006). These students, who view ability as ‘fixed’ might in theory respond more favorably to EAT (Van Dyck 2010). While the survey instrument offers a relatively superficial assessment of goal orientation, overall participants in this study reported themselves to have traits that identify as both learning goal orientated (question 1) and performance-prove goal orientated (question 2). Responses to identify performance-avoid orientation (question 3) were more evenly mixed for EMT participants with slightly more EAT participants disagreeing with a performance-avoid orientation. Those students who identify with performance-prove and performance-avoid orientation might respond more favorably to EAT (Heimbeck 2003).

Knot leaking pressure was considered to be the primary measure of performance in the adaptive transfer test. However, important limitations relate to this. By intention, participants received no directed training in how to tie knots tightly (hence adaptive transfer). Participants did however receive broad statements outlining the purpose of surgical knots during the introductory session (i.e. “knots are used to tightly secure ligatures around an artery”). The experimental design aimed to determine whether different training circumstances would produce differences in knot performance, not necessarily participants’ ability to tie knots sufficient to tightly secure a ligature. These data clearly indicate that neither EAT or EMT, as presented in this exact format, was sufficient to make trainees fully competent in this skill. Knots performed by a board certified veterinary surgeon, tested during a pilot phase of this project, tested between 250-300psi (data not shown). Both groups might have benefitted from additional training in this area. One additional limitation with using knot leaking pressure as the primary measure of performance might be that participants could have chosen to prioritize the technical elements of surgical knot tying, as opposed to tying tightly applied, strong knots. A relatively small number of errors at both the short and long-term adaptive transfer windows might support this idea.

Inherent limitations with some of the outcome measures also complicate interpretation of these data. For example, the ‘number of attempts’ and ‘number of errors’ categories did not always reflect performance, particularly for those students who elected to stop the test early, recognizing that they would be unable to complete the task. In these instances, such tests might have been associated with only one attempt or measurable error. This might partially explain the fewer errors at 7 weeks in the instrument tie. Such limitations with the outcome measures were justification for analyzing multiple measures of performance.

EMT requires that participants are given ample opportunity for exploration during training, often without clear instructions. The latter was difficult to create for tasks that required a high degree of precision, such as surgical hand ties. Another confounding factor might be the type of errors generated by students during training. While still deviating from an identified standard, violations differ from errors by virtue of their conscious intention (Frese 2015). During training instructors noted that many EMT participants would intentionally create errors. Although students were encouraged to “experiment with the technique”, purposely generated errors were not the intention of the scripted heuristics.

Performance trends identified in these data provide some support for incorporating EMT into veterinary education. Rather than avoiding mistakes, as may be the tendency among some veterinary students, instructors might utilize EMT to frame errors as part of a positive learning experience. EMT instructions have been shown to improve learning goal orientation, even among those students who do not readily identify with this mindset (Bell 2008). This strategy might be particularly applicable to circumstances where a primary goal of training is the transfer of skills to new or complex tasks.

Ch. 3 Investigating the Effects of Error Management Training for Veterinary Students Learning Blood Smear Analysis

3.1 Introduction

Creating and evaluating blood smears generates valuable clinical data that can be a critical element of overall patient assessment (Shattil 2003). Findings generated from a blood smear may provide clear evidence of a pathologic process (i.e. anemia, systemic inflammation, neoplasia) or sometimes a specific diagnosis (Bain 2005). In order to generate meaningful conclusions, elements of effective blood smear analysis are fourfold. Firstly, blood must be appropriately smeared to create a diagnostic quality slide that includes a monolayer. The monolayer is especially important for morphologic evaluation as cells in this area dry quickly, are more evenly spread, and lack artifacts commonly identified in other areas of the smear. Secondly, the dried slide must be appropriately stained to facilitate visualization of cell morphology. The third element, slide evaluation, requires cell identification and assessment of major cell types (red cells, leukocytes, platelets). The final interpretative element necessitates that evaluative findings are assimilated with a knowledge of normal and select abnormal processes. Such complexity associated with this procedure is likely to generate a broad spectrum of errors, ultimately capable of manifesting as non-diagnosis or mis-diagnosis that could impact patient management.

The aim of this study is to determine the efficacy of error management training (EMT) as it relates to canine blood smear analysis. Building upon previous findings by this group, training would be supplemented by a module designed to enhance metacognition. As compared to error-avoidance training, we hypothesized that students trained with this approach have improved performance in tasks that are structurally different from those learned in training (here after referred to as ‘adaptive transfer’). We also hypothesized that EMT students would demonstrate improved performance following a period of training quiescence. Data presented herein might influence the way in which veterinary educators chose to leverage errors as learning opportunities in both pre-clinical and clinical veterinary curricula.

3.2 Materials and Methods

3.2.1 Recruitment and ethical approval

First year students at the Oregon State University Carlson College of Veterinary Medicine (OSU-CCVM) were eligible for inclusion. This study was conducted prior to a didactic course on cytological analysis and clinical pathology. Students who self-identified as having extensive experience or formal training in making and/or analyzing blood smears were deemed ineligible. Recruitment entailed two e-mails sent to the 1st year class, as well as an in-person address in the classroom. A complimentary catered meal and gift card raffle were offered as incentives. To obtain unbiased results, the specific research objectives were not disclosed to participants. Following completion of the study, all subjects were debriefed and given the chance to have their data withdrawn. Participant data were anonymized using assigned numbers. All recruitment, incentives and experimental methods were approved by the Institutional Review Board (IRB) at Oregon State University (approval number 8193).

3.2.2 Introductory training and analogue pre-test

Upon successful enrollment, participants completed a self-assessment of blood smear expertise (1=no prior experience creating/evaluating blood smears; 2=practice creating and evaluating blood smears, but never in a real clinical setting; 3=practiced creating and evaluating blood smears and performed in ≤5 real clinical settings; 4 =practiced creating and evaluating blood smears and performed in >5 real clinical settings). Participants were subsequently assigned into either error-avoidance training (EAT) or error management training (EMT) groups using a random number generator (https://www.google.com/search?q=random+number).

All participants received a 30-minute didactic PowerPoint presentation intended to summarize the purpose of blood smear analysis. The presentation reviewed the method by which smears of normal blood are created, stained and analyzed, according to standard procedures (Weiss 2012). Content was guided by delivered by a board-certified veterinary clinical pathologist (EG). The overview also covered identification of the monolayer at 10X magnification, leukocyte identification (neutrophil, lymphocyte, monocyte, eosinophil, basophil) at 40X magnification, and justification for performing a white blood cell (WBC) differential count.

To obtain a baseline for individual participants, pre-tests directly analogous to the training method were administered within 60 minutes of the completion of the introductory session (hereafter referred to as analogue pre-test). Testing required that participants 1) create one quality canine blood smear from up to three attempts, 2) effectively stain their slide with Diff-Quik (Harleco), and 3) correctly identify 20 normal leukocytes on a premade normal canine blood smear, with one co-investigators (EG or DM) navigating the microscope at 40X magnification.

Analogue test outcomes fell into three categories. The first assessment of smear quality was the monolayer area. This variable recognized the monolayer as an especially important area of the smear where cell distribution and morphologic detail is optimally captured (Weiss 2012). Monolayer area in mm^2 was defined as a rectangle calculated as length (mm) x width (mm) using an Olympus BX46 microscope (Olympus Corporation of the Americas, Waltham, MA, USA). Width was defined as the lateral points at which the monolayer first appeared on the slide in a plane perpendicular to the axis of the smear; length measurements were taken from the start of the monolayer to the start of the feathered edge in a plane parallel to the axis of the smear (**Figure 3.1)**. The second category was a qualitative assessment of smear quality and staining, scored out of six. Scores were determined by the presence or absence of common technical errors, examples of which are shown in **Figure 3.2.** Such errors were defined as stutter (horizontal breaks in the smear resulting from uneven pressure during slide preparation), absence of a feathered edge, insufficient blood placed on slide or drying of blood prior to smearing (defined as a clear area above where the blood was spotted), absence of a monolayer (visibly apparent upon microscopic exam), feather not captured on the slide, and smear not adequately stained (inadequate red/purple definition and/or uneven color distribution due to inappropriately short/long immersion in Diff Quik or swirling of slide when in solution). For leukocyte identification, participants were required to identify 20 leukocytes in consecutive 40X fields. Scores were out of a maximum of 20 points.

