# Fundamentals of Robotic Surgery
# Consensus Conference III: Validation Study Design

## Meeting Summary Report

MITIE, Houston, TX
November 17-18, 2012

## TABLE OF CONTENTS

# EXECUTIVE SUMMARY:

## Introduction and FRS Project Background: Dr. Richard Satava

The Fundamentals of Robotic (FRS) is a joint industry (Intuitive Surgical, Inc) and Department of Defense (DoD) funded project jointly managed through Minimally Invasive Robotics Association (MIRA) and Florida Hospital Nicholson Center for Surgical Advancement (NCSA). The mission of FRS is to create and develop a validated multi-specialty, technical skills competency based curriculum for surgeons to safely and efficiently perform basic robotic-assisted surgery.

3 consensus conferences were planned in the following topics:

1. Outcomes Measures
2. Curriculum Development
3. Validation Study Design

These conferences will be followed by multi-institutional validation study with participation by multiple surgical specialties at each institution.

## Lectures about Validation

1. Validation: Tools for Improving Curriculum and Assessment: Dr. Sara Kim, ISIS, University of Washington Medical Center
    a. Essence of validation
    b. Framework of validation
    c. Types of validation study designs
    d. Implications for FRS validation
2. Toward a Course Validation Template: Dr. Wallace Judd, Authentic Testing Corporation
    a. Defining tasks and skills
    b. Looking at a typical course
    c. Characteristics of validation tasks
    d. Course validation template
3. Validation Study Design & Methods: Dr. Anthony Gallagher, University College Cork
    a. Validation is a process
    b. Metrics for an optimal training program
    c. Validation: Why is it Important?
    d. What needs to be done to advance the validation process?
    e. Possible model for robotic validation studies

## Break out sections

### Measures and Metrics Group

1. The goal of the Measures and Metrics Group is to answer the question: what are the metrics?
2. Measurements and metrics, errors, and critical errors were determined for the 7 technical skill tasks
    a. Docking & Instrument Insertion
    b. Ring Tower Transfer
    c. Knot Tying
    d. Railroad Track
    e. 4th Arm Cutting
    f. Cloverleaf Dissection

    g. Vessel Dissection/Division
3. Metrics around team/communication issues for the high stakes exam (HSE) were also discussed

### Study Design Group

1. The goal of the Study Design Group is to focus on the psychomotor component and conceptualize a validation study (or series of studies)
2. Define the research questions
3. Define an characteristics of an experts
4. What will be measured?
5. Study design
    a. Phase 1: Pilot at Florida Hospital Nicholson Center (logistics and refinements to model)
    b. Phase 2: Get face and content validity from the society leadership and boards
    c. Phase 3: Get face, content, and construct validity at test sites and society meetings
    d. Phase 4a: Get concurrent validity with video correlations
    e. Phase 4b: Get predictive validity – full research study at 10 sites
6. Validity questions

## Open Forum

Following the main meeting, the floor was opened to everyone including industry to provide input and ask any questions. This was not part of the main meeting and had nothing to do with curriculum development, so there is no real or perceived conflict (bias) from industry.

# WELCOME

## Brian Duncan, MD, The Methodist Hospital, Houston TX

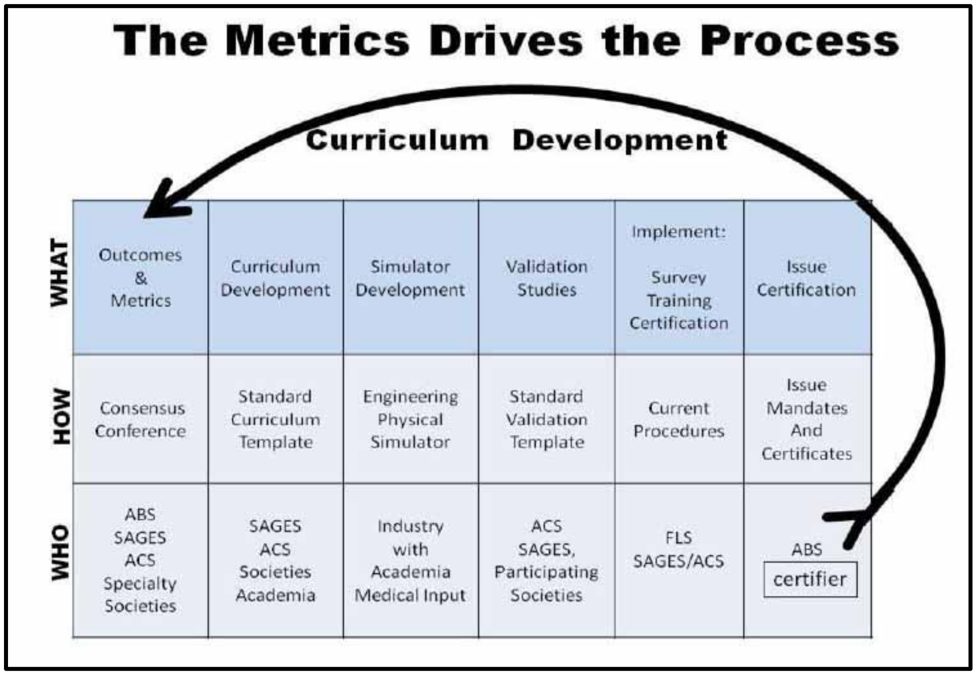
The meeting was held in The Methodist Institute for Technology, Innovation & Education (MITIE). MITIE is a comprehensive, state-of-the-art education and research center at The Methodist Hospital in Houston, Texas. Its educational mission focuses on physicians who wish to acquire new procedural skills and integrate new technologies into their practices. Its research mission is to enhance the use of image guided technology to guide procedures, incorporate robotic surgery into the image guided platform, and develop new technology and procedural techniques.

## Introduction and FRS Project Background – Richard Satava, MD

The Fundamentals of Robotic (FRS) is a joint industry (Intuitive Surgical, Inc) and Department of Defense (DoD) funded project jointly managed through Minimally Invasive Robotics Association (MIRA) and Florida Hospital Nicholson Center for Surgical Advancement (NCSA). The mission of FRS is to create and develop a validated multi-specialty, technical skills competency based curriculum for surgeons to safely and efficiently perform basic robotic-assisted surgery.

The basic curriculum for robotic surgery will be created jointly by multiple surgical specialties that use robotic systems for surgery. The curriculum will be open source and adaptable for many methods of simulation – from physical models to full virtual reality. A shortcoming of previous curricula for simulators was that they were developed in cooperation with a single surgical expert in a single specialty, who alone determined the best method for the skills or procedures; however most often, the resulting curriculum, while clearly validated, did not have the outcomes measures that were acceptable to certifying bodies. To overcome this limitation, all major stakeholders were invited to participate.

The development of the FRS is conceived as a "full life cycle" development of the curriculum (see graphic below).




