Enhanced Learning of Manipulation Techniques using Force-Sensing Table Technology (FSTT)

John Triano, Dominic Giuliano, Marion McGregor, Loretta Howard, Canadian Memorial Chiropractic College
# Table of Contents

Executive Summary ................................................................................................................. 4  
Introduction ................................................................................................................................. 7  
  Educational Theory in Health Care Neuromotor Skills Development ........................................... 8  
  Manipulation Skills ...................................................................................................................... 9  
  Teaching and Learning Gaps for Manipulation Training ............................................................... 11  
Improving Teaching and Learning of Bimanual Skills .................................................................... 13  
Research Questions ....................................................................................................................... 13  
Research Methods ....................................................................................................................... 13  
  FSTT Intervention and Evaluation ................................................................................................. 14  
  Learner Reflection ......................................................................................................................... 16  
  Conceptions of Confidence and Competence ............................................................................... 16  
  Data Analysis ................................................................................................................................. 17  
Limitations of the Study ............................................................................................................... 18  
Findings ........................................................................................................................................ 18  
  Performance Measure Stability ..................................................................................................... 18  
  Short-Term Gains from FSTT ....................................................................................................... 19  
  Retention of Gains ......................................................................................................................... 21  
  Reflection ..................................................................................................................................... 22  
  Intentional Modulation of Force .................................................................................................... 23  
  Conceptualization and Perceptions of Skill Development ............................................................ 25  
Conclusions .................................................................................................................................. 29  
References ..................................................................................................................................... 31
List of Common Abbreviations and Special Terms

- Incisure/Incisural: a sudden and brief downward dip in a line graph
- Force-Time profile: a graphical trace displaying a change in force over time
- Formative assessment: assessments that provide feedback to inform student progress
- FSTT: force-sensing table technology
- KR: knowledge of results, a form of feedback to learners
- Newtons (N): a measure of force. 1 N = 0.448 pounds
- Newtons/sec: rate of change of force or “speed” of force
- se: standard error
- Sigmoidal: “S” shaped
- Summative assessment: evaluation of learning that has occurred over time

List of Figures

Figure 1: Force-Time Profile .................................................................................................. 10
Figure 2: Project Timeline and Assessment Intervals ............................................................ 14
Figure 3: FSTT Coaching ....................................................................................................... 15
Figure 4: Force Amplitudes ................................................................................................. 20
Figure 5: Rise in Force ........................................................................................................ 21
Figure 6: Force Modulation ................................................................................................. 25
Figure 7: Perceptions without use of FSTT ........................................................................ 26
Figure 8: Perceptions with use of FSTT .............................................................................. 27

List of Tables

Table 1: Performance Stability over the Two Weeks from Baseline to the Pre-Session 1 Assessment .... 19
Table 2: Change in Performance Parameters from Participation in Session 1 .............................. 20
Table 3: Retention of Gains Represented as the Difference Scores from Post-S1 to Pre-S2 Assessments... 22
Table 4: Gains in Skill Parameters from Reflective Practice ....................................................... 23
Table 5: Mean Gains in Force Modulation ................................................................................ 24
Table 6: Focus Scores from Text Analysis on Competence and Confidence ............................... 28
Executive Summary

In July 2011, the Higher Education Quality Council of Ontario (HEQCO) issued a Request for Proposals that focused on the innovative use of technology in the classroom. The goal was to provide funding to institutions to allow them to evaluate the effectiveness of pedagogical practices that aim to enhance the quality of student learning through the introduction and integration of new technologies. Based on a novel implementation of technology within a new skills simulation laboratory, the Canadian Memorial Chiropractic College (CMCC) submitted a successful application that allowed it to evaluate the system as a means of assessing manual skills development.

CMCC was founded as a professional college in 1945 to provide the education and training of chiropractors. The four-year program was initially accredited through the Council of Chiropractic Education Canada, a status that continues through the present. In 2005, CMCC became the first private college to be given approval to award a second-entry undergraduate degree by assent from the Minister of Training, Colleges and Universities, under recommendation from the Postsecondary Education Quality Assessment Board of Ontario (PEQAB). In 2011, that consent to award the degree of Doctor of Chiropractic (DC) was renewed and extended for a ten-year period.

Integral to the administration of care to patients by Doctors of Chiropractic is the safe and effective use of complex, bimanual treatment procedures that can be collectively described as manual manipulation therapy. As a mode of treatment for complaints from the spine, neck and extremity joints, manipulative therapy spans a broad spectrum of maneuvers that have been shown to be useful to the public in managing pain from strain/sprain injuries, degenerative arthritis, disc, shoulder and knee problems, as well as other conditions (Bronfort, Haas, Evans, Leininger & Triano, 2010). Each maneuver must be competently administered in the context of the patient’s condition, with consideration given to how chronic or severe the problem is and whether there are other diagnoses that would require modification of the treatment maneuvers during their application.

Through the auspices of the Knowledge Infrastructure Program (KIP) of Industry Canada, as overseen by the Minister of Industry in consultation with the Minister of State (Science and Technology), CMCC received a grant in 2009 that established its simulation laboratory. A combined diagnosis and treatment learning lab, the simulation laboratory includes four stations of custom force-sensing table technology (FSTT) that measure the administration of treatment procedures. The FSTT stations provide the student with the opportunity to rehearse the application of treatment skills using passive foam mannequins at first, before progressing to volunteer subjects. Knowledge of Results (KR) is provided in the form of immediate feedback and as a formative assessment by presentation of force-time profiles that quantify the treatment force applications. Direct observation of results for each procedure administered allows the learner and his/her instructor to gauge the relative safety and quality of performance. Coaching and learner reflection on their performance can thus be directed toward strategies that can specifically change individual parameters positively, with tightly yoked feedback that immediately rewards effective change.

Rapidly accepted and lauded as a learning aid by the initial cohort of students who accessed the FSTT (Triano, McGregor & Giuliano, 2011), the literature on this application of KR previously was restricted to the experimental laboratory setting. As such, information on the feasibility of short-term gains in measurable skill parameters existed; however, KR feedback applied more broadly across the curriculum was lacking. As well, information on several characteristics of treatment that are believed to be relevant to enhanced effectiveness of care remained unavailable. For example, the long-term sustainability of such gains, the capacity for a learner to provide targeted levels of performance on demand or even the appropriate target levels for skilled performance were uncertain. In short, the quantification of treatment delivery was at hand, but the means to
assess it appropriately as fitting the clinical learning objectives and career needs for the learner remained undefined.

Through assistance from HEQCO, a strategy to evaluate the FSTT as a learning and assessment tool was implemented in classes of approximately 200 students. It was believed critical to the success of the FSTT implementation that the conceptualization of manual skills learning among both faculty and students be shifted from the traditional coaching model, based on subjective observation, to one that combines objectively measured outcomes with an apprenticeship experience.

Both quantitative and qualitative research methods were employed to address three fundamental questions related to safe and effective treatment delivery:

a) What short-term gains in skill parameters can be obtained and how well are they retained over time?
b) Can learners consciously modulate force applications to comply with the immediate need to respond to changes in criteria?
c) What is the conception of relative confidence and competence in manual treatment skills among learners and supervisors?

The data yielded several important findings related to the effectiveness of implementing FSTT across moderately large class sizes within the curriculum. As this report represents the first evidence of its kind with respect to measuring performance for moderately large groups of learners, multiple statistical tests were performed to identify potentially meaningful findings. Statistical significance was set at $p < 0.001$, using a correction for multiple t-tests to minimize the chance that findings might later turn out to be false.

The findings of the study are summarized as follows:

- What short-term gains in skill parameters can be obtained and how well are they retained over time?
  - On average, the cohort of learners engaged in the FSTT simulation laboratory achieved highly statistically significant gains in parameters of performance (force amplitude and speed) by the end of a two-hour session.
  - Learners not obtaining notable change in speed during FSTT laboratory time were motivated to practice procedures voluntarily ad lib during unscheduled lab time, reflecting on recorded performance and rehearsing procedures.
  - On average, learners who engaged in reflection and rehearsal achieved comparable gains in performance to their peers on re-evaluation.
  - Gains achieved at the end of the first two-hour FSTT, on average, were sustained through a seven-month (for second-year students) interval and five-month interval (for third-year students) between FSTT sessions as determined by formative assessment at the beginning of a second session.

- Can learners consciously modulate force applications to comply with the immediate need to respond to changes in criteria?
  - On average, learners were able to achieve statistically significant modulation of force amplitudes on demand.
  - Gains in ability to modulate over the FSTT session were more easily achieved when attempting to enhance force and only trended to significance for the year 2 cohort for reducing force.
• What is the conception of relative confidence and competence in manual treatment skills among learners and supervisors?
  - Ratings of confidence and competence in performance increased between initial and final clinic experience during the final year of training.
  - Supervisors rated confidence and competence significantly lower than interns rated themselves.
  - FSTT experience was associated with slightly lower ratings by interns of themselves, possibly due to greater self-awareness of shortcomings.
  - Conceptualization and attitudes toward confidence and competence in the themes of commentary became more skill focused for learners and supervisors after experience with FSTT.