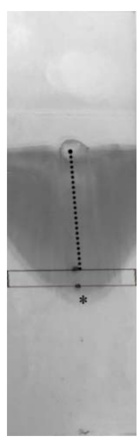

Figure 3.1. Appropriately created and stained canine blood smear. The smear contains an appropriate volume of blood, creating a width that is approximately 75% slide width. There is a feather present (*) and the slide is appropriately stained. Length boundaries of the monolayer were identified with microscopic assistance and marked with pen; the length measurement was taken parallel axis of the smear (space in box between black dots). Width boundaries were defined by the widest aspect of the monolayer perpendicular to the axis of the smear (long borders of rectangle). The area measured in mm^2 is represented by the rectangle.

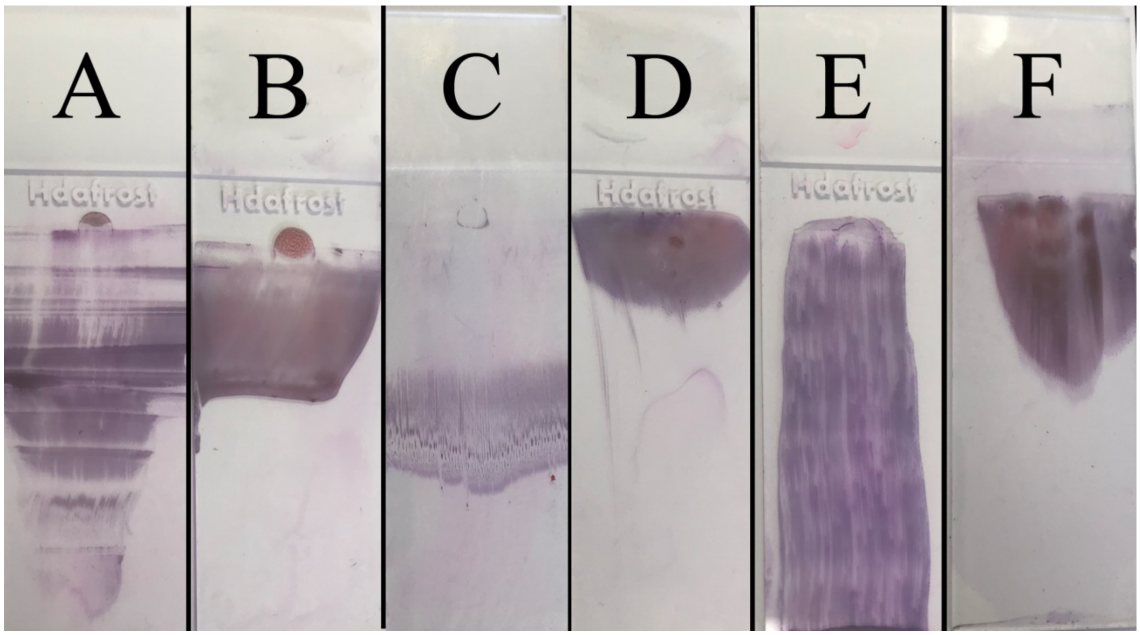

Figure 3.2. Rubric for semi-quantitative evaluation of errors commonly encountered on student-generated blood smears. Blood smears were assigned scores out of six according to commonly identified errors. A. Smear has horizontal breaks in the smear ('stutter') resulting from uneven pressure during smear preparation. The smear also lacks a feathered edge and a monolayer. B. Smear lacks a feathered edge and a monolayer. C. Insufficient blood placed on slide or drying of blood prior to smearing, resulting in a large clear area immediately beneath where the blood was spotted. D. Absence of a monolayer. E. Feathered edge not captured on the slide and the smear lacks a monolayer. F. Smear is inappropriately stained as identified by unevenly distributed red and purple staining.

3.2.3 Training and experimental interventions

The EAT and EMT groups met separately 24 or 48 hours after the introduction session, respectively. Participants individually completed four training modules totaling to 1.5hrs time. Modules were video review, create/stain normal canine blood smear, evaluation of a pre-prepared normal canine blood smear, and open time for any of the prior three modules. All participants received written and illustrated instructions that demonstrated how to create, stain and evaluate a blood smear, according to standard procedures (Weiser 2012). Instructions also included images of diagnostic quality smears, before and after staining. Static images of blood smear components were also illustrated. These included normal red blood cell distribution in the monolayer at 10X magnification, 8 mature neutrophils, 1 small mature lymphocyte, 2 monocytes, 4 eosinophils, and 2 basophils (all leukocytes at 40X magnification). Abnormal features presented in the training document included anemia (defined by reduced RBC density in the monolayer, visible from 10X magnification), erythrocyte regeneration (defined by polychromasia and nucleated RBCs), abnormal RBC distribution (rouleaux and agglutination), and inflammation (defined by band neutrophils and/or toxicity). Toxic changes were defined as neutrophils having hyposegmented nuclei, Döhle bodies, cytoplasmic vacuolization, primary granulation and/or cytoplasmic basophilia.

A list of scripted heuristics was provided in the distributed documents and repeated by co-investigators at approximately 15 minute intervals. These heuristics were devised based on previous publications and are listed in **Table 3.1** (Frese 1991). During training, participants could ask for verbal feedback from instructors as many times as they wished. For the EAT group, feedback was limited to a verbal correct/incorrect response. For the EMT group, feedback also consisted of correct/incorrect replies but with additional encouragement to experiment and not be averse to making mistakes.

Table 3.1: Scripted heuristics for Error Avoidant Training (EAT) and Error Management Training (EMT)

EAT	**EMT**
We want you to know these tasks are really challenging! Follow the instructions exactly as demonstrated.	We want you to know these tasks are really challenging! You should expect to make lots of mistakes when learning these skills.
If you closely follow the instructions you should be able to complete them easily	Although the instructions are very specific, you do not need to stick to them exactly as demonstrated! Experiment with the technique as much as you like.
Try to make as few mistakes as possible. Do everything you can to get the correct knot as efficiently as possible!	Mistakes you make are great! The more mistakes you make, the more you will learn
You got it right – well done! Keep practicing the same way	Use your mistakes to help you learn better! Call over an instructor to check your technique when you like.
That’s incorrect. Go back and review the instructions/video closely!	You got it right. But remember you can learn from getting it wrong too! You can call me over as many times as you like
Make sure you read the instructions really closely!	That’s incorrect. Remember, don’t worry about getting the tie wrong– these are a natural part of your learning experience
Aim to do the best you possibly can	That’s incorrect. Try exploring variations in technique! These instructions are just a broad guide
	You made a mistake - now try to figure out where you might have gone wrong and how you can approach it differently.
	You have made a mistake? Great! Now you can learn something new.

Video modules were included with the intention of encouraging reflection upon blood smear technique. Specifically, the video module given to EMT participants was designed to encourage metacognition, previously shown to be a mediating factor in the efficacy of EMT (Keith 2005). Video shown to the EMT group displayed common errors in creating blood smears and staining techniques. Questions posed as part of this module were designed and modified based upon prior studies using metacognitive instructions (King 1991). EMT participants were asked to 1) identify elements performed incorrectly, 2) reflect upon the goals of blood smear analysis, writing down ways in which mistakes might have influenced the final result, and 3) write down ways in which the procedure might have been performed better. Video shown to the EAT group illustrated proper technique for creating the blood smear and staining the slide. Participants were asked to write down elements performed correctly. For both groups, videos were looped on projectors throughout the entire training session.

3.2.4 Testing and Surveys

Short term assessments were performed 24-72 hours after training. Investigators were blinded as to the training method received. Students were given up to 45 minutes to complete an analogical post-test (identical to the pre-test described above) and one adaptive transfer test (unfamiliar task). The adaptive transfer test used a sample unknown to the participants. To create additional unfamiliarity, the test used normal canine blood, diluted with 0.9% sterile NaCl solution, with the intention of mimicking a non-regenerative anemia. Anemic blood has different smearing properties, thus creating a scenario different from training. Participants were asked to create and stain a smear as previously. To ensure that the test adequately tested adaptive transfer, participants were required to evaluate their own prepared slide (in contrast to the pre-prepared slide used for the training and analogue tests), and also calculate absolute leukocyte (neutrophil, lymphocyte, monocyte, eosinophil, basophil) numbers from a provided total leukocyte count with reference ranges. This latter skill requires that sufficient cells are counted, white blood cells are accurately identified, and that percentages are used to calculate the absolute cell count for each leukocyte (Weiser 2012). This particular skill was designed to be entirely unfamiliar to participants. While participants were provided with a cell counter and calculator, they received no prior instructions as to how an absolute leukocyte count might be generated. In addition, participants were asked the following yes/no interpretive questions for a total of five marks: anemic, abnormal RBC distribution, RBC regeneration, inflammatory, and neoplastic. Further outcomes were assigned for leukocyte percentages equaling 100% (score 0 or 1), and sum of leukocyte absolute counts equaling the total WBC count (score 0 or 1). These ordinal data received a total score out of seven marks. Performance in leukocyte differential counts was defined as the difference between the participant's leukocyte count (neutrophil, monocyte, lymphocyte, eosinophil, basophil) and that of a single expert reference point. This reference point was generated from a manual count of 200 cells, performed and calculated by a board certified veterinary clinical pathologist (EG). All negative differences were converted to positive values.

