The impact of model fidelity on acquisition of abdominal incision closure skills in novice veterinary students

Williamson J1, Brisson B2, Skowron C1, Farrell R3, Anderson S1, Spangler D1
1Lincoln Memorial University
2Ontario Veterinary College
3Ross University School of Veterinary Medicine

Purpose
To evaluate a lower-fidelity (LF) model made of foam and fabric and a higher-fidelity (HF) model made of silicone, for teaching novice veterinary students to perform abdominal incision closure.

Introduction
Surgical skills are learned through deliberate practice.1 Simulation can be used to increase student competency prior to live animal surgery.

Research suggests that the fidelity, or realism, of a model should increase as a trainee’s experience level rises.2-4 Third-year veterinary students with previous training in abdominal incision closure achieved higher scores on the task in live surgery after practicing on a high-fidelity silicone model rather than on a low-fidelity foam and fabric model.5 The value of model fidelity in training novice veterinary students to perform abdominal incision closure remains unclear.

Materials & Methods
Expert Review: Experienced veterinarians (n=10) were surveyed about their opinions of the models.

Learning Outcomes: Third-year graduate-level veterinary students (n=38) who had never performed abdominal closure on a live animal were randomized to lower-fidelity (LF) and higher-fidelity (HF) groups. Students were recorded performing a 3-layer abdominal incision closure on a canine cadaver to establish pre-training scores. They participated in four 3-hour teaching sessions using their model. Students were again recorded performing the task on a cadaver to obtain post-training scores. Blinded raters at three veterinary schools scored the recordings using a task-specific rubric.

Student Review: Student participants were surveyed to obtain their opinions of the models.

Post-training Raw Scores

Discussion
Even for novice students, increasing model fidelity improved learning outcomes for abdominal incision closure in this study. This goes against what the literature suggests. Why?

• Higher-fidelity (HF) group students may have practiced more than lower fidelity (LF) group students, as seen in another study;6 or
• During the study students may have moved past being novices, when gains from LF and HF would have been equivalent, into intermediate skill level, when HF students gained more; or
• For this specific task, fidelity may matter even for a novice.

These possibilities could have been evaluated by assessing skills more frequently and having students report practice hours. Other limitations include the rubric’s inability to differentiate global flaws (e.g. suture handling) from checklist item flaws. Using both a checklist and global rating scale may have better clarified where students’ skill differed. Tissue layers and spacing were assessed only via recording, not measured on the cadaver at the time of suturing; thus we evaluated the reliability of the raters’ scores but not their accuracy which may have impacted the results.

Veterinarians thought both models would be helpful for teaching, but students disagreed. However, nearly all students had better post-training scores, suggesting that practice on models improved skill level regardless of student perception. Experts may ultimately be better suited than novices to determine the educational value of a model.

Selecting the best model for teaching is a multifactorial decision. One must consider ease of use, availability, initial and replacement costs, durability, student and educator acceptance, and the specific learning objectives to be met.

Fidelity increases cost. More research is needed to understand when educational outcomes justify the higher cost. Necessary fidelity may vary by complexity of the task, experience of the student, and the instructor’s educational settings.

References

Expert Review: Veterinarians believed both the lower-fidelity (LF) and higher-fidelity (HF) models were suitable for training and assessment (median 'agree'), 5-point Likert scale, LF and HF) and found them easy to use (median 'agree', LF and HF).

Rubric Scores: The internal consistency of scores was 0.82. A quarter of the videos (19) were scored by two raters; inter-rater reliability via intraclass correlation was 0.64.

Learning Outcomes: The groups had identical mean pre-test scores (11.0). After training the LF group had lower post-test scores (M=15.1) than the HF group (M=18.9, p=0.02).

Training effect score, or post-training score minus pre-training score, was positive for 78% of LF and 95% of HF students (p=0.12). Training effect scores were higher for the HF group (M=7.9) than for the LF group (M=4.1, p=0.04).

Student Review: When asked if their models were good for learning the task, students in both groups reported that they were not (median 'disagree', LF and HF).

The higher-fidelity model

The lower-fidelity model