There are some limitations to these data, and they are discussed in greater detail later in the report.

The unique skill set required for the application of complex, bimanual tasks associated with manipulation is critical for safe and effective service to the public. FSTT simulation laboratory experience is an effective way to enhance skill development prior to the time a learner faces the need to administer care in a clinical encounter. This project built on experimental work to identify and validate stages of learning and properties of manual treatments. CMCC has now shown that such approaches may be integrated successfully in teaching curricula. Future work must continue with these methods to extend the application, identify ways to enhance the precision of skilled performance and to optimize the care provided to patients who can benefit from these services.
Introduction

In July 2011, the Higher Education Quality Council of Ontario (HEQCO) issued a Request for Proposals that focused on the innovative use of technology in the classroom. The goal was to provide funding to institutions to allow them to evaluate the effectiveness of pedagogical practices that aim to enhance the quality of student learning through the introduction and integration of new technologies. Based on a novel implementation of technology within a new skills simulation laboratory, the Canadian Memorial Chiropractic College (CMCC) submitted a successful application that allowed it to evaluate the system as a means of assessing manual skills development.

CMCC was founded as a professional college in 1945 focusing on the education and training of chiropractors. The four-year program was initially accredited through the Council of Chiropractic Education Canada, a status that continues through the present. The graduate received a diploma after completing course streams in basic science (years 1, 2), clinical science, diagnosis and therapies (years 1, 2, 3, 4) and an apprentice-style clinical internship (year 4). In 2005, CMCC became the first private college to be given approval to award a second-entry undergraduate degree by assent from the Minister of Training, Colleges and Universities, under recommendation from the Postsecondary Education Quality Assessment Board of Ontario (PEQAB). In 2011, that consent to award the degree of Doctor of Chiropractic (DC) was renewed and extended for a ten-year period.

Integral to the administration of care to patients by Doctors of Chiropractic is the safe and effective use of complex, bimanual treatment procedures that can be collectively described as manual manipulation therapy. As a mode of treatment for complaints from the spine, neck and extremity joints, manipulative therapy spans a broad spectrum of maneuvers that have been shown to be useful to the public in managing pain from strain/sprain injuries, degenerative arthritis, disc, shoulder and knee problems, as well as other conditions (Bronfort, Haas, Evans, Leininger & Triano, 2010). Each maneuver must be administered competently in the context of the patient's condition, with consideration given to how chronic or severe the problem is and whether there are other diagnoses that would require modification of the treatment maneuvers during their application.

Through the auspices of the Knowledge Infrastructure Program (KIP) of Industry Canada, as overseen by the Minister of Industry in consultation with the Minister of State (Science and Technology), CMCC received a grant in 2009 that established its simulation laboratory. A combined diagnosis and treatment learning lab, the simulation laboratory houses four stations consisting of computerized and interactive manikins that mimic the presentation of clinical decision-making dilemmas in diagnosis. In addition, there are four stations of custom force-sensing table technology (FSTT) that measure the administration of treatment procedures. The FSTT stations provide the student with the opportunity to rehearse application of treatment skills using passive foam mannequins at first, before progressing to volunteer subjects. Knowledge of Results (KR) is provided in the form of immediate feedback and as a formative assessment by presentation of force-time profiles that quantify the treatment force applications. Direct observation of results for each procedure administered allows the learner and his/her instructor to gauge the relative safety and quality of performance. Coaching and learner reflection on their performance can thus be directed toward strategies that can specifically change individual parameters positively, with tightly yoked feedback that immediately rewards effective change.

Rapidly accepted and lauded as a learning aid by the initial cohort of students who accessed the FSTT (Triano, McGregor & Giuliano, 2011), the literature on this application of KR was restricted previously to the experimental laboratory setting. As such, information on the feasibility of short-term gains in measurable skill parameters existed; however, KR feedback applied more broadly across the curriculum was lacking. Information on several characteristics of treatment that are believed to be relevant to enhanced effectiveness


of care remained unavailable. For example, the long-term sustainability of such gains, the capacity for a learner to provide targeted levels of performance on demand or even the appropriate target levels for skilled performance were uncertain. In short, the quantification of treatment delivery was at hand, but the means to assess it appropriately as fitting the clinical learning objectives and career needs for the learner remained undefined.

Through the assistance from HEQCO, a strategy to evaluate the FSTT as a learning and assessment tool was implemented in classes of approximately 200 students. It was believed critical to the success of the FSTT implementation that the conceptualization of manual skills learning among both faculty and students be shifted from the traditional coaching model, based on subjective observation, to one that combines objectively measured outcomes with apprenticeship experience. Both quantitative and qualitative research methods were employed to address three fundamental issues related to safe and effective treatment delivery:

a) What short-term gains in skill parameters can be obtained and how well are they retained over time?

b) Can learners consciously modulate force applications to comply with the immediate need to respond to changes in criteria?

c) What is the conception of relative confidence and competence in manual treatment skills among learners and supervisors?

Educational Theory in Health Care Neuromotor Skills Development

Theories are the core foundation upon which health care delivery is built. The choice of theory influences the way practitioners engage in both research and practice. Ironically, the conscious application of educational theory within health care is a relatively recent development (Triano, Descarreaux & Dugas, 2012). Most faculty members have tacit models of teaching and assumptions about learning but are unable to articulate a coherent educational theory from which they operate (Morcke & Eika, 2009, p. 642). Modern health care educators (Palter, 2011; Dennick, 2012; Sadideen & Kneebone, 2012) argue that learning theories have much to offer health care education and practice. Dennick (2012) notes that educational psychology, sociology and neuroscience provide overlapping perspectives on the motivation for learning and on how learning occurs that apply specifically to health care providers. Theoretical frameworks should inform practical pedagogical interventions, including the specialized applications involving development of manual neuromotor skills required for graduates over the course of their careers.

A number of health care disciplines (e.g., surgery, dentistry, obstetrics, chiropractic) rely on the administration of manual maneuvers to patients (Shalev, Royburt, Fite, Mashiach, Schoenfeld, Bar, Ben-Rafael & Meizner, 2002; Reznick & MacRae, 2006; Hauser & Bowen, 2009; Triano, Descarreaux & Dugas, 2012; Dennick, 2012). Dennick (2012) draws on constructivist, experiential and humanistic learning theories linked with evidence from neuroscience (Krakauer & Mazzoni, 2011) in suggesting a foundation for teaching sensorimotor skills required within health care practices. Dentistry requires manual dexterity reacting to images reflected in mirrors to examine the inside of the mouth, tasks that are counterintuitive to daily experiences of the novice. Explicit teaching of vital manual skills and emphasizing the continual practice of the basic maneuvers translates to a more successful provision of clinical patient care in the future (Chambers, 1987). Once learned, the expert technical skills become more automatic and require less conscious control to accomplish. As such, active use of expert skills ensures their retention even with aging (Duong, Gardner & Rucker, 2010). Surgical skills are learned via the same principles (Reznick & MacRae, 2006). Dunkin and colleagues (2007) have noted challenges related to the objective and consistent assessment of surgical technical skills that may be resolved through the use of technological tools. Technology-based surgical education has been cited as a promising addition to training (Dunkin, Adrales, Apelgren & Mellinger, 2007), translating to greater clinical proficiency. Kornorffer’s group (2005), for example, used simulation techniques to augment learning in wound suturing. Technology-trained residents showed significantly higher proficiency.
in actual suture performance in the operating room than training with traditional methods (Korndorffer Jr., Dunne, Sierra, Stefanidis, Touchard & Scott, 2005).