Upon immediate completion of the short-term assessments, participants took a survey aimed to evaluate their perception of errors, attitudes towards unfamiliar scenarios, their ability to reflect upon their own learning (metacognition), and tendency towards performance goal or learning goal orientation (Meritet 2019). Answers were scored on a five point Likert scale (1=strongly disagree to 5= strongly agree). Space for open comments was provided at the end of the survey.

For the long-term assessment (7 weeks after the short-term assessment), participants completed an adaptive task similar to that described for the short-term adaptive task with a different unknown canine blood sample. In this instance the unknown sample had an elevated white blood count representing an inflammatory leukogram. Outcome measures were the same as noted previously. Investigators were blinded to participant training method.

3.2.5 Statistical analysis

All data sets were tested for normality using the D'Agnostino and Pearson's normality test. Data were presented as mean (standard deviation, SD) if normally distributed and median (interquartile range, IQR) if non-normally distributed. To compare performance differences between normally distributed groups, Welch's t-tests were used for monolayer area and cell differential counts (Delacre 2017). Mann Whitney tests were used for ordinal data (smear quality, cell identification, application and interpretation) and other non-normally distributed data. To evaluate performance differences within individuals between two different time points (i.e. analogue pre-test to post-training tests; short term adaptive transfer and long term adaptive transfer tests), paired Student's t-tests (normally distributed data) and Wilcoxon matched-pairs signed rank tests (non-normally distributed data and ordinal data) were used. For all test outcomes, pre-analyses were performed to determine if data could be combined or stratified for subsequent hypothesis testing. Statistical significance was set at two-tailed $P \leq 0.05$. In an effort to avoid type I errors, false discovery rate (FDR) analysis was performed using a two-stage step-up Benjamini-Krieger-Yekutieli method and a desired FDR of 5% (Benjamini 1995; Chen 2010). Sample size and distribution precluded statistical analysis of the survey data.

3.3 Results

Twenty-six first year OSU-CCVM students enrolled in the study, with 13 students assigned to each group. Results of the first questionnaire handed out during the introductory session found both the median and 75th percentile at 2 (practiced creating and evaluating blood smears, but never in a clinical setting).

3.3.1 Performance in analogue pre-test and post-training test

Performance in the analogue tests are shown in **Table 3.2**. Following training, participants created slides with a significantly greater monolayer area (P = 0.0014). Smear quality scores also significantly improved (P < 0.0001), and participants correctly identified a greater number of cells (P = 0.0002).

Table 3.2. Performance in analogue tests, before (pre-test) and shortly after training (post-training). Results are shown as the mean (SD) if normally distributed and median (IQR) if non-normally distributed. Two tailed *P*-values marked with * are Wilcoxon matched-pairs signed rank test for all except monolayer area where a paired t-test was used. Two tailed *P*-values marked with ** are Mann Whitney tests (smear quality and cell identification) or Welch's t-test (monolayer area). Bolded *P*-values are those that are statistically significant after performing a false discovery rate (FDR) controlling procedure.

	Pre-test	Post-training	*P*-value*
Monolayer area (mm^2)			
All	0 (0-52.5)	69.9 (37.9)	**0.0014**
EAT	22.2 (32.8)	73.8 (39.5)	**0.0079**
EMT	36.0 (44.8)	66.0 (37.4)	**0.053**
P-value**	-	0.61	
Smear Quality (/6)			
All	3.7 (1.4)	5.3 (0.8)	**<0.0001**
EAT	4.1 (1.0)	5.3 (1.8)	**0.0078**
EMT	3.3 (1.8)	5.2 (0.9)	**0.001**
P-value**	-	1	
Cell Identification (/20)			
All	14.1 (2.6)	16.7 (1.8)	**0.0002**
EAT	13.2 (2.7)	16.3 (1.5)	**0.0098**
EMT	15 (14-16)	17.1 (2.0)	**0.02**

P-value**	-	0.26

3.3.2 Performance in short- and long-term adaptive transfer tests

Results of the adaptive transfer tests are shown in **Table 3.3**. After sustained training quiescence, cell differential counts had greater deviation from the expert reference point (median of 4418 as compared to a mean of 3972). This change was statistically significant (P = 0.0009). Over the same time period, neutrophil and eosinophil counts were more similar to the expert reference point. Monocyte and lymphocyte counts showed greater deviation from the expert reference points. Application and interpretation scores increased over this time period; this difference was statistically significant (P = 0.0026).

Table 3.3. Performance in adaptive transfer tests, shortly after training (short term) and 7 weeks after training (long term). Results are shown as the mean (SD) if normally distributed and median (IQR) if non-normally distributed. Two tailed *P*-values marked with * are Wilcoxon matched-pairs signed rank test for all except cell differential counts (EAT only), where a paired t-test was used. Two tailed *P*-values marked with ** are Mann Whitney tests with the exception of short-term differential counts, where Welch's t-test was used. Bolded *P*-values are those that are statistically significant after performing a false discovery rate (FDR) controlling procedure.

	Short Term	Long Term	*P*-value*
Cell differential counts (difference from expert reference; calculated # of cells)			
All (total difference)	3972 (2418)	6122 (4418-9208)	**0.0009**
Neutrophil	2593 (2263)	1486 (685-3667)	
Eosinophil	232 (134-235)	100 (0-525)	
Monocyte	314 (184-730)	1289 (583-2130)	
Lymphocyte	654 (523)	1949 (1243)	
EAT (total difference)	3992 (2316)	7866 (5412)	**0.015**
Neutrophil	2558 (2241)	1942 (567-6508)	
Eosinophil	235 (157-235)	100 (0-577)	
Monocyte	614 (561)	1473 (692-2248)	
Lymphocyte	594 (530)	1576 (1147)	
EMT (total difference)	3952 (2610)	5084 (4420-8498)	**0.033**
Neutrophil	2628 (2376)	1107 (677-3145)	
Eosinophil	196 (85-235)	394 (0-550)	
Monocyte	310 (140-586)	1355 (921)	
Lymphocyte	712 (529)	2322 (1267)	
P-value**	0.97	0.77	
Application and interpretation (/7)			
All	3.9 (1.4)	5.0 (1.2)	**0.0026**
EAT	3.6 (1.3)	4.8 (1.5)	**0.032**
EMT	4 (4-5)	5.2 (0.9)	**0.049**
P-value**	0.22	0.57	

3.3.3 Performance comparisons between EAT and EMT

Paired data for the analogue and adaptive transfer tests is shown in **Figures 3.3** and **3.4**, respectively. There were no significant differences between EAT and EMT for the analogue post-training test, short term adaptive transfer tests, and long term adaptive transfer tests (P all $\geq$ 0.22).