"Full life cycle" development includes 3 consensus conferences in the following topics:

1. Outcomes Measures : The first FRS consensus conference (FRSCC#1) brought together subject matter experts (SME) from multiple surgical societies, surgical educational societies, surgical boards and other governing organizations who agreed upon the critical skills, tasks, and most common errors that needed to be included in a comprehensive basic curriculum. The result was a table that defined the skills/tasks/errors, the desired outcome measures, and the metric(s) that should be measured. The table was rank ordered both as to sequence in which these occurred, as well as a second table that rank ordered the measurements in terms of their priority.

2. Curriculum Development: The second consensus conference (FRSCC#2) had four specific goals that will lead up to the completion of a curriculum for the FRS and the methods of training and assessing the full range of technical skills (cognitive, psychomotor, team training/communication) that are necessary to safely use a robotic surgery system. The goals are to:
    a. Review the Outcomes Measures Tables
    b. Select from those measures the ones which can be included into the curriculum development (and add any other critical measures that may have been missed)
    c. Review and adapt the curriculum template from ASSET (developed and published a curriculum template with wide consensus for surgical training)
    d. Complete the actual curriculum, including the metrics for each skill/task/error and assessment tools.

    Once this curriculum is completed and accepted by the participants of this conference, it will be distributed for comments.

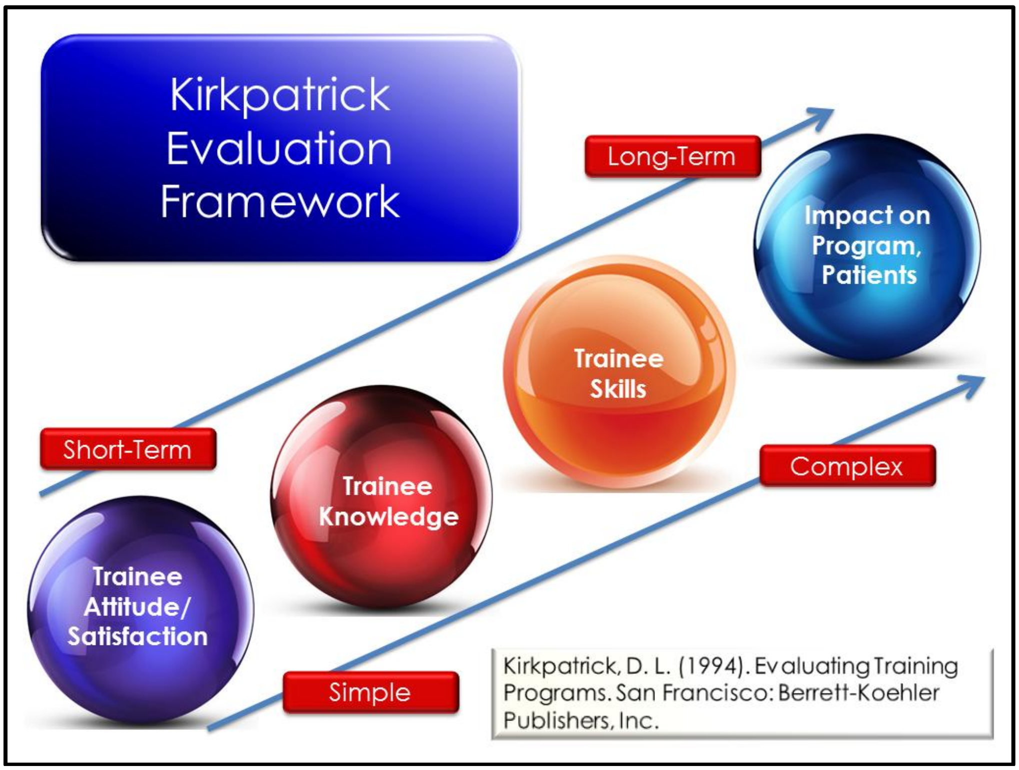
3. Validation Study Design: The third consensus conference will be for the design of the Validation Study, to meet the most rigorous evaluation that would meet criteria for high stakes testing and evaluation.

These conferences will be followed by multi-institutional validation studies with participation by multiple surgical specialties at each institution.

# VALIDATION LECTURES

## Validation: Tools for Improving Curriculum and Assessment by Dr. Sara Kim, ISIS, University of Washington Medical Center

1. Essence of Validation
   a. Collecting evidence from data in order to make an inference about what you are assessing.
   b. Two critical components that affect the quality of evidence you collect include: the range of behaviors you are observing based on tasks and how you measure and score trainees' behaviors.
2. Framework of Validation




   a. Generalization: Generalizability of the scores (association between the score a person receives on an assessment and the universe score—the theoretical score he/she would receive if taking the assessment an infinite number of times), or reliability, must be ascertained keeping the complexity of the assessment in mind.
   b. Extrapolation: Assessment scores are closely linked to the "construct" and there is a relationship between simulation assessment and patient care.

   c. Decision/Interpretation: Need evidence that assessment methods are defensible in interpreting scores and that there is a consequential impact of assessment (i.e. curricular change, healthcare efficiency).
3. Types of Validation Study Designs
   a. Task Based Studies
   b. Proficiency Based Studies
   c. Training Model Based Studies
4. Implications for FRS Validation
   a. Number, type and sequencing of tasks are critical to ensure adequate sampling of behaviors in validation studies.
   b. Paucity of training scenarios of graded complexity and difficulty levels (Sum of technical proficiencies across tasks = competency?)
   c. Team-based approach to robotic surgery should be an integral part of training to avoid collision of "multiple learning curves"
   d. Lack of consistent definitions for novices vs. experts, which impedes cross-study comparisons in validation literature.

## Toward a Course Validation Template by Dr. Wallace Judd, Authentic Testing Corporation

1. Tasks and Skills
   a. Definition of a task: When you do something... there is an outcome.
   b. Don't confuse tasks and skills. The following are skills (not tasks):
      i. Knot tying
      ii. Docking
      iii. Instrument exchange
      iv. Suturing
      v. Multi-arm control
      vi. System settings
      vii. Ergonomic positioning
   c. Task mastery does not equal skill comprehension
   d. Paradigm task: A task that generalizes to other tasks
2. Looking at a Typical Course
   a. Within a course, one conducts a series of skills and verifies each of them
   b. The sum of the skills does not mean there has been mastery (see bike example below)
   c. The sum of the verifications does not mean there has been validation
   d. Bike example
      i. The sum of certain skills including pedaling, braking, steering, and balance do not necessarily mean mastery of bike riding
      ii. Are the skills listed above really comprehensive for bike riding?