The early work of Fitts and Posner (1967) on a theory of staged motor skill acquisition is widely accepted in both the motor skills and the health care literature (Reznick & MacRae, 2006; Hauser & Bowen 2009). The approach postulates distinct yet overlapping stages of learning – the cognitive, the integrative or associative, and the autonomous phases – in which a contextual framework for learning content may be structured (Reznick & MacRae, 2006). During the cognitive stage, meaning of the manipulation task must be developed by learners assembling and integrating basic knowledge in the context of a theoretical patient’s circumstances. The physical performance is conceptualized, together with a plan of steps formulated to complete a manipulation task safely. These processes are grounded in the constructivist view of learning and, from a neuroscience perspective, involve the prefrontal and motor cortical segments of the brain to identify task components and the cerebellum and brain stem to anticipate motor pathway coordination, including boundaries of performance (Krakauer & Mazzoni, 2011; Turner & Desmurget, 2010). The integrative/associative phase is an experiential learning stage in which the learner is immersed in a simulation or working environment filled with many important and relevant experiences to affect knowledge, skills and attitudes (Dennick, 2012). It begins with explicit directions under standardized, often simplified, circumstances. Practice and repetition in manual skills development are designed to gradually approximate the learning outcomes. Knowledge of results, as a formative assessment, is key to the systematic advancement of performance on practice trials (Triano, Scaringe, Bougie & Rogers, 2006). In the seminal work of Adams (1971), learners were theorized to understand errors in performance by comparing KR with peripheral sensory feedback information and coaching related to task trials. Extrinsic coaching and feedback (Hauser & Bowen, 2009; Triano, Descarreaux & Dugas, 2012) focus on the rehearsal of isolated skill components (Krakauer & Mazzoni, 2011; Triano et al., 2012; So, Proctor, Dunston & Wang, 2013) to provide a foundation for later performance. Progression in skill is observed as the development of some judgment and recognition of the need to modify performance in response to coexisting factors. The learner’s brain centers undergo operant reinforcement through sensory monitoring and extrinsic reward during practice sessions. Cognitive brain centers, along with the brain stem and cerebellum, adapt the coordination of movement and force application to correct errors in response to circumstances. The teacher engages humanistic theory by invoking prior learning in relation to potential underlying conditions that may require procedure modification. Coaching and feedback foster trial and error with the objective to identify individual strategies to accomplish skilled performance that accounts for the unique physical attributes of each learner.

Attributed originally to Ericsson (Reznick & MacRae, 2006), milestone states of manual skills development – novice, advanced beginner, competent, proficient and expert – are widely accepted (Dunphy & Williamson, 2004; Hauser & Bowen, 2009). Autonomous learning begins as performance becomes automatic and smooth, forming skill competence. The learner is ready to begin independent practice under supervision and has a range of capacity to reflect and accurately self-assess. Expanded experience provides proficiency with fluid performance that is easily modified, conforming to context. Expert status is achieved over time as performance becomes independent and is intrinsically rewarded. In theory, the pinnacle of expertise is reached when skills are able to be applied successfully under unfamiliar or novel circumstances.

**Manipulation Skills**

As with the application of educational theory to health care teaching in general, the systematic investigation of manipulation of the spine is also relatively recent, with the first scientific conference on the topic held in 1975 in Bethesda, Maryland by the National Institute of Neurological and Communicative Disease and Stroke (Goldstein, 1975). Spinal manipulation is a complex, bimanual motor skill involving various levels of interlimb coordination and postural control combined with a timely weight transfer. Moreover, in a clinical setting, execution is highly adaptive and context-dependent, taking into account any coexisting pathology, structural
weakeness or other factors that may limit the approach to delivering manual maneuvers. Students rely on prior learning in anatomy, biomechanics, physiology and pathology to intellectualize the task and then to translate it into appropriate motor behaviour (Triano et al., 2012).

Performance of a manipulation procedure must simultaneously manage patient and caregiver postures, motions and stability for a safe and effective treatment. Each procedure can be partitioned into a sequence of distinct maneuvers or phases which, when carried out competently, flow smoothly together to deliver a therapeutic force to a local region of the body. The inherent nature of the difference in the interactions between two individuals during movement can be demonstrated by envisioning dance partners. The caregiver acts as lead in the maneuvers and must adapt her or his stature and use of strength to the stature and flexibility of the partner. With a change in partners, there must be an accommodation in the maneuvers performed. Several authors (Kawchuk & Herzog, 1993; Triano, 2001; Cambridge, Triano, Ross & Abbott, 2011) have now contributed to widely accepted descriptions for the primary phases of a procedure through laboratory experiments using force-sensing technology that measures the forces acting on the patient. Figure 1 displays an idealized force-time profile from a procedure and defines the typical phases.

**Figure 1: Force-Time Profile**

![Force-Time Profile Diagram](image)

A typical profile for a manipulation procedure: 1) defines the pre-load force applied during the initial positioning and control of the patient's posture, 2) indicates the rate of rise or “speed” of force production, 3) signifies the peak amplitude of force and 4) represents the duration of the impulse component of the procedure. The duration of the entire procedure is generally less than 300 milliseconds (see Kawchuk & Herzog, 1993).

Explanation of the profile elements are summarized by Triano et al. (2012), herein using “baseline” and pre-load as interchangeable terms:

Typically, quantities that are derived from the profile include pre-load, rate of rise, peak force and duration. When present, a downward incisural point between the baseline and the rate of rise may be quantified. The pre-load is a quasi-static load applied to the surface overlaying the targeted segment of interest. Its purpose is to...
compress the soft tissue and to move the joint toward the limit of its voluntary range of motion. The downward dip or incisural point is a low amplitude reduction from the baseline that may occur just prior to the increase of force leading to a maximum point. Rate of rise, sometimes referred to as speed of delivery, is defined as the change in amplitude from the baseline or incisural point to the peak amplitude divided by the interval of time between those two points. It represents the average rate of force increase. Duration, in most cases, defines the interval from the incisural point to when the fall of peak amplitude crosses the pre-load value. (p. 734)

Laboratory studies have evaluated performance in groups of experts and of learners at different stages in their training and experience. The emphasis for learners is on mastering adequate and appropriate force development while attending to patient comfort and their own sense of confidence. The variables considered most representative of skill acquisition – time to peak force, peak force, rate of force rise – are all interrelated. Under traditional coaching with observation-based feedback, learners show a sigmoidal learning curve, accelerating over the first two years of training with a steep rate of improvement in year 3 (Triano et al., 2012). The development of skilled performance seems to be characterized by an increase in pre-load force amplitudes and in rate of force production to its peak (Cohen, Triano, McGregor & Papakyriakou, 1995; Descarreaux, Dugas, Raymond & Normand, 2005; Triano et al., 2012). KR, using accurate and reliable force-sensing technology, has been applied in isolated experiments to show that skills can be rapidly enhanced by immediate contrast of quantitative performance against the quantitative reference standards of experts (Triano et al., 2006). These changes are retained over short intervals even when challenged with distraction by intense intellectual tasks. Indicators of automaticity are less well studied but appear to be related to a measure of global coordination and the amount of variability seen in a procedure performance, trial to trial (Descarreaux & Dugas, 2010). These are considered to reflect a higher level of expertise and seem to require training and experience to refine. As expressed by Krakauer and Mazzoni (2011), knowing what you have to do at the global task level improves the precision of component movements that are already practiced to a high level.

As the learner begins to act more independently during apprenticeship in the fourth year, there is a levelling off of performance that gradually reaches a plateau by the fifth year in regulated practice after graduation. The course of experience beyond five years is unknown other than what can be extrapolated by direct comparison of novices to peer-selected experts. In an early study by Cohen et al. (1995), experts retained advanced skills only for procedures that they routinely used. For those that were not a common part of their practice, their performance metrics were no better than those of novices who had recently been trained in the procedures.

Teaching and Learning Gaps for Manipulation Training

Instructional engineering principles (analysis, design, evaluation) in designing learning for psychomotor skills indicate that sequencing of content, rehearsals and assessments are important to facilitate more rapid progression to competence (Hauser & Bowen, 2009; McGregor & Quam, 1996). There are currently four alternate strategies for physical practice (Triano et al., 2012): a) choreography for patient-doctor positioning using simulated patients, without the dynamic rehearsal of procedure execution; b) procedure execution with mechanical simulators, without the coordination dynamics of patient transfer and posture management; c) use of simulation mannequins that offer a limited approximation of sensorimotor feedback and patient transfer dynamics; and d) full application of procedures using colleague student volunteers as simulated patients.

Teaching of manual manipulation skills has relied heavily on the evaluation of a learner’s progress largely through the subjective, observational aptitude of tutors’ verbal coaching and on students’ reflective abilities without direct KR or quantitative feedback. Some of the relevant understandings to teaching and learning of complex tasks beyond manipulation skills are informative. James (2012) notes that verbal coaching of
choreography for turning body movements may assist with early improvements, but gains are not well retained. Coaching that physically corrects body posture and relies on the learner’s perceptions of joint positions for improved performance of complex bimanual tasks similarly results in rapid degradation in execution later (Beets, Macé, Meesen, Cuypers, Levin & Swinnen, 2012). Use of augmented feedback involving repetition seems to fair better for retention of gains. In experimental settings, gains can be retained for the short term even when the learner is challenged by intense activity requiring recall of prior knowledge (Triano et al., 2006). Laufer (2008) reports data that suggest that the use of demanding cognitive tasks during physical skill training may actually enhance retention and transfer of skill by forcing learning to involve more automatic processes. As learners reflect on performance through recorded media, self-observation appears to accelerate progress (Ste-Marie, Vertes, Law & Rymal, 2012). Similarly, feedback related to more accurate performance produces more effective physical task learning (Badami, BaezMousavi, Wulf & Namazizadeh, 2012).