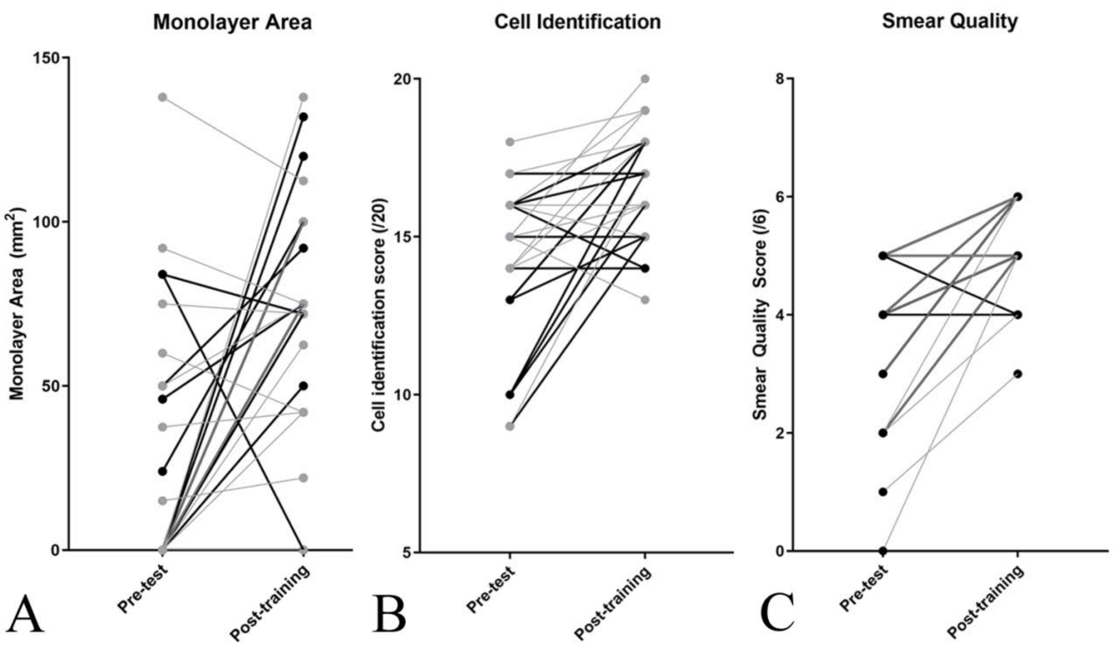

Figure 3.3. Performance changes in an analogous test, before and after training. EAT participants are shown in black and EMT participants are shown in gray. There were no significant differences between EMT and EAT at the post-training test ($P \geq$ 0.26).

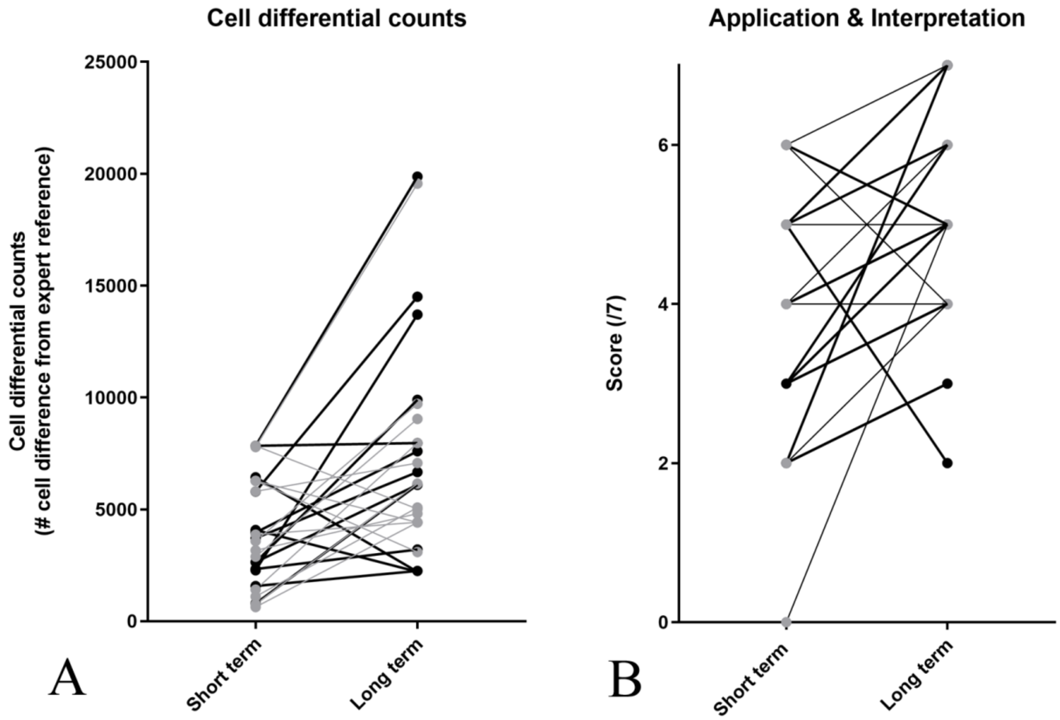

Figure 3.4. Performance changes in an adaptive transfer test at two different time points. The short-term test was performed 24-72 hours after training (short term) and the long term test was performed 7 weeks after training. EAT participants are shown in black and EMT participants are shown in gray. There were no significant differences between EMT and EAT at either the short term or long-term tests ($P \geq 0.22$).

3.3.4 Survey

Survey results are shown in **Figure 3.5**; the rationale for asking these questions has been discussed in another publication (Meritet 2019). Both groups indicated a strong preference for learning goal orientation (**Question 1**), with 83-85% of participants agreeing or strongly agreeing. Many students also indicated a preference towards performance goal orientation (**Question 2)**, although responses were more balanced. Responses to a question evaluating performance-avoid orientation (**Question 3**; tendency to evade negative judgement) were mixed. In **Question 5**, many participants (50-61%) indicated feeling unprepared for adaptive transfer, even though many indicated feeling prepared to recognize or avoid problems related to blood smear analysis (**Question 4**). Participants generally felt that mistakes were an important part of their training (**Question 6**). Many participants positively reflected on their ability to self-evaluate **(Questions 7-10**), with EMT scores slightly trending beneath that of EAT.

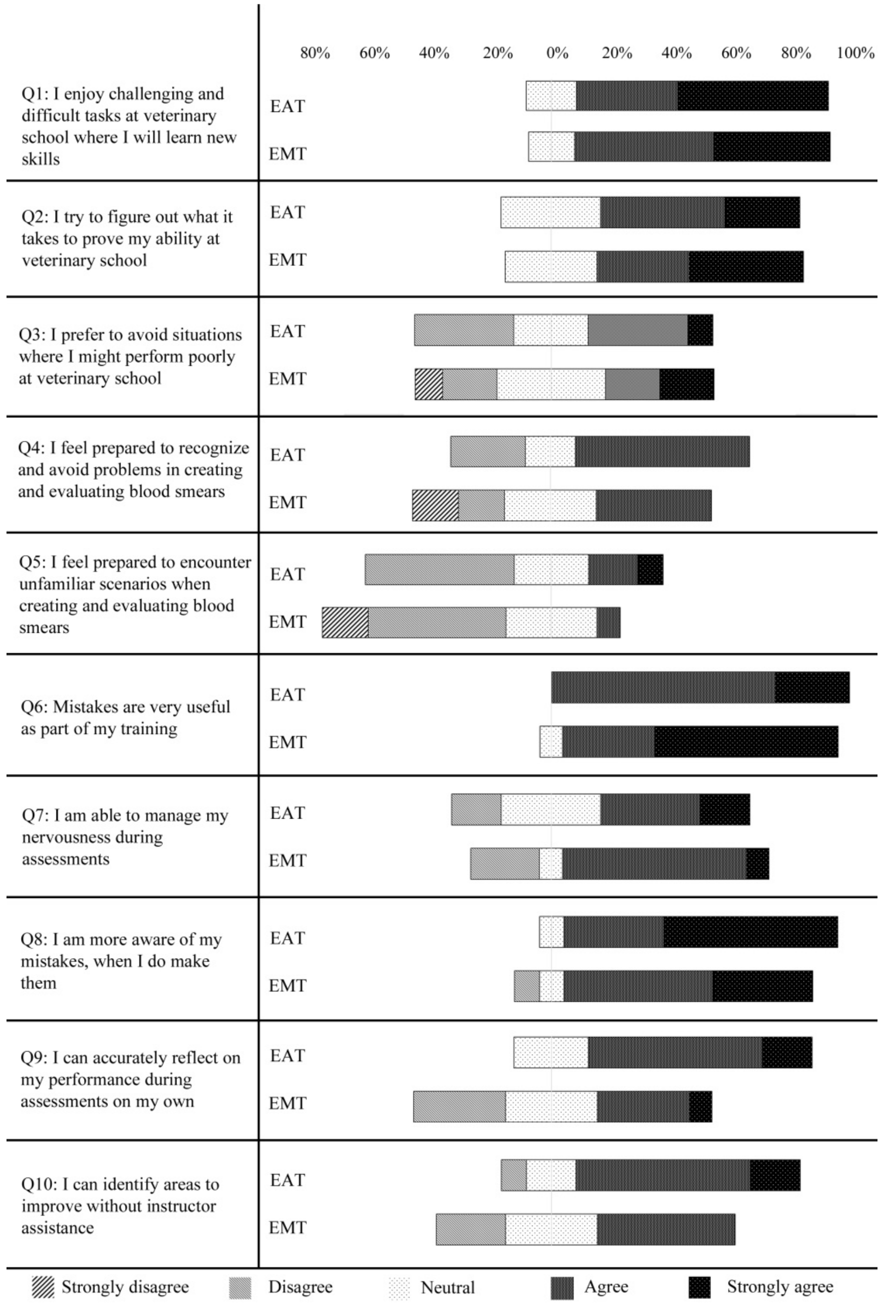

Figure 2.5. Results of 10 survey questions taken immediately after short-term testing. Data are expressed on a five point Likert scale (strongly disagree, disagree, neutral, agree, strongly agree). The Y-axis is centered on the median neutral responses, with agree/strongly agree responses on the right half and disagree/strongly disagree responses on the left half. EAT participants are shown above EMT participants.