3. Validation Tasks <u>Must Be</u> ...
    a. Identical for all candidates
    b. Singular
    c. Reusable
    d. Stable
4. Course Validation Template
    a. Define scope of course
    b. Define blueprint
    c. Determine paradigm tasks
    d. Define skills required to do tasks
    e. Create alternative tasks
    f. Define overlay of complexity
    g. Write items
    h. Define item scoring
    i. Administer test tasks
    j. Rate candidates in course success
    k. Correlate course ←→ test score
    l. Correlation establishes course validity

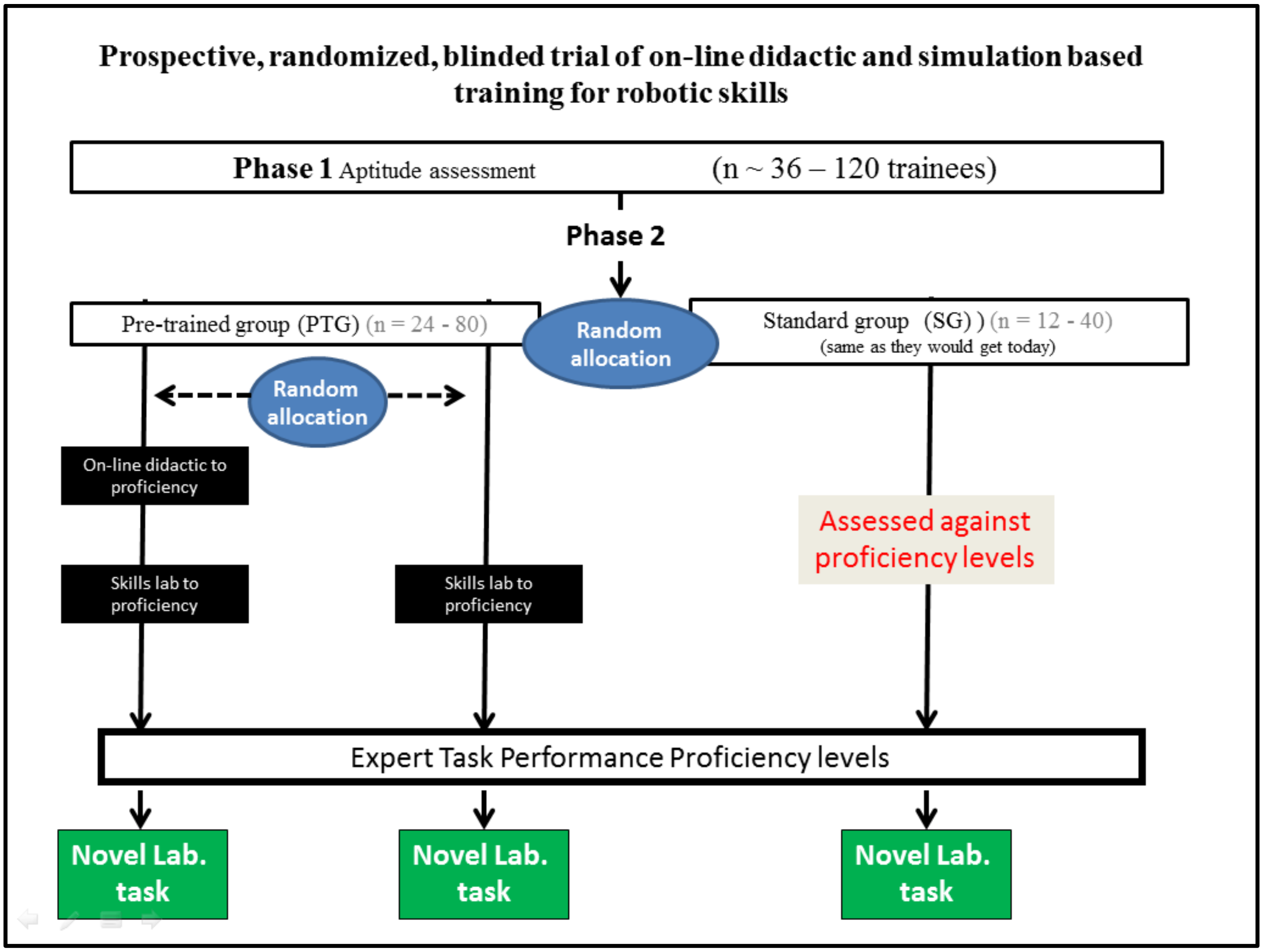
## Validation Study Design & Methods by Dr. Anthony Gallagher, University College Cork

1. Validation is a Process
    a. Validation is about making it easy for individuals reviewing the studies to believe the results (i.e. making it easy for ‘them’ to say yes).

   b. Validation is also about creating compelling enough evidence to sway nay-sayers who are inclined not to believe the results (i.e. making it difficult for ‘them’ to say no)
   c. Minimally invasive surgeons are well versed in the process of validation
2. Metrics for an Optimal Training Program
   a. What to measure?
      i. Time (context dependent)
      ii. Errors (events)
      iii. Performance variability (consistency)
   b. Metrics for ‘proficiency-based progression’ training are not complex. Simply put, metrics are:
      i. What makes you cringe when you see another operator (or yourself) doing something that <u>really</u> shouldn’t be done!
      ii. What you try and train your trainees to do and not to do
3. Validation: Why is it Important?
   a. Different types of validation (Validity of an assessment is the degree to which it measures what it is supposed to measure.)
      i. Face-Content Validity
         - Danny Scott has laid the groundwork for robotic surgery in his article: Dulan G, Rege RV, Hogg DC, et. al. Content and face validity of a comprehensive robotic skills training program for general surgery, urology, and gynecology. *The American Journal of Surgery* Volume 203, Issue 4 , Pages 535-539, April 2012.
      ii. Concurrent Validity
      iii. Construct Validity
         - Danny Scott has laid the groundwork for robotic surgery in his article: Dulan G, Rege RV, Hogg DC. Developing a comprehensive, proficiency-based training program for robotic surgery. *Surgery* Volume 152, Issue 3 , Pages 477-488, September 2012.
         - Showed that tasks that look like they should be appropriate for training robotic skills are rated as such by robotic ‘experts’
         - Also showed they can differentiate between an expert and a novice
      iv. Predictive Validity
   b. Reliability
      i. Describes the overall consistency of a measure. A measure is said to have a high reliability if it produces similar results under consistent conditions.
      ii. Reliability includes test-retest, internal, inter-rater (a discussion for another day)
      iii. If a valid test is unreliable it is almost useless
   c. Assessments should be:
      i. Objective
      ii. Transparent
      iii. Fair
4. What Needs to be Done to Advance the Validation Process?
   a. On-line didactic
      i. Need to demonstrate validity with ‘appropriate’ numbers
         - N = >40 *(estimated from Berryhill et al., Urology 2008 and Gallagher et al. Ann Surg ToT/TER data; more is better*)

b. Skills lab
   i. Need to demonstrate construct validity with 'appropriate' numbers
      - N = >40 (more is better)
   ii. Establish construct validity and thus proficiency
      - What should be done with critical errors (e.g., forgetting about clutching status after resuming from 'parked')?
   iii. Who is going to be the 'guardian' of the proficiency levels?
   iv. What do your proficiency levels mean?
      - Predictive validity

5. Possible Model for Robotic Validation Studies



# MEASURES & METRICS GROUP

## Group Members

*Group Leaders: Brian Dunkin & Martin Martino*
Todd Larson
Tony Jarq
Tim Brand
Rodney Dockter
Naz Siddiqui
Arnie Advincula
Victor Wilcox
Paul Rose
Tom Lendvay
Jan Cannon-Bowers
Wally Judd
Alvin Goh
Paul Neary

## Meeting Goals:

The goal of the Measures and Metrics Group is to answer the question: what are the metrics? The goal is not to decide success/non-success. The system developed should:

- Be easily produced and measured
- Not be reliant on manual counting/validating (electronic metrics are preferable)
- Should take a maximum of 1 hour to complete.