Different educational programs have adopted some combination of rehearsal methods, ranging from the strict use of mannequins to that of simulated patients (Triano et al., 2012). Debate on which performance training method is preferable centers on how a conflict in priorities is resolved within the program. On the one hand, there is concern for the safety of volunteers versus an obligation of the apprentice in year 4, and of course for the graduate, who requires full understanding and experience in the procedures, including their relative risks and benefits. Physical risks to both the apprentice learner and the volunteer simulating a patient, while minor in severity and self-limiting, are real. Leading a group of faculty engaged in training using volunteer simulated patients, Kuehnel et al. (2008) took on the question of risk by monitoring unintended side-effects of procedure rehearsal. Incidents tended to occur in the early stage of maturation of skill in academic years two and three. Kuehnel’s group (2008) found that minor adverse events, depending on the training program sampled, occur at rates ranging from 33% to 150% of those seen among practicing graduates and in the general public (Senstad, Leboeuf-Yde & Borchgrevink, 1996). Clearly, the nature of scheduled laboratory practice brings the volunteer into more frequent exposure than any patient would experience. While no serious events have been reported, side-effects should be minimized.

All of these methods pose significant challenges to reaching a learning outcome of confidence and competence in ministering to patients. Currently, psychomotor outcomes are aimed to produce competence, but without the benefit of any objective evaluation methods. Instructor observation is unable to discern the student production of force during procedure application. Thus, directly relevant and explicit feedback on this key factor for safe and effective delivery cannot be provided to assist the learner in understanding errors in performance. The tangible conceptualization of the desired result is absent. Procedures may be demonstrated and coached but not directly experienced or measured. During simulated patient participation, the learner can estimate with some fidelity (Triano et al., 2006) the relative performance of her/his partner. They are then left to reflect on how to compare and translate that experience as a patient to their own performance as operator in administering a procedure. In the case of mannequin use alone, the sensory feedback during the performance has very low fidelity with real world experience. During the time when skill development accelerates the most, at the end of year 3 (Triano, Gissler, Forgie & Milwid, 2011; Descarreaux & Dugas, 2010), there is currently no widely accepted formative or summative tool in the curriculum that can objectively rate the quality of performance and competence of the learner. Regardless of the curricular design, at some point, the learner must confront the realities of providing safe and effective manual care to the public.
Improving Teaching and Learning of Bimanual Skills

Health care education has begun to shift to simulation tools for contextual learning and data-driven feedback. CMCC has responded to the challenges presented by teaching and learning of manipulation skills by linking the underlying educational theory of motor skills development to empirical results in a simulation teaching laboratory outfitted with force-sensing table technology (FSTT). As a curricular tool, the laboratory experience allows the learner to use KR as immediate feedback to reflect on performance. Coupled with traditional coaching feedback, learners hopefully can construct meaning, which they may utilize in turn to inform changes to subsequent performance. Four force-sensing tables were implemented in our institution in 2010. Force-sensing tables provide direct measure of the forces acting through patient tissues with high fidelity (Rogers & Triano, 2003). Pilot work with students has demonstrated an increased engagement of student interest and active participation (Triano, McGregor & Giuliano, 2011). The curriculum for teaching skills was hybridized to embed formative experiences, scheduled for two hours in the laboratory, twice each for years 2 and 3, separated by a five- to seven-month interval. Students rotated in small groups through a KR technology-based portion of the course where their coached skills were quantified directly and performance compared to expert force-time profiles. During the interval between laboratory sessions, students could elect to schedule personal time to rehearse and reflect on digitally recorded performances from the laboratory sessions. The subject of this report is the evaluation of the effectiveness of this hybrid approach to enhance student motor skills.

Research Questions

With support from HEQCO, a strategy to evaluate the FSTT as a learning and assessment tool was implemented in CMCC classes of approximately 200 enrolled students. It was believed critical to the success of the FSTT implementation that the conceptualization of manual skills learning among both faculty and students be shifted from the traditional coaching model, based on subjective observation alone, to one that combines objective outcomes with apprenticeship experience. Both quantitative and qualitative research methods were employed to address three fundamental issues related to safe and effective treatment delivery:

1. What short-term gains in skill parameters can be obtained and how well are they retained over time?
2. Can learners consciously modulate force applications to comply with the immediate need to respond to changes in criteria?
3. What is the conception of relative confidence and competence in manual treatment skills among learners and supervisors?

Research Methods

The research project was superimposed on the standard curriculum, which was modified only to the extent of the pre-planned introduction of a simulation laboratory intervention in years 2 and three. Following a four-year educational curriculum, the learner at CMCC focuses primarily on basic sciences (e.g., anatomy, physiology, microbiology, pathology) in years 1 and 2. Manual skills development, while initiated in year 1, accelerates in earnest during years 2 and 3. Content for clinical sciences, diagnosis and therapies span all years, culminating in an apprentice-style clinical internship in year 4, when skills are applied in supervised health care delivery to the public.
A mixed-methods research design incorporating both qualitative and quantitative data was established to address the research questions. Controlled within-group comparisons were created for the study’s quantitative skill measures (i.e., research questions 1 and 2, above) by taking advantage of the differences arising in the natural progression of the curriculum (Figure 2). Baseline measures of procedure performance were obtained during the second week of the year after orientation to the respective manual skills development course in years 2 and 3. The intervention was defined by the systematic scheduling of learners into the simulation laboratory on two occasions for skills training using FSTT. The first laboratory session followed a baseline evaluation by two weeks. A seven- (year 2) or five- (year 3) month interval separated the lab sessions. The difference in interval length was dictated by unrelated curriculum scheduling. Formative evaluations were carried out at the beginning and end of the initial laboratory session, as well as at the beginning of the second session. A final formative evaluation was carried out on completion of lab session 2.

For quantitative and qualitative components of the research question concerning learners’ conception of relative confidence and competence (question 3 above), a control group comparison was created through a survey of the year 4 class and its clinical supervisors, who had not been exposed to the new simulation laboratory intervention in their respective second and third years of study. These results were obtained three months after the beginning and at the end of their clinical internship year. The initial three-month interval was planned to allow adequate familiarization of the various intern groups’ skill levels by their respective supervisors. Comparison was made with the same measures obtained from the year 3 group, which had just completed the simulation laboratory interventions upon its promotion to year 4.

**Figure 2: Project Timeline and Assessment Intervals**

The timeline shows the intervals and assessments of learners according to the type of data collected. For years 2 and 3, B = Baseline (precedes S1 by two weeks), S1 = FSTT simulation lab 1, S2 = FSTT simulation lab 2, Yr 2 = Year 2, Yr 3 = Year 3. For year 4, conceptualization questionnaires were administered at Q1 and Q2.

<table>
<thead>
<tr>
<th>June</th>
<th>July</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Force-Time Profiles</strong></td>
<td><strong>Summer break</strong></td>
<td>B</td>
<td>S1</td>
<td>S2</td>
<td>Yr3</td>
<td>S2</td>
<td>Yr2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Confidence/Competence</strong></td>
<td><strong>First quarter of internship year 4</strong></td>
<td>Q1</td>
<td>Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FSTT Intervention and Evaluation**

The FSTT consists of a standard treatment table that has been modified to embed a sensing platform capable of recording forces and moments (Triano et al., 2012). The system has the elements necessary for accurate estimation of the loads transmitted through the targeted spinal region. The characteristics of the procedural performance are preserved under configurations consistent with realistic applications of manual treatment methods while respecting the biomechanical recording requirements for accurate measures. Force-time
profiles are recorded electronically. Profiles were post-processed using MatLab software to represent the force-time profiles in anatomically meaningful formats. The display may be used to assess performance against a gold standard or to provide KR feedback as a formative assessment and as a teaching aid (Figure 3). Fidelity of the system in accurate reporting of forces in this type of application was first established by Rogers and Triano (2003). The parameters of the force-time profile on which the evaluation of the effect of FSTT was conducted were pre-load force, rate of rise in force and peak force from the force-time profile, as described in Figure 1. Force-time profiles may be obtained by the learner applying manipulation procedures to a mannequin on the table, a volunteer acting as a simulated patient or a real patient in a clinical setting.

Using a mannequin exercise, the FSTT display is used by the instructor and student to obtain immediate feedback that guides reinforcement of success or the correction of error in force application during clinical maneuvers (Figure 3).

All students completed the initial course orientation and familiarization with materials, reference manuals and treatment tables. The reference manual is common to training in years 1 to 3. Students were divided into groups and then paired by nearest match in stature. Pairing remained constant for the laboratory interventions throughout the course of the project and served two purposes. First, it prevented variation in measures of manipulation skill parameters that would arise solely because of difference in stature, for example, of different simulated patients. Second, stature was considered in the assignment of treatment tables to provide a more ergonomically correct height for the learner administering the procedure. Further demographic information on individuals was not recorded since the experimental design is essentially a pre-test, post-test comparison of individual performance.