3.4 Conclusions

Graduating veterinarians are required to apply their expertise to unfamiliar scenarios and respond to mistakes that inevitably occur as part of clinical practice. EMT has been proposed as a way of improving adaptive transfer, in part by modifying attitudes towards errors encountered during low-stakes training environments (Keith 2008). Our group has previously demonstrated the utility of EMT as it relates to a dexterous veterinary medical skill (Meritet 2019). To our knowledge, the efficacy of EMT in a veterinary task demanding a complex mixture of technique, knowledge, microscopy and cognitive function is previously undetermined. These findings may help educators seeking to encourage students to learn from and build upon their errors. Students successful in this realm are likely to benefit from an enriched educational experience that might help prepare them for adversity and life-long learning.

We hypothesized that students trained with EMT have improved performance in tasks that are structurally different from those learned during training. Adaptive transfer performances at both the short- and long-term testing windows were similar, leading us to reject our central hypothesis (P all ≥ 0.22). Seven weeks after training, we noted a statistically significant decline in the accuracy of cell differential counts among all participants ($P = 0.0009$) but an improvement in application and interpretation ($P = 0.0026$). Both EMT and EAT were responsible for these changes leading us to conclude that, in this training model, EMT trained students perform no differently after a period of training quiescence. Greater errors in cell differential counts were expected given the interruption in training and practice. However, the contrasting trend for application/interpretation might reflect some degree of prior familiarity for the long-term adaptive transfer test (i.e. calculating cell differentials and applying calculations to interpretive possibilities), even though the blood sample was unknown at both testing intervals.

Our findings differ from published studies where EMT has been associated with considerable improvements as compared to proceduralized/error-avoidance training – particularly with regards to adaptive transfer tasks (Frese 1991; Ivancic 2000; Frese 2015).

These data suggest that at a minimum, EMT performs comparably to error avoidance for the purposes of analogue performance. For all participants, training produced a significant improvement in all three analogue outcomes ($P \leq 0.0014$). Statistically significant differences were not noted between EMT and EAT at the post-training test. These data are similar to that of another study which tested EMT by a surgical knot tying model (Meritet 2019). Other non-veterinary EMT studies have also found that students trained by proceduralized methods might also perform comparably to EMT (Keith 2005; Heimbeck 2003). While the benefits of EMT might be particularly applicable to adaptive transfer (Keith 2008), positive effects relating to analogous transfer have been associated with enhanced awareness generated by errors. Such awareness might make skills more accessible for similar problems at a later time point (Ivancic 2000).

While error-avoidance and error-management force students to interact with errors very differently (Keith 2008), we noted possible differences between the testing/survey data and training observations that were not directly quantified as part of this experiment. Survey question #6 indicated that participants in both groups considered errors to be a useful part of training. Yet during the training modules themselves, there was a noticeable contrast in the way students went about training and interacted with instructors. The proceduralized group, while still engaged, was intent on quickly creating the perfect slide and repeating their practice exactly as indicated in the instructions. Many participants only sought out instructor feedback for “correct” performance. In contrast, the EMT group were more relaxed and visibly open to experimentation. Examples of this included creating slides while kneeling instead of standing, using a “pull” instead of “push” method during slide preparation, changing the order of slide staining, and varying the times immersing the slide in each stain. Such procedural experimentation almost certainly led to a greater number of errors within the EMT group. Often instructors were called over to determine whether a variation in technique might be acceptable, or simply to verify “incorrect” performance. A mechanism by which such errors might benefit learning is prediction error theory, whereby discrepancies between expectation and the actual outcome serve to enhance learning (Friston 2005). Despite encouraging errors, EMT is likely to generate a large number of self-generated correct outcomes. This active (as opposed to passive) approach to incorrect and correct strategies is likely to have a powerful effect on learning (Bertsch 2007).

Recognizing the importance of metacognition in EMT (Keith 2005, Bell 2008), training was supplemented by a module intended to reinforce reflective thinking in students. Stimulating students to reflect upon their errors is an effective strategy for ensuring sustained motivation to improve, in part by offsetting overconfidence which can be an impediment to learning (Metcalfe 2017). The scripted heuristics themselves helped focus attention upon errors, encouraging students to monitor their own performance and design corrective strategies (i.e. “try to figure out where you might have gone wrong and how you can approach it differently”)(Cannon-Bowers 1998). Survey responses to questions 8-10 suggested that both groups engaged in metacognition, with EAT participants perhaps agreeing to a slightly greater extent. During training, feedback given to both groups was limited to “correct/incorrect”, maybe compelling EMT and EAT participants to explore different approaches to the task on their own. In this way, metacognition may have been inadvertently fostered for both training scenarios.

One element of successful individual performance is the ability to apply “emotion control” – specifically the ability to self-manage performance anxiety associated with errors (Kanfer 1996). Such skill could be valuable, especially during early training, because it ensures that attention is directed towards the task as opposed to the self (Kluger 1996). Responses to question 7 indicate emotion control in both groups of participants, with modestly stronger responses for EMT. Improved adaptive transfer performances seen with EMT are in part mediated by helping participants to manage negative emotions associated with errors (Keith 2005). A reasonable proportion of participants (31-50%) reported feeling neutral or negatively towards their ability to manage nervousness during assessments. EMT might be one particularly useful strategy for helping this subset of students – particularly when encountered with tasks requiring adaptive transfer. The relationship between individual learning styles and attitudes towards errors has been discussed elsewhere (Meritet 2019)

Primary limitations of this study are the relatively small number of participants, the limited length of training, and some degree of overlap between the analogous and two adaptive transfer tasks. One difficulty with implementing EMT in this particular setting is that the impact of some errors might not have been apparent to participants. For instance, training was designed such that Diff-Quik immersion technique was not compared to microscopic staining patterns, or that smear technique could not be directly compared to the microscopic appearance of the monolayer. Corrective feedback after making errors is especially important; (Moreno 2004, Anderson 1972) as there is also literature to suggest that feedback should provide the correct answer (Pashler 2005). EMT instructions were intentionally designed such that feedback did not specify an explicit error remediation strategy, in lieu of encouraging discovery (Keith 2008).

Despite regularly repeating heuristics throughout the training, it is possible that feedback was not adequately processed by participants.

Findings presented herein are preliminary justification for helping veterinary students to manage errors generated during low-stakes training environments. These data contribute to a growing argument for utilizing EMT, or other broadly similar training strategies, as a possible framework for improving adaptive transfer and potentially navigating mistakes encountered during clinical practice. Encouraging veterinary students to contemplate their interaction with error, and leveraging this to create dynamic learning opportunities, could be one part of a multi-faceted strategy to enhance student learning and professional success.

Ch. 4 Discussion

The preceding studies describe the novel implementation of error management training (EMT) for veterinary student learning and analyze its effects on task performance compared to more traditional error-avoidance training (EAT) method outcomes. Overall, these data do not find a significant advantage of one training method over the other with regard to task performance. Restated, these findings suggest that EMT is, at the minimum, a comparable training strategy to EAT. The first study – integrating EMT into the initial instruction of surgical knot tying – did find that the overall decline in task performance in the long-term setting after a period of training acquiescence was more greatly attributed to EAT participant performance. This finding was not repeatable in the blood smear study, though is still worth mentioning as potential advocacy for the ability of EMT to improve skill/knowledge retention, particularly regarding adaptive tasks.

Other studies have found similar outcomes for EMT trainees (Dyre 2017; Gardner 2015). Possible explanations for the superior long-term performances in adaptive tasks after using EMT include the use of self-regulatory skills of metacognition and emotion control as well as dispositional attributes like goal orientation. By encouraging students to reframe their perspective of error as a learning opportunity instead of a negative reflection of competency, EMT helps reduce the anxieties and cascade of emotional thoughts that can arise around errors, allowing for enhanced learning and enduring understanding (Bell 2008). Emotion control strategies are especially useful when introduced early in training programs, the period in which trainees are most likely to encounter errors (Bell 2008). Our studies attempted to follow this suggestion by recruiting junior veterinary students who lacked formal training in the tasks introduced. While not statistically significant, the positive effect of emotion control via EMT on more junior trainees may be inferred from the survey data (**Figures 2.5 and 3.5**). Comparing the answers to question 7, a greater proportion of the 1st year students in the blood smear study felt they could effectively manage their nervousness during assessments, with approximately 80% agreement compared to 60% agreement amongst the 2nd year students in the knot tying study.