## General Scoring Guidelines for Skill Drills

Adapted from Siddiqui et al. for the Robotics Training Network. Validity and reliability of R-OSATS: a novel assessment tool for robotic surgical training. Presented at the 2013 CREOG and APGO Annual Meeting, Pheonix, AZ.

| Depth Perception/Spatial Orientation/Accuracy | | | | |
|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** |
| Constantly overshoots target, slow to correct | | Some overshooting but quick to correct | | Accurately directs the instruments to target |

| Force/Tissue Handling | | | | |
|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** |
| Breaks model, ring, or suture; damages needle | | Moves or bends wire; minor trauma to model or needle, frays suture | | Handles model, suture, and/or needle well; traction is appropriate |

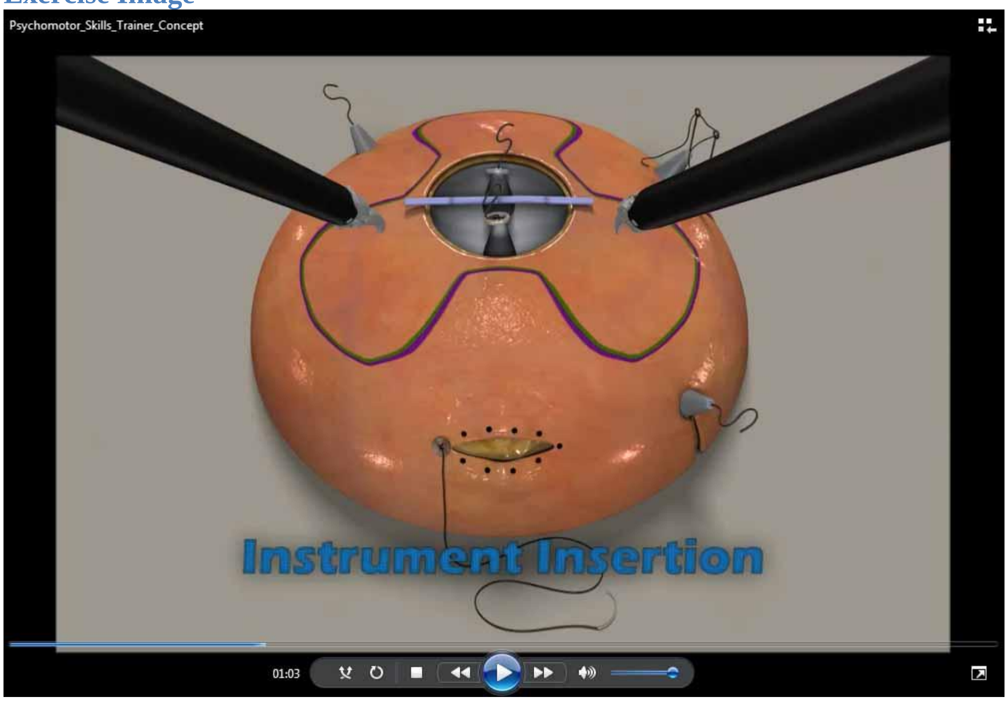
| Dexterity | | | | |
|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** |
| Poor coordination of hands; repetitively drops ring or band; inappropriately drops needle or poor suture management | | Suboptimal interaction between hands, any drops of ring or band. Suboptimal suture or needle management. | | Expertly uses both hands; always transfers rings or bands without dropping. Optimal needle or suture |

management.

| Efficiency | | | | |
|---|---|---|---|---|
| **1** | **2** | **3** | **4** | **5** |
| Uncertain movements with little progress | | Slow, but movements seem reasonably organized | | Confident, fluid progression, adjusts quickly |

## 1) Docking & Instrument Insertion

### Exercise Image




### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception /Spatial Orientation/ Accuracy | 1 | 2 | 3 | 4 | 5 |
| Force of Insertion | 1 | 2 | 3 | 4 | 5 |
| Dexterity (no instrument collision) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (speed of entry - time in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

a. Instrument collisions occur
b. Instrument tips not in view
c. Insertion of instrument not visualized

### Critical/Fatal Errors

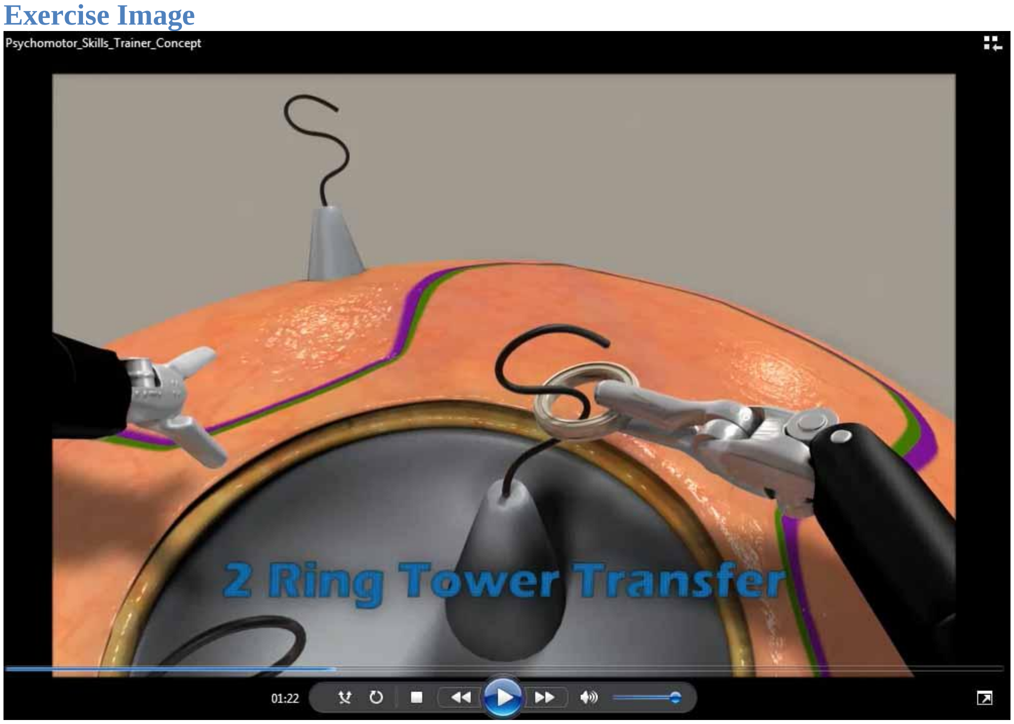
a. Instrument inserted into box
b. Inability to complete the exercise