In order to determine the stability of learners’ performance prior to the first FSTT laboratory session, a baseline measure was obtained using a convenience sample from available students. Baseline measures were carried out on a subgroup of 140 from the overall sample of participants for year 2, involving 188 students, and year 3, for 186 students. The protocol for formative assessments was identical for baseline and laboratory sessions. Whether the learner’s first exposure to FSTT was at the baseline or at the initial laboratory session, their participation in the project began with them selecting the procedure they preferred to learn from one of three alternatives for treatment to the thoracic spine. Learners had been trained in all three manual maneuvers during the first academic year, where the time spent on these constituted 4.9% of the course content. The choice of procedure remained constant for assessment purposes over the duration of the project. After procedure selection, the learner was encouraged to consult any reference materials as desired before proceeding with a formative evaluation using the FSTT. One member from the pair of learners initially acted as the simulated patient, assuming a prone position on the FSTT standardized for each procedure, while their partner applied the treatment maneuver. Instructions were provided by the laboratory tutors to administer the procedure representing the learners’ perception of a “typical” maneuver. The digital data for the force-time profile was displayed for review and saved to disc. Once data were saved, the pair reversed roles, providing for an evaluation of both members.
The simulated patient was then replaced on the FSTT with a foam mannequin. Using a random sampling method without replacement, the learner then administered a sequence of three procedures to the mannequin attempting to produce: a) a replica of the typical maneuver; b) a similar maneuver to typical but at one-half the amplitude force; and c) a similar maneuver but at twice the amplitude force. These variations created the opportunity to address the research question regarding learners’ ability to modulate force application. Contrasts with baseline performance provided a means to assess stability of performance prior to the use of FSTT.

Following the initial formative assessment during the simulation laboratory sessions, the learner worked with the instructors over a one-hour interval, rehearsing and refining performance on the selected procedure. The FSTT was used to capture each new effort and compare it to the stored profiles. The objective of coaching during this time was to accelerate the learning of how to increase the rate of rise of force and the amplitudes of pre-load and peak force in comparison to the reported maturation rates given in the literature (Descarreaux et al., 2010; Triano et al., 2011). At the end of the lab, the formative assessment protocol was repeated. Comparisons between the first and second assessment were used to evaluate the research question related to short-term gains associated with use of the FSTT.

**Learner Reflection**

Prior to implementation of FSTT, any student wishing to engage in additional reflection on or coaching in manipulation procedures could do so by making appointments with supervising faculty. No systematic reflection experience was available. With inception of the FSTT laboratory, a daily open hour schedule was made accessible for reflection under supervision. A five- to seven-month interval, depending on whether considering year 2 or year 3 participants, was scheduled between the two simulation lab sessions. During this interval all students were offered the opportunity to access the simulation lab at will, including the stored digital data on their individual recorded performances. Students were encouraged to reflect on their earlier efforts using self-observation of prior performance and to attempt to identify errors and develop strategies to reduce them. Tutors were available to coach and respond to learner inquiries and learners were able to control the frequency and timing of feedback to develop a new conceptualization of procedure performance. The maneuvers were rehearsed again using FSTT. Data were stored digitally and kept available for continued reflective practice as desired by the learner. The number of hours spent using the reflection time was recorded for each participant. Results of assessments at the end of FSTT lab session 1 and the beginning of lab session 2 were used to evaluate and compare the performance of those who elected to use reflection time versus those who did not.

For participants who chose not to engage in reflection and practice, the contrast of the second formative evaluation of lab session 1 with the pre-lab formative assessment from lab 2 allowed a gauge of how well skills developed using the FSTT were retained over prolonged periods. In the case of those who took advantage of reflection, the same comparison gauged the impact of retention and reflection on performance gains. Finally, contrast of performance between the final formative assessment of session 1 to the initial formative assessment at the beginning of session 2 provided a determination of long-term gains over the course of the full intervention.

**Conceptions of Confidence and Competence**

In order to assess the impact of the simulation laboratory experience on the conception of relative confidence and competence in manual treatment skills among learners and supervisors, it was necessary to identify and measure a control group sample. Looking again to the natural progression within the curriculum, concurrent with the implementation of the FSTT itself in years 2 and 3, quantitative and qualitative data on the
perceptions of confidence and competence were captured from year 4 interns and their supervisors. None of the learners in year 4 were familiar with or experienced in the use of FSTT.

To provide for adequate exposure between interns and their supervisors to warrant a grounded opinion of intern confidence and competence, sampling was carried out at an interval of four months after the fourth-year class commenced its internship in June. The interns are assigned to supervisors in cohorts of six to eight. Both interns and supervisors were given questionnaires, copies of which are available in the appendix. Interns responded with respect to their individual perceptions of the recall of their own confidence and competence in administering manual treatment procedures on entering the internship. Supervisors scored their perception of the cohort on average on entering their clinical internship year.

Sampling took two forms. In the first, two 10 cm visual analogue scales were created to quantify the responder’s perception of confidence and competence in performance of manipulation. Each scale was anchored to the left with the appropriate respective phrase “Not confident/competent at all”, while the right was based on “Completely confident/competent.” Quantitative scores were obtained by measuring the distance in millimeters of the respondent’s mark on the scale from the left hand margin and dividing by 100. The second form of response was qualitative, where the respondent was allowed to make open-ended comments on their perceptions. The text from each respondent was transcribed and pooled across the respondents for text analysis (Corman, Kuhn, McPhee & Dooley, 2002) using a commercially available software entitled Crawdad™ (version 1.2). A second sampling was obtained in the same fashion from the year 4 class at the end of internship in May.

Evaluation of the effect of the FSTT laboratory was conducted by obtaining a sample of the learners who experienced the lab in their third year upon reaching the four-month milestone after promotion to year 4.

Data Analysis

In the analysis of gains in and control of force parameters, the force-time profiles originally obtained from the learner’s performances at baseline and laboratory sessions 1 and 2 served to provide quantitative measures of the procedures exemplified in Figure 1. Each profile was displayed and then divided into segments defining the parameter boundaries for baseline, rate of rise in force and peak force amplitude and the respective values were calculated. Descriptive statistics were calculated for each assessment. Student’s paired t-tests were used to compare stability of performance before implementing the FSTT laboratory sessions by contrasting measures at baseline with the initial formative assessment in session 1. Paired t-tests were also used to evaluate short-term gains from the laboratory experience by contrasting the pre-session and post-session formative assessments for session 1, as well as retention of gains by comparing post-session 1 with pre-session 2. Effects of reflection practice were assessed by evaluating for differences between the reflection group and the non-reflection group at the pre-session 2 time point.

Perceptions of confidence and competence were quantified as visual analogue scores. Comparisons were made using two-way analysis of variance (ANOVA) to evaluate the effects of FSTT between interns and supervisors, by year and exposure to FSTT. A secondary analysis examined the change in scores as a result of internship even without FSTT. Finally, the themes of conversation about manual procedures and their skilled performance were assessed through Crawdad™ analysis of open-ended commentary by interns and their supervisors.
Limitations of the Study

There are a number of limitations that should be considered when interpreting the study results. Perhaps the main factor constraining conclusions in several parts of this project is the inability to physically isolate the learners and supervisors, separating them from other influences in the academic environment. General knowledge of the implementation of the FSTT was available throughout the academic community. Learners continued with other relevant course work. Skill development classes were not focused specifically on the three optional maneuvers available to participants in the study. However, reviews of these procedures were a part of the spectrum of treatments within the course outlines for year 2, making up 1.7% of the course content. Such overlap and persistence of coaching in related content may have had an influence on the reflection and practice and retention of gains observed in the study.

Supervising clinicians overseeing the internship in year 4 were aware of the institutional plan to implement the FSTT laboratory and had the expectation that future learners would be reaching them who would have experience with the laboratory. Such awareness may well have influenced participants’ conceptions and attitudes. Likewise, the number of participants who contributed commentaries on confidence and competence was small, which may have skewed the interpretation of the analysis of the conversation.

As with any new program, there is an initial level of enthusiasm that may change on repetition or if the lab were to be conducted by different faculty. While unlikely, the collective group of learners involved in the simulation laboratory experience may have been unique in their manual dexterity and ability to learn complex motor skills. Consequently, these findings should be replicated in the hands of others to confirm the effects of the FSTT simulation experience before assuming broad ability to generalize.

Findings

The data yielded several important findings related to the effectiveness of implementing FSTT across moderately large class sizes within the curriculum. Summaries of the main findings are provided below, followed by a discussion of their implications. As this report represents the first evidence of its kind with respect to measuring performance for moderately large groups of learners, multiple statistical tests were performed to identify potentially meaningful findings. Statistical significance was set at p < 0.001, using a correction for multiple t-tests to minimize the chance of findings that might later turn out to be false.