In spite of their perceived emotion control, the EMT participants of the blood smear study did not perform better in adaptive tasks than EAT trainees. As previously mentioned, the lack of significant performance differences between the two groups may be due to an unintended lapse in training to portray clinical consequences of errors for the tasks being presented. Contextual support has been found to improve individual's performance by encouraging them to engage in higher cognitive capacities (King 2004). The intention of creating low-stakes tasks was meant to alleviate anxieties that surround errors. However, this was not intended to ignore the potential clinical implications of errors. Placing instruction within context for students helps the learning process by enabling them to make meaningful, relevant connections between their education and their life and/or future career (Merriam 2013a).

This concept of meaning-making is prevalent in pedagogy and psychology in that both intellectual and personal growth involves the interaction of an individual's prior knowledge and pre-existing assumptions with their current experiences (King 2004, Baviskar 2009). Hindrances to growth or learning can occur when there is discrepancy between prior knowledge/assumptions and current experience; some refer to this incongruent state as "cognitive dissonance." (Baviskar 2009, Hartle 2012). Veterinary students are enrolled in higher education and constitute a particular cohort of adult learners. As adult learners, veterinary students enter their veterinary education with pre-established sets of beliefs, assumptions and knowledge. It follows that there may be an opportunity, if not multiple, for a veterinary student to experience cognitive dissonance during their veterinary education. Errors can serve as a source of cognitive dissonance. How the veterinary student – or any learner – reacts towards cognitive dissonance, towards errors, may be linked to motivation which can be guided by contextual learning (Baviskar 2009, Hartle 2012). Motivational theory links motivation with an individual's values and expectancies (Ambrose 2010b). Framing instructional assessments within a relevant context for learners can help create both intrinsic and incremental value (Ambrose 2010b) While trainees in the knot tying study were informed of the various purposes of surgical knots (e.g. closing wounds, ligating vessels), the consequences of ineffective knots was not relayed (e.g. infection, exsanguination). Similarly, participants of the blood smear study were told of the importance of a monolayer, but the clinical impact of its absence – namely, non- or mis-diagnosis – was not emphasized during training. It is possible this additional context could have led to increased value, improved learning and task performance, and is worth consideration in future investigations.

Relating back to motivational theory, the second factor affecting motivation is expectancy which reflects an individual's belief they will (or will not) succeed in a task (Ambrose 2010b). As such, expectancies can be positive or negative, shaping an individual's behavior and approach towards a task. In addition, expectancies can be altered by an individual's self-efficacy and by the learning environment (Ambrose 2010b). Although not an explicit objective of EMT studies, EMT inherently cultivates positive expectancies in learners by creating supportive environments and increasing self-efficacy. The former is largely achieved by providing encouraging and informative feedback while the latter is obtained through altered causal attribution of errors.

Improved self-efficacy can be attained through EMT by modifying an individual's causal attribute towards errors. For any outcome, there are two variables that are considered when an individual attributes cause: stability and locus of causality (Van Dyck 2010). The cause for an outcome is stable or unstable and internal or external. The locus of causality is linked with self-efficacy in that internal factors are contained within oneself and, so, are manageable while external factors are (at least perceived to be) outside an individual's control. It is plausible to consider that learners with a sense of high self-efficacy likely attribute outcomes to internal influences. Van Dyck et al., found that learners who attributed errors as internal and unstable often adapt an error mastery approach compared to an error avoidant approach. While not directly studied in this body of research, error-management may help encourage learners towards an error mastery approach by shifting their causal attributions of errors towards being internal and unstable (i.e. errors are within their power and are able to be changed), thereby increasing self-efficacy. Self-efficacy may also be increased by utilizing low-stakes assessments (Ambrose 2010b). With that in mind, low-stakes assessments can help stir motivation through positive expectancies. Moreover, by promoting nontechnical skills like emotion control and metacognition, EMT helps students cope and handle errors more effectively, thus increasing their self-efficacy by enabling them to overcome obstacles in a productive manner.

While both studies contained survey questions relating to metacognition, only the blood smear study implemented a learning tool meant to enhance metacognition. Participants were asked to watch videos of task completion done either correctly (EAT training session) or incorrectly (EMT training session). EAT trainees were asked to identify correct technique while EMT trainees were asked to identify incorrect technique as well as comment on how they would have addressed the errors shown. In spite of this additional training module, trainee engagement in metacognition did not greatly differ between the groups according to survey data (**Figure 3.5**, questions 8-10). In fact, a greater proportion of EAT participants agreed they engaged in metacognition than did EMT trainees, albeit not to significance. Possible reasons for this unanticipated finding have been previously addressed in this document. When comparing the survey data between studies (**Figures 2.5 and 3.5**), there is a general right-skewed (i.e. greater agreement) trend for questions 8-10 relating to metacognition in the knot tying study compared to the blood smear study, in spite of the former not having the benefit of a metacognition module. This may in part be a result of the different training levels of participants. Having an additional year of veterinary medical education, the knot tying participants likely have more experience in higher-order thinking than their junior colleagues. Studies have shown that more experienced individuals exhibit a greater ability to think abstractly and engage in metacognition (Etelapelto 1993; Hinds 2001; McKeithen 1981). Evaluating the effects of EMT on veterinary students at varying training levels was beyond the scope of these studies but may be worth future investigation.

An alternative interpretation of the survey data relating to metacognition is that junior veterinary students perceive themselves to have high-efficacy. Yet, their performances in adaptive tasks do not necessarily demonstrate that fact. In regard to the surgical knot study, the median knot bursting pressure ranged from 10-15psi while an expertly-tied knot held at 300psi (see page 33). Again, this may be a consequence of not providing sufficient context. It is possible that long-term performances may have improved if participants had received the results of their short-term performance. Feedback was only provided in the training sessions and not during/after testing phases. This was also intentional as these tasks were meant to be low-stakes and were not assigned a “grade” that would suggest competency or lack thereof. In retrospect, if students were aware of their short-term performance and were told that expert performance equates to a leaking pressure of 300psi, they may have reflected more on their previous performance prior to completing the long-term adaptive task. Productive self-reflection is predicated on receiving informative feedback (Ambrose 2010a, Merriam 2013b). Feedback can take many forms; it can be dictated or sensed (e.g. auditory prompts), human or instrument-generated. Interactive learning through "serious games" is gaining favor in the medical field and is in experimental-stages in veterinary curricula (Graafland 2012, Franson 2010, Parkes 2009). These computer-based simulations provide multiple advantages over traditional training methods, including real-time feedback. As opposed to being evaluated at the end of the task, trainees are evaluated during the task and receive feedback throughout task completion. We attempted to provide real-time feedback during training sessions, though the low ratio of trainee: co-investigator rendered that difficult. The addition of equipment that sensed errors in real-time may have enhanced trainee learning and improved task performance in our studies. Inclusion of such equipment may be beneficial in future error-management research.

A third factor influencing the effect of EMT – goal orientation – was not thoroughly investigated in these studies. Several of the survey questions (1-3) were intended to gather information regarding participant goal orientation. For both studies, there were no overt differences in distribution of learning vs. performance goal oriented participants in the two training groups (not tested statistically). This is ideal since other studies found EMT to impose greater benefit on those of a learning goal orientation and placing more learning goal oriented individuals in one of the groups could have confounded results.