### Additional Discussion by the Group

- Should the trocars in abdomen be placed in a specific manner? Yes, need standard method. Measure that they can insert/remove in defined position (should be timed)
- How involved is proctor to correct movement before progression? Should have independent steps. Need to be able to reset and continue.
- Fatal errors vs. errors – discuss difference and acceptability. Is that part of test or pre-test practice?
- Time is one variable; touching is another variable to create a metric. Can we use plumb lines, known angles? Where are robot arms? What are the probabilities for internal/external collisions?
- Is camera in sweet spot? Can someone take a picture to determine optimal position? May need different kind of measuring tool – proctor or on-screen.
- Steps to dock (camera and at least two arms): close both flanges; determine pitch and yaw, clutch, and set up joint
- Task time starts from fixed starting position, not necessarily inserting ports at specific angles.

## 2) Ring Tower Transfer

### Exercise Image




### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception/Spatial Orientation /Accuracy | 1 | 2 | 3 | 4 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| Force (ring and contact handling/force) | 1 | 2 | 3 | 4 | 5 |
| Dexterity (movement of rings and hand transfer) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (time in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

- Dropping ring
- Touching contacts
- Failing to transfer hands

### Critical/Fatal Errors

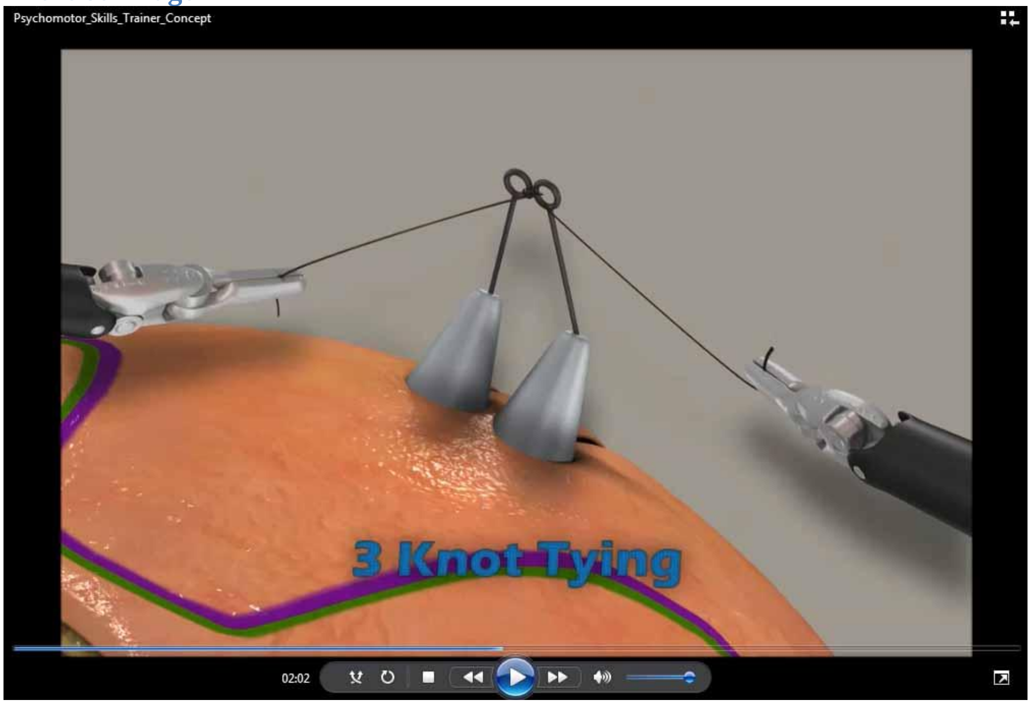
- Losing the ring
- Breaking the ring
- Popping off the wire/tower
- Inability to complete the exercise

### Additional Discussion by the Group

- Different shape tower plug-ins – spiral and non-spiral ring contacts. Shape should be more complicated to force slow down.
- When do you reset? Lost ring? How should we define when a ring is “lost”? Set up should give access to every part of the box to allow for ring retrieval.
- How many ring drops do you count? Clock does not stop during drops. Non-recoverable errors are any that happened outside the box or outside the field of vision (blindspots) inside box.

## 3) Knot Tying

### Exercise Image




### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception /Spatial Orientation/ Accuracy | 1 | 2 | 3 | 4 | 5 |
| Force of knot tying/force on sutures | 1 | 2 | 3 | 4 | 5 |
| Dexterity (two handed knot tying) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (speed measured in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

- Air knot – two rings do not touch
- Breaking suture (standardized 2.0 silk)
- Put dome on contacts so it reports lift BEFORE the suture breaks

### Critical/Fatal Errors

- Knocking off contacts with force
- Inability to complete the exercise

### Additional Discussion by the Group

- Constrained by type of knot: square knot, surgeons knot, slip knot?
- Must do first double through (surgeon's knot) and 2 half hitches (single throw)

## 4) Railroad track

### Exercise Image


### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception /Spatial Orientation/ Accuracy (going through designated circle) | 1 | 2 | 3 | 4 | 5 |
| Force on sutures and tissue | 1 | 2 | 3 | 4 | 5 |
| Dexterity (two handed suturing) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (speed measured in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

- Going outside of designated circles
- Tearing through tissue
- Slack in railroad track suture
- Improper two hand transfer of needle