Performance Measure Stability

Accurate measure of performance requires good reliability and fidelity from the instrumentation used and an understanding of the inherent variability in the performance itself. The FSTT instrumentation has been previously validated (Rogers & Triano, 2003). To assess the inherent variation in performance among learners, baseline measures were obtained from a group of 140 learners across both years 2 and 3 two weeks prior to the first FSTT simulation laboratory. During the interval, routine delivery of course content appropriate to the academic year continued, unrelated to the FSTT simulation lab and exclusive of the procedures related in this study. Results were paired between the pre-session (S1) typical force-time profiles and the baseline to look for differences in peak amplitude and rate of rise in force application (Table 1). Learner performance of the manual treatment maneuvers proved to be quite stable over the interval between measures for both those in year 2 and in year 3. For year 2, the S1 typical peak force measure mean was 430 Newtons (N), and was not statistically significantly different from the baseline measure of 444 N (p = 0.3209). Similarly for year 3, the baseline mean was 464 N, with S1 at 452 N. Again, there was no statistically
significant difference in peak force between time points ($p = 0.3936$). For rate of rise in force, year 2 performed on average at 2182 Newtons per second (N/s) at baseline, with 2184 N/s at S1 ($p = 0.9867$). The rates were slightly higher in the mean for year 3 learners, with baseline of 2294 N/s and S1 at 2383 N/s; however, again there was no difference in rates between time points ($p = 0.3287$).

Table 1: Performance Stability over the Two Weeks from Baseline to the Pre-Session 1 Assessment

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Baseline (se)</th>
<th>Session 1 Pre (se)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peak Force (N)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>68</td>
<td>444 (12)</td>
<td>430 (14)</td>
<td>0.3209</td>
</tr>
<tr>
<td>Year 3</td>
<td>71</td>
<td>464 (17)</td>
<td>452 (12)</td>
<td>0.3936</td>
</tr>
<tr>
<td><strong>Rate of Rise (N/s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>68</td>
<td>2182 (131)</td>
<td>2184 (101)</td>
<td>0.9867</td>
</tr>
<tr>
<td>Year 3</td>
<td>71</td>
<td>2294 (119)</td>
<td>2383 (97)</td>
<td>0.3287</td>
</tr>
</tbody>
</table>

Short-Term Gains from FSTT

Based on earlier work (Triano et al., 2006) using visual feedback KR from force-time profiles, gains were expected to occur in both year 2 and 3 learners following a two-hour session inclusive of formative assessments with KR from FSTT and one hour of guided coaching. At S1, the pre-laboratory assessment was compared to the post-lab assessment. Data on peak force and rate of rise in force were paired and tested for differences (Figures 4, 5 and Table 2). Learners in year 2 demonstrated a highly significant ($p = 0.000$) gain of 13.9% in their ability to generate peak force amplitude and an increase in rate of rise by 10.5% ($p = 0.0003$). For year 3, peak amplitude gains were 9% ($p = 0.0000$), though rate of rise in force was less clear, only approaching significance ($p = 0.0893$), with an increase in the mean of 3.9%. All together, short-term gains were strong in force production, while increase in rate of rise or “speed” of performance was more evident in year 2.

Short-term changes for the second administration of FSTT and KR (S2) did not show significant change associated with the two-hour laboratory in either peak force (year 2, $p = 0.3343$; year 3, $p = 0.0196$) or speed measures (year 2, $p = 0.9420$; year 3, $p = 0.2226$).
Table 2: Change in Performance Parameters from Participation in Session 1

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Pre-S1 (se)</th>
<th>Post-1 (se)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peak Force (N)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>186</td>
<td>454 (7.3)</td>
<td>517 (6.5)</td>
<td>0.000</td>
</tr>
<tr>
<td>Year 3</td>
<td>186</td>
<td>485 (8.6)</td>
<td>529 (6.9)</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Rate of Rise (N/s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>185</td>
<td>2263 (67.1)</td>
<td>2501 (60.1)</td>
<td>0.0003</td>
</tr>
<tr>
<td>Year 3</td>
<td>186</td>
<td>2584 (69.7)</td>
<td>2649 (61.4)</td>
<td>0.0893</td>
</tr>
</tbody>
</table>

Figure 4: Force Amplitudes

Changes in force amplitude from formative assessments before and after the first FSTT session for each group:
Figure 5: Rise in Force

Changes in rate of rise in force from formative assessments before and after the first FSTT session for each group:

In summary, for short-term gains, the first session showed gains for both year 2 and year 3 students in peak force development. In the case of rate of rise in force, only year 2 students gained significantly. Session 2 showed no additional gains for either set of learners.

Retention of Gains

Learning retention of manual skills has been notably problematic as it may be influenced by the rate and frequency of feedback (Pringle, 2004), as well as by the demand of simultaneous cognitive tasks (Laufer, 2008). From feedback with KR, the longest duration of successful retention of manipulation skills following an intellectually distracting exercise has been for short intervals of ten minutes (Triano et al., 2006). The embedding of simulation lab exposures, interspersed over five to seven months, allowed comparison of results from formative assessments post-S1 versus pre-S2. Short-term gains arising during S1 could have been reinforced or degraded as a result of the time lag and intervening course content exposure. The results of these interactions may be observed in the apparent retention of performance as defined by force-time parameter differences (Table 3).
Retention of gains was assessed for the year 2 and year 3 groups as a whole, not accounting separately for individuals who may or may not have taken advantage of opportunity for reflective practice. Those results are described under the section on reflection. For year 2 learners as a whole, no significant change \((p = 0.2223)\) was seen in paired peak force measures from the end of the S1 to the onset of S2. However, an 8% gain in performance was noted for year 3 \((p = 0.0000)\). For the parameter of speed as the rate of rise in force, year 2 again showed no significant difference \((p = 0.0107)\) although, in the mean, an increase of 11.2% was noted. Also, for year 3, there was an increase effect with a change in speed of 12.6% \((p = 0.0000)\).

On the whole, then, gains achieved from S1 experience were retained across the five- to seven-month interval up to S2. For year 3, further gains were experienced during the interval. The data themselves do not permit a specific evaluation of any influence of the fact that each learner continued in the standard curriculum and training in related but different procedures than those used for the assessments in S1 and S2. The work of Cohen et al (1995) suggests that the amount of crossover between different procedures may be limited.

**Table 3: Retention of Gains Represented as the Difference Scores from Post-S1 to Pre-S2 Assessments**

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Post-S1 (se)</th>
<th>Pre-S2 (se)</th>
<th>Difference (se)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peak Force (N)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>179</td>
<td>528.7 (6.6)</td>
<td>533.1 (11.1)</td>
<td>14.3</td>
<td>0.2223</td>
</tr>
<tr>
<td>Year 3</td>
<td>183</td>
<td>528.6 (7.0)</td>
<td>572 (8.4)</td>
<td>43.4</td>
<td>0.0000</td>
</tr>
<tr>
<td><strong>Rate of Rise (N/s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>177</td>
<td>2497 (61.8)</td>
<td>2777.6 (113)</td>
<td>280.5 (108.8)</td>
<td>0.0107</td>
</tr>
<tr>
<td>Year 3</td>
<td>183</td>
<td>2648 (62.3)</td>
<td>2983.0 (70.9)</td>
<td>335 (66.)</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**Reflection**

All students were provided the option of spending additional time in the simulation lab between S1 and S2. A total of 48 out of 181 year 2 learners (27%) took advantage of this opportunity, accumulating an average of 51.8 \((\pm 51.9)\) minutes reflective practice time. A total of 61 out of 183 year 3 students (33%) made the same choice, using average reflection time of 45.3 \((\pm 45.3)\) minutes. Upon evaluating the differences between those who chose to use reflection opportunities versus those who did not, it was found that learners who sought reflection time after S1 had made lesser percentage gains during their first FSTT lab with respect to speed. That is, students choosing reflection in each of years 2 and 3 made only an insignificant 2.9% gain in speed from the beginning of S1 to the end of S1 \((p = 0.6065\text{ and } p = 0.4023\text{ respectively})\). Learners who did not choose reflection opportunities in year 2 made a highly significant 13.4% gain in speed on average during the S1 laboratory \((p = 0.0001)\). In year 3, those who did not choose reflection had a 4.4% gain, although this was not statistically significant \((p = 0.1405)\).