Interestingly, across both studies and regardless of training group, participants often identified as being both learning and performance-prove goal oriented with mixed responses towards performance-avoid goal orientation. The influence of such “dual” goal orientation on EMT efficacy is unknown and may further explain the lack of significant differences between EAT and EMT performances in our studies. Research by Button et al. discovered that goal orientation is not static; it is fluid with both demographic and situational factors affecting its direction in a given learning task (Button 1996). This inquiry found that an individual can adopt a learning goal orientation if there is intrinsic or incremental value in a learning task. While goal orientation is dynamic, individuals typically approach a task with one orientation or another, not both simultaneously. In general, veterinarians are perceived as high-achievers and have been found to most closely align with a performance goal orientation, slightly favoring a performance-avoid goal orientation (Arnold 2012). However, the intrinsic and incremental value veterinarians and veterinary students are likely to find in learning new clinically applicable tasks could shift them towards a learning goal orientation, helping to explain the survey information on goal orientation for these present studies (Button 1996). Regardless of how errors are broached during training, individuals with a learning goal orientation tend to adopt an error mastery approach (Van Dyck 2010). This reasoning may account for the comparable performances between the EAT and EMT participants noted in these investigations. Participants possibly adopted a learning goal orientation and, so, an error mastery approach, regardless of the training they received. Future studies may wish to include a cohort with less intrinsic value towards the skills being taught (e.g. graduate students in unrelated fields). The effect of EMT on the performance of this third cohort may help further define the role of goal orientation and motivation in error-management training as it pertains to veterinary students.

There is a proposed theory that connects motivation with goal orientation called motivation action theory (DeShon 2005. This theory postulates that actions are goal-driven; they are directed towards the attainment of a desired state, expanding goals to include needs, values and principles. The motivation to engage in a certain action depends on the goal as defined above. With that in mind, it reasons that goal orientation would be situationally specific and dynamic, further lending support to Button et al findings (1996). When designing studies to investigate goal orientation, the investigator should determine if they wish to either manipulate or measure goal orientation. Our survey questions 1-3 were intended to measure goal orientation, but we failed to realize that error-management could be a means to manipulate goal orientation. This discrepancy in our intention and understanding may have, in hindsight, failed to capture an important additional benefit of error-management: a method to push learners towards a learning goal orientation.

Lastly, this publication on motivation action theory as an explanation of goal orientation touches upon employee well-being. Most studies regarding goal orientation heretofore have focused their investigation on the relationship of goal orientation with trainee/employee performance and not on trainee/employee development. Similarly, our study focused on the effect of EMT on trainee performance, though its effect on trainee well-being is a noteworthy factor to consider. As previously mentioned, some studies measuring goal orientation found that veterinarians more often align with a performance-avoid goal orientation (Arnold 2012). This type of orientation originates when a certain task threatens personal goals such as self-esteem and self-efficacy, leading to debilitating emotions (e.g. anxiety and stress). In an effort to evade such emotions, while also preserving self-esteem and self-efficacy, a performance-avoid goal orientation is adopted (DeShon 2005). At the organizational level, performance-avoid goal orientation is less desirable in that it could potentially decrease employee productivity. DeShon et al. proposes organizations adopt strategies that help both the organization and individual achieve their respective goals. By helping students learn both technical and nontechnical skills while attempting to reduce and manage errors, error-management training provides an ideal training strategy that helps reach both institutional and individual goals.

In summary, what advantage could EMT offer in the veterinary educational setting? From a purely pedagogical standpoint, EMT encourages active learning, placing the student at the center of the learning process, through exploratory learning (Bell 2008). We are currently experiencing an evolution in higher education, one in which the learner becomes an active participant, trending away from traditional training strategies that embrace highly structured environments and limit the learner’s control and creativity (Ivancic 1995). The latter approach may be effective for routine expertise and helps performance in analogous tasks, but it does not facilitate the transfer of knowledge to novel tasks or promote discovery and innovation (Frese 1994).

The role of veterinary institutions is to prepare future generations of veterinarians for a challenging, ever-changing profession. To succeed in this role, educators need to encourage their students to become self-directed, lifelong learners. Active learning employs self-regulatory processes considered vital to developing complex skillsets and effecting adaptive transfer of knowledge (Ivancic 2000). Current educational research seeks out interventions that support active learning and self-regulatory practices to promote enduring understanding, enhance performance and facilitate adaptive knowledge/skill transfer (Kozlowski 2001; Frese 1991; Heimbeck 2003; Keith 2005). In veterinary medicine, problem-based learning and simulation learning are becoming more prevalent though require curricular reform and/or substantial costs (Lane 2008, Graafland 2012). EMT could provide an alternate, inexpensive and feasible intervention in veterinary medicine. The implementation of EMT would require instructor training. Such training is expected to have minimal expense with regard to time and finances, as well as minor curricular change, namely by the addition of modules that encourage self-reflection. Through such modules, EMT can provide trainees with self-regulatory skills sorely needed in the veterinary profession. Both metacognition and emotion control, as fostered through EMT, enable learners to adapt to unfamiliar situations and regulate their emotions so their cognitive resources are centered on the task and not the self. Such skills can have enduring positive effects, reducing performance-induced anxieties and improving productivity as well as overall wellbeing.

This body of work outlines two introductory investigations on the effect of EMT for veterinary students learning relevant clinical skills. Our findings suggest that EMT is, at minimum, comparable to traditional error-avoidance approaches with regard to task performance, warranting further investigation. We argue that the nontechnical skills fostered through EMT merit consideration of its integration into veterinary educational systems.

Bibliography

Ambrose SA, Bridges MW, DiPietro M, Norman MK. (Eds.) How do students become self-directed learners? In *How learning works: Seven research-based principles for smart teaching.* San Francisco: Jossey-Bass. 2010a:188-216.

Ambrose SA, Bridges MW, DiPietro M, Norman MK. (Eds.) What factors motivate students? In *How learning works: Seven research-based principles for smart teaching*. San Franciso: Jossey-Bass. 2010b:66-90.

Anderson RC, Andre T, Kulhavy RW. Conditions under Which Feedback Facilitates Learning from Programmed Lessons. *J Educ Psychol*. 1972;63(3):186-188.

Arnold K. *Effects of goal orientation profile types on organizational outcomes in veterinarians.* [dissertation]. College of science and health theses and dissertations. DePaul University; 2012.

Bain BJ. Current concepts: Diagnosis from the blood smear. *New Engl J Med.* 2005;; 353(5):498-507.

Baviskar SN, Hartle RT, Whitney T. Essential criteria to characterize constructivist teaching: Derived from a review of the literature and applied to five constructivist-teaching method articles. Int J Sci Educ. 2009;31:541-550.

Bell BS, Kozlowski SW. Active learning: effects of core training design elements on self-regulatory processes, learning, and adaptability. *J Appl Psychol.* 2008;93(2):296-316.

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B.* 1995;57(1):289-300.

Bertsch S, Pesta BJ, Wiscott R, McDaniel MA. The generation effect: A meta-analytic review. *Mem Cognition.* 2007;35(2):201-10.

Boothe HW. Chapter 17: Instrument and Tissue Handling Techniques. In: Johnston SA, Tobias KM, eds. *Veterinary Surgery: Small Animal* 2nd ed. St. Louis, MO: Elsevier; 2018.

Buljac- Samardžić M, van Woerkom M, Paauwe J. Team safety and innovation by learning from errors in long-term care settings. *Health Care Management Review.* 2012;37(3):280–291.

Button SB, Mathieu JE, Zajac DM. Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes.* 1996; 67(1):26-48.

Cannon-Bowers JA, Rhodenizer L, Salas E, Bowers CA. A framework for understanding pre-practice conditions and their impact on learning. *Pers Psychol.* 1998;51(2):291-320.

Chillarege KA, Nordstrom CR, Williams KB. Learning from our mistakes: Error management training for mature learners. *J Bus Psychol.* 2003;17(3):369-85.

Chen JJ, Roberson PK, Schell MJ. The False Discovery Rate: A Key Concept in Large-Scale Genetic Studies. *Cancer Control.* 2010;17(1):58-62.

Debowski S, Wood RE, Bandura A. Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology.* 2001;86(6):1129-41.

Delacre M, Lakens D, Leys C. Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test. *Int Rev Soc Psychol.* 2017;30(1):92-101.

DeShon RP, Gillespie JZ. A motivated action theory account of goal orientation. J Appl Pyschol. 2005; 90(6):1096-1127.

Dweck, C. S. Motivational processes affecting learning. *American Psychologist.* 1986;41(10): 1040–1048.

Dweck CS, Leggett EL. A Social-Cognitive Approach to Motivation and Personality. *Psychol Review*. 1988;95(2):256-273.

Dyre L, Tabor A, Ringsted C, Tolsgaard MG. Imperfect practice makes perfect: error management training improves transfer of learning. *Medical Education.* 2017;51(2):196-206.

Edmondson, AC. Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *J Applied Behavioral Sci*. 1996;32:5–28,.

Elliott ES, Dweck CS. Goals: an Approach to Motivation and Achievement. *J Pers Soc Psychol.* 1988;54(1):5-12.