### Critical/Fatal Errors

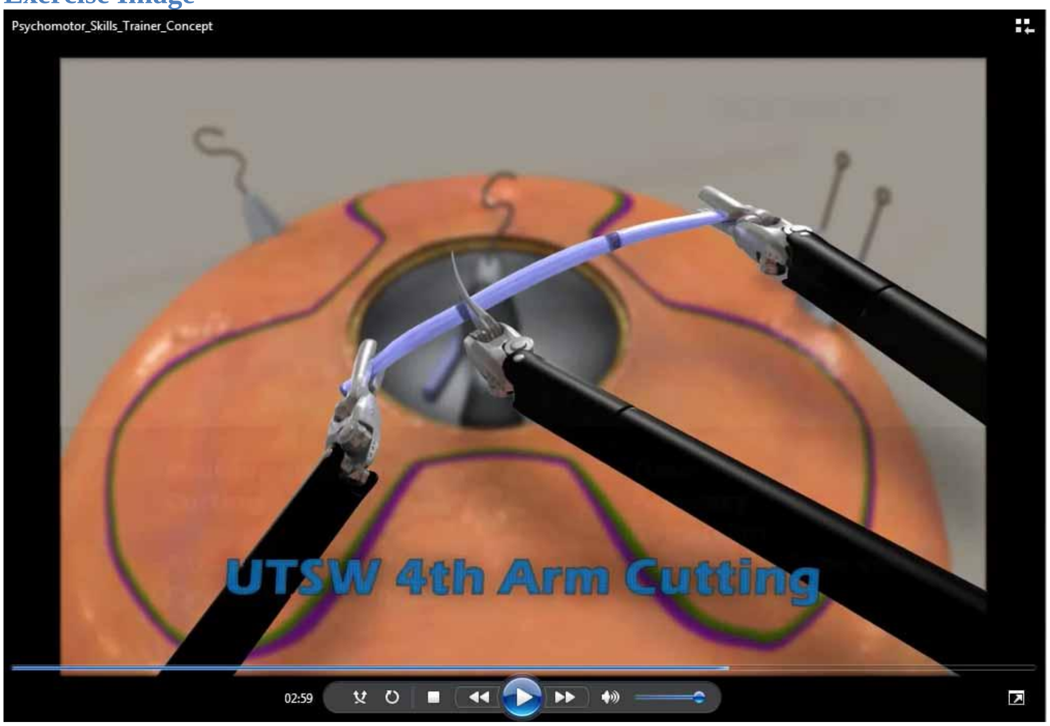
- Break needle
- Inability to complete the exercise

### Additional Discussion by the Group

- Should there be left and right hand throws?  Can there be backhand throws?
- Should the railroad track be in the vertical or horizontal orientation?
- Must come out of the dot (standard size dot and position)
- Measure closure, not the knot
- Should there be a button to start or start with a knot? If test runs long, use the button to start; if time allows, start with a knot

## 5) 4th Arm Cutting

### Exercise Image




### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception /Spatial Orientation/ Accuracy (cutting in the black mark) | 1 | 2 | 3 | 4 | 5 |
| Force on tube/simulated vein | 1 | 2 | 3 | 4 | 5 |
| Dexterity (proper use of 4th arm) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (speed measured in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

- Not cutting in the black mark
- Not visualizing 4th arm before moving

- Breaking/tearing vein

### Critical/Fatal Errors

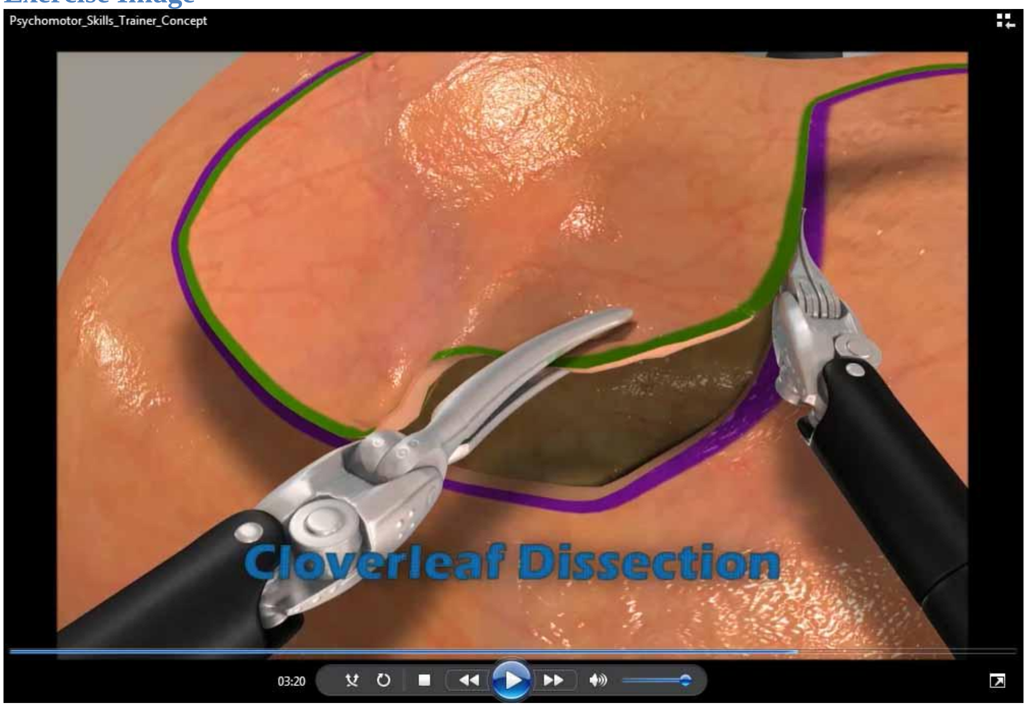
- Inability to complete the exercise

### Additional Discussion by the Group

- Start with 4th arm out of view and then move it into view to begin cutting (Sequence icon to show camera movement before arm? Video recording to capture visualization of 4th arm before moving?)

## 6) Cloverleaf Dissection

### Exercise Image




### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception /Spatial Orientation/ Accuracy (cutting on the line) | 1 | 2 | 3 | 4 | 5 |
| Force on superficial tissue | 1 | 2 | 3 | 4 | 5 |
| Dexterity (Not cutting underlying tissue) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (speed measured in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