Interestingly, there was a strong and significant gain in skill between the end of S1 and the beginning of S2 for students who chose reflection experiences (Table 4). Peak force for these learners in years 2 and 3 increased by 14% and 16%, respectively, with \(p = 0.0000\) in both cases. For speed, students selecting reflection in year 2 gained 21% \((p = 0.0014)\). Those selecting reflection in year 3 gained 19.6% \((p = 0.0000)\). Although learners not choosing reflection maintained their gains from the end of S1, they failed to show the substantial changes between S1 and S2 of those choosing reflection. On the formative reassessment for Pre-S2, comparison of
data in Tables 3 and 4 demonstrates that those who used reflective practice achieved as much gain or better than their other colleagues. It should be recalled that the natural maturation of skill in manipulation procedures occurs inherently in the year 3 program (Triano et al., 2012). This likely explains the trend to improve between Post-S1 and Pre-2 assessments for year 3 in the group not choosing reflective practice.

**Table 4: Gains in Skill Parameters from Reflective Practice**

<table>
<thead>
<tr>
<th>Reflective Practice</th>
<th>Sample</th>
<th>Post-S1 (se)</th>
<th>Pre-S2 (se)</th>
<th>Difference (se)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peak Force (N)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>48</td>
<td>514.3 (11.4)</td>
<td>587.4 (14.9)</td>
<td>73.1 (15.3)</td>
<td>0.0000</td>
</tr>
<tr>
<td>Year 3</td>
<td>61</td>
<td>527.4 (10)</td>
<td>611.4 (12.7)</td>
<td>84.0 (12.1)</td>
<td>0.0000</td>
</tr>
<tr>
<td><strong>Rate of Rise (N/s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>48</td>
<td>2418 (103)</td>
<td>2928 (161.4)</td>
<td>510.4 (150.7)</td>
<td>0.0014</td>
</tr>
<tr>
<td>Year 3</td>
<td>61</td>
<td>2802.6 (104.9)</td>
<td>3353.5 (111.1)</td>
<td>550.9 (109.6)</td>
<td>0.0000</td>
</tr>
<tr>
<td><strong>No Reflective Practice</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Peak Force (N)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>131</td>
<td>520.3 (8.1)</td>
<td>513.2 (13.9)</td>
<td>-7.2 (14.6)</td>
<td>0.6238</td>
</tr>
<tr>
<td>Year 3</td>
<td>122</td>
<td>529.1 (9.2)</td>
<td>552.3 (10.5)</td>
<td>23.2 (10.7)</td>
<td>0.0329</td>
</tr>
<tr>
<td><strong>Rate of Rise (N/s)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>131</td>
<td>2526.6 (75.7)</td>
<td>2721.5 (143.6)</td>
<td>194.9 (137.9)</td>
<td>0.1598</td>
</tr>
<tr>
<td>Year 3</td>
<td>122</td>
<td>2570.8 (76.8)</td>
<td>2797.8 (86.3)</td>
<td>227 (81)</td>
<td>0.0059</td>
</tr>
</tbody>
</table>

For those learners with lower performance after experiencing the S1 session, reflective practice on the test procedures resulted in improved performance comparable to that of their colleagues by the time of Session 2. The cumulative time to achieve these results over the five- to seven-month interval was relatively small, at less than an hour on average.

**Intentional Modulation of Force**

One of the desired outcomes of skill development is the ability to modulate force application to match the clinical context. As each learner produced their individual amplitude for typical peak force, no set quantity allowed comparison for the groups as a whole. In the learning context, the protocol design modeled force modulation as a change on demand from the individual typical peak force to one targeting half of typical or one that was double. Given that the findings related to short-term gains were more evident in S1, comparisons for the determination of modulation in force were limited to the first session.
Difference scores were created by subtracting the targeted force amplitude (e.g., half/twice of typical) from the typical force amplitude (Table 5). This was done for the data from both the pre-S1 and the post-S1 assessments. Difference scores (i.e., half from typical/twice from typical) were paired for each subject and the gain in performance due to participation in S1 was determined by subtracting the difference scores of the pre-S1 assessment from the post-S1 assessment. Gains were tested by one-sided Student's t-test, since the desired direction of effect (reduction/enhancement) for the force amplitude from the S1 coaching and KR rehearsals was known ahead of time.

Table 5: Mean Gains in Force Modulation

<table>
<thead>
<tr>
<th></th>
<th>Sample</th>
<th>Pre-S1 Mod (se)</th>
<th>Post-S1 Mod (se)</th>
<th>Gain (se)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typ - Half</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>185</td>
<td>-130.4 (6.1)</td>
<td>-148.1 (6.8)</td>
<td>-17.7 (8.3)</td>
<td>0.0167</td>
</tr>
<tr>
<td>Year 3</td>
<td>186</td>
<td>-149.3 (5.8)</td>
<td>-151 (6.9)</td>
<td>-1.7 (7.8)</td>
<td>0.4149</td>
</tr>
<tr>
<td>Typ - Twice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>184</td>
<td>105.8 (6.9)</td>
<td>162.4 (9.9)</td>
<td>56.6 (9.7)</td>
<td>0.0000</td>
</tr>
<tr>
<td>Year 3</td>
<td>186</td>
<td>78.2 (6.8)</td>
<td>123.4 (8.0)</td>
<td>45.2 (8.5)</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

As was expected from the short-term gains data, overall increases in peak force application were observed (Table 2 and Figure 6) during the S1 Lab. For year 2, the mean typical force increased from 454 N to 517 N. A smaller increase from 485 N to 528 N was observed in year 3. As shown in Figure 6, learners were able to modulate force in the direction of targets (Half-typical/Twice typical) but were unable to accurately achieve the target values. Regardless of academic year, learners found the requirement to decrease force amplitude, targeting half of typical, more challenging. For year 2, the learners showed a trend (p=0.0167) toward a gain in downward modulation of force amplitude, while the year 3 group was unchanged (Table 2). The results were markedly different when augmented effort targeting double of typical was required. The additional demand force amplitude accomplished over the S1 lab experience for year 2 was 53.5%, in the mean (p=0.0000). For year 3, the additional demand improved by 57.7% (p < 0.0000).
In summary, learners are able to modulate the forces they apply in desired directions, although the targets of half and double may be unrealistic. The laboratory session appeared to be more effective for learning further how to increase force than to decrease it.

**Conceptualization and Perceptions of Skill Development**

The teaching and learning of manual skills development, like that of other technical skills (Rutherford & Ahlgren, 1991), grew out of personal experience with coaching techniques and was handed down from generation to generation. CMCC believed it critical to the success of the FSTT implementation that the conceptualization of manual skills learning among both faculty and students be shifted from the traditional coaching model, based on subjective observation, to one that enfolds objective outcomes into an apprenticeship experience. As a result, both quantitative and qualitative research methods were employed in an effort to identify the conversation among learners and supervisors about manual treatment skills. Three inquiries were made, engaging two sets of learners with their supervisors for consecutive classes entering the fourth-year internship. The first set consisted of different individuals than those who participated in the S1 or S2 sessions precisely because they had no experience with the FSTT technology and could act as a control group. These were year 4 students during the time that the research was on going with the year 2 and year 3 students in the FSTT labs. This year 4 group of 183 interns and their supervisors were sampled using visual
analogue scales for their sense of confidence and competence in manual treatment skills of the interns. Visual analogue scales allow the respondent to estimate their sense of skill level on a 10 cm line labeled “Not confident/competent at all” at “0” measure and “Fully confident/competent” at “10”. The relative distance along the line gives a quantitative estimate that is used in Figures 7 and 8 to represent descriptively the mean and variation of perceived skills. The related themes were captured by open-ended commentary on a questionnaire at two time points (i.e., Q1/Q2) during the one-year internship. Samples were obtained from the first cohort after the first four months into the twelve-month clinical internship and again nine months later at the termination of the internship. The second cohort of 186 year 4 interns who had exposure to the FSTT simulation laboratory (the previous year 3 students from the FSTT study as described above) was sampled only after the first four months into the internship. The sampling strategy allowed CMCC to evaluate the difference in sets of cohorts but also provided a chance to see if the clinical internship itself carried an influence.

Not surprisingly, a significant difference (p < 0.0000) in the perceptions of intern confidence and competence was observed between learners and their supervisors (Figure 7) based on the visual analogue scores. While experience in the clinic raised the confidence of both parties in the first cohort over time, the supervisors provided a significantly lower rating, even at the end of the internship year, than was felt by the learners. Regardless, supervised experience in an apprenticeship role over the year appeared to make a difference in its own right.

**Figure 7: Perceptions without use of FSTT**

Relative perceptions of confidence and competence for the cohort with no exposure to FSTT at the initial assessment in clinic compared to the final assessment based on visual analogue scores:
Perceptions at the initial assessment from the cohort that had experienced FSTT are shown in Figure 8. Small increases in the mean estimates of intern confidence and competence level were reported by supervisors when compared to the prior cohort of students who had not had experience with the FSTT (mean confidence: 45.22 and 50.44, respectively, p=0.1171; mean competence: 50.35 and 53.91, respectively, p=0.2249). However, these differences did not reach statistical significance using one-tailed t-tests. Learners expressed a slightly lower sense of both confidence and competence after FSTT (mean confidence: 64.72 and 62.03, respectively, p=0.0807; mean competence: 67.36 and 65.74, respectively, p=0.1514). Again, these differences did not reach statistical significance. The possible trends toward greater confidence and competence scores by supervisors, despite the lack of statistical significance, may speak to some sense of relative improvement perceived. Also, the trends toward lesser confidence and competence scores by students, despite the lack of statistical significance, may speak to some sense of self-realized shortcomings. Both subjective differences warrant consideration for future work.

