

COMPARISON OF REAL VERSUS VIRTUAL SURGICAL TRAINING MODELS

Cao, C.G.L., Waxberg, S., Park, B., Schwartzberg*, S.

Department of Mechanical Engineering, Tufts University, Medford, MA

*Department of Surgery, Tufts University School of Medicine, Boston, MA

The purpose of this study was to compare the MIST-VR (Minimally Invasive Surgical Trainer-Virtual Reality) with a real (physical) surgical skills training system, modelled after the MISTELS (McGill Inanimate System for Training and Evaluation of Laparoscopic Skills), in order to assess their relative potential effectiveness as training models. A controlled experiment was conducted with 10 medical students/interns, 13 residents (PGY 2/3/4), and 2 fellows in general surgery and gynaecology. A mixed design was used to compare their motor skill performance on various basic laparoscopic tasks. Contrary to earlier reports (Gallagher et al, 2001), our analyses of variance showed that the MIST-VR system was not able to differentiate skill levels of the subjects based on scores, errors, time, or economy of motion. In particular, the task of manipulation diathermy did not reproduce reported results. The only parameter to suggest superior performance as a function of increasing experience was the decreasing variance in the performance measures. On the other hand, the physical system was able to differentiate experience levels based on performance scores ($F=3.49$, $df=5$, $p=0.02$). Post-hoc analysis showed that some tasks were more sensitive than others. For example, in the peg transfer task, the medical students/interns, and PGY2 were not significantly different from one another; the PGY3, PGY4, PGY5, and fellows performed equally well, but significantly better than the medical students/interns, and PGY2 residents. This could be due to a ceiling effect as a result of the low task difficulty. These results suggest that further investigation is needed before adopting either system as a standard evaluation tool in surgical competency, and in particular, a virtual training system. The performance metric used may be inadequate to assess skill acquisition. In addition, the simulated tasks, while useful for training purposes, may not be appropriate for performance evaluation.