- Cutting off line
- Tearing superficial tissue
- Cutting underlying tissue

### Critical/Fatal Errors

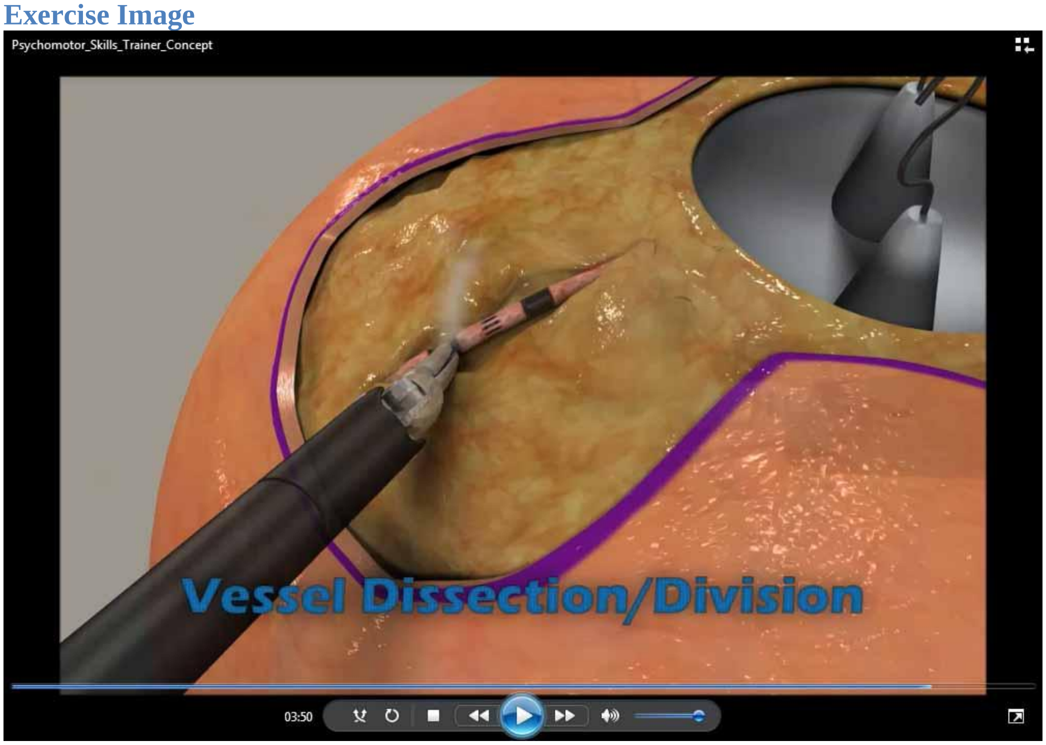
- Inability to complete the exercise

### Additional Discussion by the Group

- Does the cloverleaf shape take too much time? Is it too complicated? The group felt the circle is not complex enough.
- It was discussed to change model to not remove the skin but lift and hold partial dissection with 4th arm. The group decided to only use two arms.
- Make sure to cut the whole skin off. The skin shouldn't stick to the underlying tissue.
- Can you automate the review of the accuracy of cutting the superficial the way that FLS does? But the FLS is not the best way to do it either, this is often said there needs to be a better way.
- How to decide if you lose points for a tear versus a cut. What is the measure that can be used to say you created a tear that was harmful? Is it a real time assessment on video?
- Could you put score lines/semi-perforated lines that would show a standard place to look for tears? The engineers will have to figure out how to do this. The number of tears should be counted.
- Is there a standard nomenclature to refer to the arms, so all the testing centers will be set up in the same way?

## 7) Vessel dissection/Division

### Exercise Image




### Measurements and Metrics

| | | | | | |
|---|---|---|---|---|---|
| Depth Perception /Spatial | 1 | 2 | 3 | 4 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| Orientation/ Accuracy (correct application of energy within black line) | | | | | |
| Force on tissue/vessel for dissection and division | 1 | 2 | 3 | 4 | 5 |
| Dexterity (Not cutting underlying tissue) | 1 | 2 | 3 | 4 | 5 |
| Efficiency (speed measured in seconds) | 1 | 2 | 3 | 4 | 5 |
| Total Score: | | | | | /20 |

### Errors

- Application of bi-polar or mono-polar energy to inappropriate section
- Tearing/dislodging the vessel
- Improper use of foot pedals

### Critical/Fatal Errors

- Inability to complete exercise

### Additional Discussion by the Group

- We should consider a Bluetooth option for pedals and pressurized vessels.
- Will there be use of video icons to tell difference between energy sources, or sensors on pedal to notify proctor?
- Is it possible to get conductivity of materials for signaling?
- Some participants were worried that the cutting task would not capture real knowledge about energy use. Need to link the cutting action to energy use and audio cues.
- Should we have timestamps linked to video?
- Is it worth the added cost to make vessels beating? It is not that expensive and makes the exercise very similar to real surgery.
- The pedal configuration may not always be the same, but the task should be to "push the right pedal at the right time."

## Metrics Discussions Concerning Team Issues for High Stakes Exams (HSE):

The Measurement and Metrics group then began discussing the valid metrics of team training.

### Who is the team?

This section should test the surgeon's ability to work in a "team" that was defined as:

1) The Surgeon (the main focus is on the surgeon)
2) Bed-side assistant
3) Circulator
4) Anesthesiologist
5) Scrub nurse
6) 2nd assistant (after discussion it was decided that since the 2nd assist is not always common for all procedure, for the purposes of FRS, the 2nd assist would not be included)

### What communication needs to be scored?

1) The communication being scored revolves around the commander's intent
2) There should be an agreed upon verbal nomenclature
3) Then the scenarios should be established for the individuals being tested

### Tasks and team training

Do we need to add metrics to each of the 7 tasks that relate to team training?

1) It was decided to leave communication and team management as a separate entity since the team training might be a distraction from the psychomotor skills tasks.

### Communication challenges of the surgeon

When the surgeon is sitting at the console what are some of the things that are frustrating and challenging in communication?

2) Assistants switching in the middle of surgery without communication
3) Relying on verbal communication through speakers
4) Person specific communication impaired/chatter
5) No verbal feedback (need to use personnel names, agreed upon nomenclature, and closed loop communication)
6) Lack of situation awareness/analysis

### Possible Team-based scenarios that can be implemented and tested

Set up scenario so proctor scores on specific behaviors to occur within intra-, inter-, and post-operative times

1) Instrument guided exchange (system has awareness of what is happening-memory will remain if you don't touch the clutch inappropriately)
2) Emergent undocking
3) Movement of bed position (impact on patient and anesthesiologist)
4) Unguided instrument insertion
5) Port problems (hemorrhage and port management)
6) Loss of pneumoperitoneum (ports out)
7) System/robot malfunction (error code)
8) Camera switching

Since there will be 3 parts to the FRS HSE (cognitive, psychomotor skills and team training/communication), the question was raised, "What if the learner does not pass all 3 parts? Can he/she come back and only do that part that wasn't passed?"

### Additional discussion from the Group about team training

- Make a matrix of required measures per metric (automated vs. human elements to scoring)
- Use computer adapted testing – assess online instead of using proctor who may not be as knowledgeable. Keep proctor training to a minimum, but we will need to validate humans vs. computer so we know how productive one is over the other