Ellis S, Davidi I. After-event reviews: drawing lessons from successful and failed experience. *J Appl Psychol.* 2005;90(5):857-71.

Etelapelto A. Metacognition and the expertise of computer program comprehension. Scandinavian Journal of Educational Research. 1993;57:243-254.

Fransson B, Ragle C. Assessment of laparoscopic skills before and after simulation training with a canine abdominal model. J Amer Vet Med Assoc. 2010;236(10):1079-1084.

Frese MB, F.C.; Heinbokel, T.; Mooser, C.; Schleiffenbaum, E.; Thiemann, P. Errors in computer training skills: On the positive function of errors. *Human-Computer Interaction.* 1991;6:77-93.

Frese M, Zapf D. Action as the core of work psychology: A German approach. In: Triandis HC, Dunette MD, Hough LM, eds. *Handbook of industrial and organizational psychology*. Palo Alto, CA: Consulting Psychologists Press.; 1994.

Frese M, Keith N. Action errors, error management, and learning in organizations. *Annu Rev Psychol.* 2015;66:661-87.

Friston KJ. A theory of cortical responses. *Philos T R Soc B.* 2005;360(1456):815-36.

Gardner A, Rich M. Error management training and simulation education. *Clin Teach*. 2014;11(7):537-40.

Gardner AK, Abdelfattah MD, Wiersch J, Ahmed RA & Willis RE. Embracing errors in simulation-based training: The effect of error training on retention and transfer of central venous catheter skills. *J Surg Educ.* 2015;72(6):158-162.

Graafland M. Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Brit J Surg. 2012;99:1322-1330.

Gully SM, Payne SC, Koles KLK, Whiteman JAK. The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective. *J Appl Psychol*. 2002;87(1):143-55.

Hartle RT, Baviskar S, Smith R. A field guide to constructivism in the college science classroom: Four essential criteria and a guide to their usage. Bioscene. 2012:38(2):31-35.

Heimbeck D, Frese M, Sonnentag S, Keith N. Integrating errors into the training process: The function of error management instructions and the role of goal orientation. *Pers Psychol.* 2003;56(2):333-61.

Hinds PJ, Patterson M, Pfeffer J. Bothered by abstraction: The effect of expertise on knowledge transfer and subsequent novice performance. Journal of Applied Psychology. 2001;86:1232-1243.

Ivancic K, Hesketh, B. Making the best of errors during training. *Training Research Journal*. 1995;1:103–125.

Ivancic K, Hesketh B. Learning from errors in a driving simulation: effects on driving skill and self-confidence. *Ergonomics.* 2000;43(12):1966-84.

Johnson S. *Where good ideas come from: the natural history of innovation*. New York, NY: Riverhead books; 2011.

Kanfer R, Ackerman PL, Heggestad ED. Motivational skills & self-regulation for learning: A trait perspective. *Learn Individ Differ.* 1996;8(3):185-209.

Keith N, Frese M. Self-regulation in error management training: emotion control and metacognition as mediators of performance effects. *J Appl Psychol.* 2005;90(4):677-679.

Keith N, Frese M. Effectiveness of error management training: a meta-analysis. *J Appl Psychol.* 2008;93(1):59-69.

King A. Effects of Training in Strategic Questioning on Childrens Problem-Solving Performance. *J Educ Psychol.* 1991;3(3):307-17.

King PM, Kitchener KS. Reflective judgement: Theory and research on the development of epistemic assumptions through adulthood. Educ Psychol. 2004;39:5-18.

Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol Bull.* 1996;119(2):254-84.

Kozlowski SWJ, Gully SM, Brown KG, Salas E, Smith EM, Nason ER. Effects of training goals and goal orientation traits on multidimensional training outcomes and performance adaptability. *Organizational Behavior Human Decision Processes.* 2001;85(1):1-31.

Kuhn TS. *The Structure of Scientific Revolutions.* 4th ed, Chicago, IL:Univ of Chicago Press; 2012

Lane EA. Problem-based learning in veterinary education. J Vet Med Educ. 2008;35(4):631-636.

Lawton R, McEachan RRC, Giles SJ, Sirriyeh R, Watt IS, Writght J. Development of an evidence-based framework of factors contributing to patient safety incidents in hospital settings: a systematic review. *BMJ Quality & Safety.* 2012;21:369–380.

Mangels JA, Butterfield B, Lamb J, Good C, Dweck CS. Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Soc Cogn Affect Neurosci*. 2006;1(2):75-86.

McKeithen KB, Reitman JS, Rueter HH, Hirtle SC. Knowledge organization and skill differences in computer programmers. Cognitive Psychology. 1981;13:307-325.

Mellanby RJ, Herrtage ME. Survey of mistakes made by recent veterinary graduates. *Vet Rec*. 2004;155(24):761-5.

Meritet DM, Townsend KL, Gorman ME, Chappell P, Kelly L, Russell DS. Investigating the Effects of Error Management Training on Veterinary Students Learning to Tie Surgical Knots. 2019.*(Submitted to J Vet Med Educ – in final stages of review)*

Merriam SB, Bierema LL (Eds.) Experience and Learning. In *Adult learning: Linking theory and practice* San Franciso: Jossey-Bass. 2013a:104-126.

Merriam SB, Bierema LL (Eds.) Self directed learning. In *Adult learning: Linking theory and practice* San Franciso: Jossey-Bass. 2013b:61-81.

Metcalfe J. Learning from Errors. *Annu Rev Psychol.* 2017;68:465-89.

Moreno R. Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instr Sci.* 2004;32(1-2):99-113.

Muraven M, Baumeister RF. Self-regulation and depletion of limited resources: does self-control resemble a muscle? *Psychol Review.* 2000;126(2):247-259.

Nordstrom CR, Wendland D, Williams KB. "To err is human": An examination of the effectiveness of error management training. *J Bus Psychol*. 1998;12(3):269-82.

Oxtoby C, Ferguson E, White K, Mossop L. We need to talk about error: Causes and type sof error in veterinary practice.*Vet Rec*. 2015.

Parkes R, Forrest N, Baillie S. A mixed reality simulator for reline abdominal palpation training in veterinary medicine. Stud Health Technol Inform. 2009;142:244-246.

Pashler H, Cepeda NJ, Wixted JT, Rohrer D. When does feedback facilitate learning of words? *J Exp Psychol Learn.* 2005;31(1):3-8.

Richards JM, Gross JJ. Emotion regulation and memory: the cognitive costs of keeping one's cool. *J Personality Soc Psychol*. 2000;79(3):410-424.

Shattil SJ. A (blood) smear campaign. *Blood*. 2003;101(7):2453.

Sitkin, SB. Learning through failure: The strategy of small losses. In M. Cohen & U. Sproull (Eds.), *Organizational learning.* Thousand Oaks, CA: Sage; 1996:541-577

Taylor-Adams S, Vincent, C, Street, P. Systems analysis of clinical incidents: the London protocol. *Clinical Risk.* 2004;10:211–220.

van der Linden DS, Frese M, van Dyck C. Exploration strategies, performance, and error consequences when learning a complex computer task. *Behaviour and Information Technology*. 2001;20(3):189-98.

van der Schaaf T, Habraken, M. PRISMA-Medical: a Brief Description. Eindhoven, The Netherlands: Eindhoven University of Technology. 2005.

van Dyck C, Van Hooft E, De Gilder D, Liesveld L. Proximal antecedents and correlates of adopted error approach: a self-regulatory perspective. *J Soc Psychol.* 2010;150(5):428-51.

van Vuuren W, Shea CE, van der Schaaf TW. The development of an incident analysis tool for the medical field. Eindhoven: Technische Universiteit Eindhoven. 1997.

Weiser G. Hematologic Techniques. In: V*eterinary Hematology and Clinical Chemistry*. 2nd ed. Ed Thrall MA, Weiser G, Allison R, Campbell T. Wiley-Blackwell; 2012.

Weiss DJ, Tvedten H. The Complete Blood Count, Bone Marrow Examination, and Blood Banking. In: Willard MD, Tvedten H, editors. *Small Animal Clinical Diagnosis by Laboratory Methods*. 5th ed. St. Louis, MO: Elsevier Saunders; 2012:26-51.

Zakay D, Ellis S, Shevalsky M. Outcome value and early warning indications as determinants of willingness to learn from experience. *Exp Psychol*. 2004;51(2):150-7.

Zapf D, Maier GW, Rappensperger G, Irmer C. Error-Detection, Task Characteristics, and Some Consequences for Software-Design. *Appl Psychol-Int Rev*. 1994;43(4):499-520.