**Figure 8: Perceptions with use of FSTT**

Relative perceptions of confidence and competence for the cohort with FSTT experience, in the academic year prior to clinic, at the initial assessment in clinic compared to the prior cohort without FSTT:

In addition, the conversation around the development of skill appears to have shifted with the implementation of the simulation laboratory experience. Although, only a small percentage of supervisors and/or students elected to provide comments at each time point, trends toward greater focus on confidence and competence were observed. Using textual analysis, comments given by respondents were pooled. Nouns and noun phrases, considered centers of meaning in communication (Corman et al., 2002), were organized into word networks. Words were scored for influence, ignoring articles of speech, based on how they formed a coherent linkage. Word pairs were ranked for resonance, which describes the mutual relevance of words. Networks were mapped to show the strength of connection in themes throughout the communication. Focus of the
conversation was scored and defined by how tightly the network was organized around the centering theme. In analysis of the comments, a cut-off for influence scores was set at 0.10, considered highly significant by researchers in the field (Corman, 2005). Words above the influence score threshold were tallied and grouped to describe the emphasis within the conversation.

Table 6 provides a summary of the data on the open-ended commentary given by interns and supervisors based on text analysis (Crawdad™). The first column describes the time point (i.e., Q1 at three months after entering clinic and Q2 for nine months after entering clinic), as well as the role of the commentator all under the heading “Group”. Under “Sample”, the quantities represent the percentage of interns or supervisors providing comments upon which the text analysis was based. Finally, the overall focus score for each network of nouns and noun phrases, highlighting how tightly organized the network was, is provided under the title “Focus”.

Thus, for interns commenting on confidence and competence, who had no FSTT training and who were first entering the clinic in 2011, 16% provided comments and the Focus score for these pooled comments was 0.26. For interns entering the clinic in 2012, who had FSTT training, only 11% chose to provide comments, yet their focus was 0.40. The focus of discussion on confidence and competence then moved from the “slightly low” rank to the “slightly high” rank.

Table 6: Focus Scores from Text Analysis on Competence and Confidence

Focus scores from the analysis of open-ended comments submitted by learners and their supervisors regarding intern confidence and competence with benchmarks for interpretation (Crawdad™):

<table>
<thead>
<tr>
<th>Group</th>
<th>Sample</th>
<th>Focus Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoFSTT-Q1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intern</td>
<td>16%</td>
<td>0.26</td>
</tr>
<tr>
<td>Supervisor</td>
<td>26%</td>
<td>0.31</td>
</tr>
<tr>
<td>NoFSTT-Q2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intern</td>
<td>20%</td>
<td>0.22</td>
</tr>
<tr>
<td>Supervisor</td>
<td>20%</td>
<td>0.26</td>
</tr>
<tr>
<td>FSTT-Q1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intern</td>
<td>11%</td>
<td>0.40</td>
</tr>
<tr>
<td>Supervisor</td>
<td>7%</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Benchmarks for interpreting scores

<table>
<thead>
<tr>
<th>Benchmarks for interpreting scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;.60 Very High</td>
</tr>
<tr>
<td>.45-.55 High</td>
</tr>
<tr>
<td>.30-.45 Slightly High</td>
</tr>
<tr>
<td>.15-.30 Slightly Low</td>
</tr>
<tr>
<td>.05-.15 Low</td>
</tr>
<tr>
<td>&lt;.05 Very Low</td>
</tr>
</tbody>
</table>

Overall, the focus scores for both supervisors and interns without FSTT experience (Table 6) were lower and were stable across the internship year than for those with FSTT experience. Themes of the conversation provided by learners without FSTT experience at their initial evaluation during their clinical year were retrospectively focused on the prerequisite classroom experience. General terms about clinical “technique” were the strongest emphasis, comprising 30% of terms in the group. The terms “tutor” and “time” followed. A limited mention of competence/confidence was made, at only 10%, despite it being the topic of the questionnaire inquiries. Supervisors had a different emphasis, with specific terms on clinical procedures equaling 68%, followed by competence, skill and confidence to fill out the remainder. At the end of the internship, a lower focus was observed for both learners and their supervisors. Learners’ voices remained more general in scope of clinical technique and time, summing to 39%. A small shift was noted in terms of
adding “skill”, with a tally of 17.2%, and “clinic”, “patient” and “student”, totaling 8.7% each. The supervisors, on the other hand, were voicing “high/confident/competent” for 60%, with specific clinical techniques and clinical administrative terms distributing the balance.

The conversation following the FSTT simulation lab experience was notably different for both the learners and supervisors. Focus scores increased for commentary from the learners by 55%. Supervisors were also more tightly focused, showing a lesser increase of 21%. Learners had a split in emphasis between general clinical technique, but a reduction on “time” and an increased weight given to both confidence and competence. The terms used by supervisors of this cohort showed a clear shift. Emphasis, in terms of tallies, was 43% on “competent”, with the next strongest reference to clinical technique at 21%. The remaining emphases trailed, divided equally by “patient” and time.

Conclusions

By employing the several analyses of the data presented in this report, we can draw several conclusions, which are grouped according to the original hypotheses for convenience.

- What short-term gains in skill parameters can be obtained and how well are they retained over time?
  a) The baseline performances of learners on assessment maneuvers where they have already received instruction were stable. The procedures used in this study were all a part of the year 1 curriculum. Baseline tests of performance were unchanged over the two-week interval prior to the first FSTT laboratory experience. As a result, it is unlikely that any important improvement would have occurred in the baseline performance of the selected procedure until the learner reached the clinic internship, where practice occurs with live patients.
  b) The short-term gains achieved in the performance of the selected procedure were attained primarily during the first FSTT session. The use of sessions with FSTT, provided as an optional additional resource to students, may be better focused on advancing skill in additional maneuvers to expand the breadth of the learner skill set in manipulation.
  c) Learners who successfully show short-term gains with FSTT appear to have long-term retention of their improvements, provided that they continue in related skill development course activity. This suggests some transfer of skill, at least toward retention of gains. Other evidence suggests that continuous use of the skills in clinical settings may create a lifelong expertise that persists even through aging.
  d) Learners not obtaining notable change in speed during FSTT laboratory time were motivated to practice procedures voluntarily ad lib during unscheduled lab time, reflecting on recorded performance and rehearsing procedures.
  e) On average, learners who engaged in reflection and rehearsal achieved comparable gains in performance to their peers on re-evaluation.

- Can learners consciously modulate force applications to comply with the immediate need to respond to changes in criteria?
  a) On average, learners were able to achieve statistically significant modulation of force amplitudes on demand.
  b) There appears to be a natural baseline performance that can be readily increased but is more difficult to decrease. Data from this study were not designed to achieve a specific force target.
That work and the work to identify strategies to limit force amplitude on demand more easily remain to be done.

- What is the conception of relative confidence and competence in manual treatment skills among learners and supervisors?
  
  a. Ratings of confidence and competence in performance increased between initial and final clinic experience during the final year of training.
  b. Supervisors rated confidence and competence significantly lower than interns rated themselves.
  c. FSTT experience was associated with slightly lower ratings by interns of themselves, possibly due to greater self-awareness of shortcomings.
  d. Conceptualization and attitudes toward confidence and competence in the themes of commentary became more skill focused for learners and supervisors after experience with FSTT.

The unique skill set required for the application of complex, bimanual tasks associated with manipulation is critical for safe and effective service to the public. FSTT simulation laboratory experience is an effective way to enhance skill development prior to the time a learner faces the need to administer care in a clinical encounter. This project built on experimental work to discover and validate stages of learning and properties of manual treatments. CMCC has now shown that such approaches may be integrated successfully in teaching curricula. Future work must continue with these methods to extend the application, identify ways to enhance the precision of skilled performance and to optimize the care provided to patients who can benefit from these services.

What remains as a significant challenge that previously could not have been adequately addressed without FSTT is the determination of the appropriate level of specific manual treatment parameters that should be the target for delivery under select clinical conditions. Future work to apply this technology as a fully summative assessment integrating clinically relevant factors associated with a patient’s diagnosis and comorbid conditions may now be engaged.
References


Morcke, A. M., & Eika, B. (2009). Medical faculty and curriculum design - 'No, no, it's like this: You give your lectures ...': *Medical Teacher, 31*(7), 642-648.