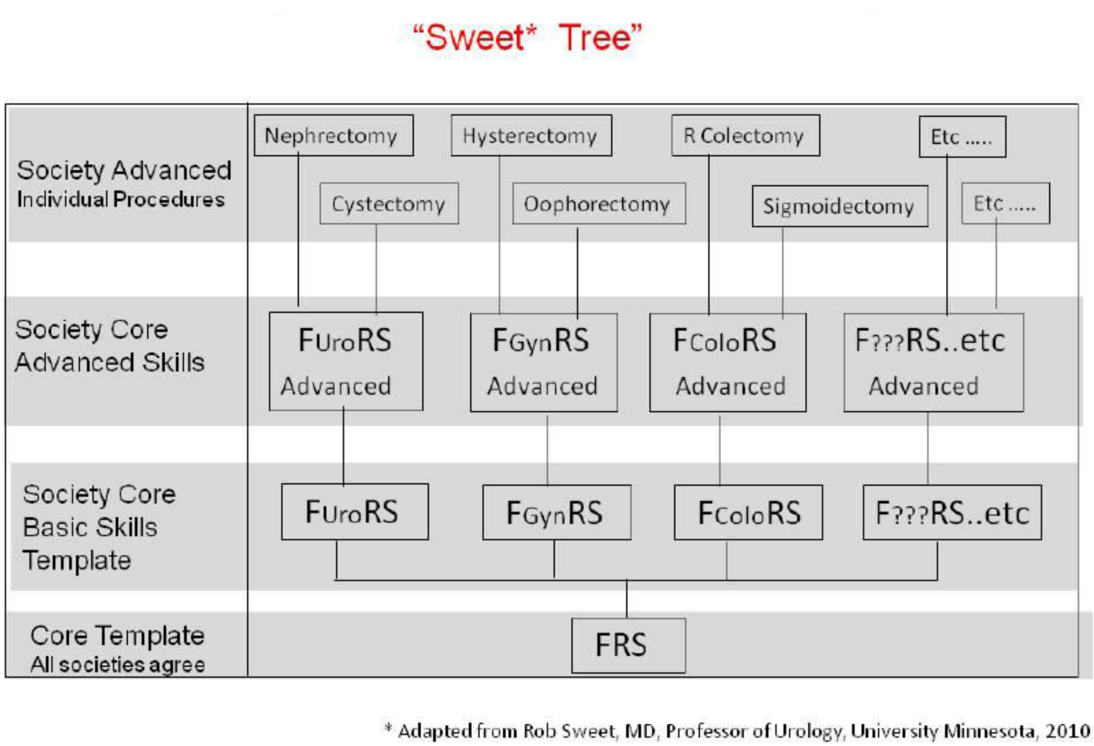
# STUDY DESIGN GROUP

## Group Members

Group Leaders: Jeff Levy & Rob Sweet
Rajesh Aggarwal
Garth Ballantyne
Sanket Chauhan
Tony Gallagher
Jacques Hubert
Sara Kim
Manuela Perez
Cyril Perrenot
Judith Riess
Rob Rush
Brendan Sayers
Mika Sinanan
Roger Smith
Dimitrios Stefanidis

## Goals of the Breakout Group

- Focus on psychomotor component (not talking about curriculum any longer)
- Conceptualize a validation study (or series of studies)
- Look at the model as a simulator
- Focus on the trunk of the Sweet Tree (i.e. focus on the lowest level of the chart below and not on a clinical procedure)



## Define the Research Questions

- Is there a core set of abilities that a robotic surgeon should have?
- Are FLS trained surgeons better on the robot?
- Are there core abilities (validated pre-screen) that predict technical proficiency differences?
- Are their fundamental abilities that impact the safety and effectiveness of a robotic surgeon?
- Are there certain abilities that improve the learning curve for robotic surgeons?
- Is FRS a valid assessment tool?
- Does FRS certification (standards we define) improve (translate to) the clinical performance? (i.e. surgeons trained in FLS made fewer and less severe errors)
- How should the face, content and construct validity and the proficiency level be defined for the benchmark while ensuring inter-rater reliability

## Specific Questions/Hypotheses within Construct Validity

- Do expert robotic surgeons perform above the benchmark performance?
- Does the performance on the FRS model correlate with expertise?
- Novice surgeons do not reach the performance metrics (hypothesis)
- Does PGY level correlate with performance?
- Does previous laparoscopic experience correlate with performance?
- Does previous microsurgery experience correlate with performance?
- Does the type of specialty correlate with performance? (hopefully not)

## Definition of Experts (and other groups):

- Set a minimum required criteria for expert robotic surgeons as a pre-screen
    - More than 5 cases a month
    - More than 50% of cases are completed robotically
    - Clinically active
    - Could have fellowship training in robotics
- Ask experts who have met the required minimum criteria to submit a video of a procedure
- Relatively straightforward multi-specialty tasks for assigning groups (based on expert performance)
- Train the raters of the videos in objective parameters that verify expertise in robotic surgery. This will ensure inter-rater reliability.
- Make sure the raters are blinded
- Focus on safety not "how I do it" (objective criteria)
- Groups will be based on performance metrics rather than PGY levels, or experience levels

## What Will Be Measured?

### List of validation types

- Face
- Content
- Construct
- Concurrent

- Predictive (not until we have a curriculum)

### Reliability
- Inter-rater
- Test-retest

### Other Parameters
- Usability
- Acceptability

## Study Design
- Phase 1: Pilot at Florida Hospital Nicholson Center (logistics and refinements to model)
- Phase 2: Get face and content validity from the society leadership and boards
- Phase 3: Get face, content, and construct validity at test sites and society meetings
- Phase 4a: Get concurrent validity with video correlations
- Phase 4b: Predictive validity – full research study at 10 sites (IRB will be needed for every site)

### Notes
- Curriculum validation will join in phase 4
- Train-the-trainer happens during phase 1-3 in preparation for phase 4

## Pilot Study Protocol
- Settings
    - Set the robot to standardized motion scaling settings
    - The person being tested can change ergonomic related factors
    - The person being tested can warm up for a maximum of one minute after the docking and instrument insertion task
    - There was a discussion if the S, Si, or both could be used for testing purposes. It was determined that using just the Si would be best.
- Number people for the pilot study
    - 10 experts, 10 intermediate, 10 novices (total 30)
- Inclusion criteria:
    - Must be a surgeon or surgeon-in-training
    - Must have done the online course and passed the online test
    - If intermediate/expert, must have Si experience
- Exclusion criteria:
    - Non-surgeon
    - Didn’t complete online course
    - Medical students
    - Failed the online test
    - Experienced only with only S model and no Si experience
- Dimitrios Stefanidis and Sarah Kim volunteered to develop the FRS cognitive test
    - This test will also solve the purpose of face validity of the cognitive portion

## Questions/Purpose for the Pilot Study
1) How long does it take to complete the exercises?
2) Are there differences between skills when using the daVinci S vs. Si models?

    a. The individual participating in the pilot will be asked what model they normally use.
    b. All testing will be done on Si.
3) How many individuals are necessary to administer the final study?
4) Power analysis data to help inform the "n" and resources necessary for construct study (existing studies can inform effect size)
5) Calibrate global rating score
6) Inform the curriculum design
7) Inform "expertise"

### Demographic Data Collected

1) Age
2) Gender
3) Specialty
4) Hand dominance
5) Number of robotic cases
6) Number of robotic cases per month
7) Number of robotic cases in last 6 months
8) Number of robotic cases in the last 6 months that have involved robotic suturing (stapling/clipping/etc)?
9) Length of time in years/months doing robotic surgery
10) Greatest familiarity of robotics system (Si/S/Standard)
11) Involvement in fellowship/resident training including robotics
12) Number of laparoscopic cases
13) Involvement in simulator training (robotics)
    a. Number of hours spent on robotic simulators in the last 6 months (0, 0-10, etc)?
    b. Reason for use robotic simulators (course/warm up/research)
    c. Simulator used most
14) Past/present experience with video games (quantify)
15) Number of years in surgical practice

## Validity Questions

- The questions was raised whether face and content validity questions should be asked after completing all tasks, or after each individual task

### Face Validity:

- Question 1: Does the model appear at face value to represent the skills necessary to safely perform basic robotic surgery?
- If content is valid, is it is representative of the skill?

### Content

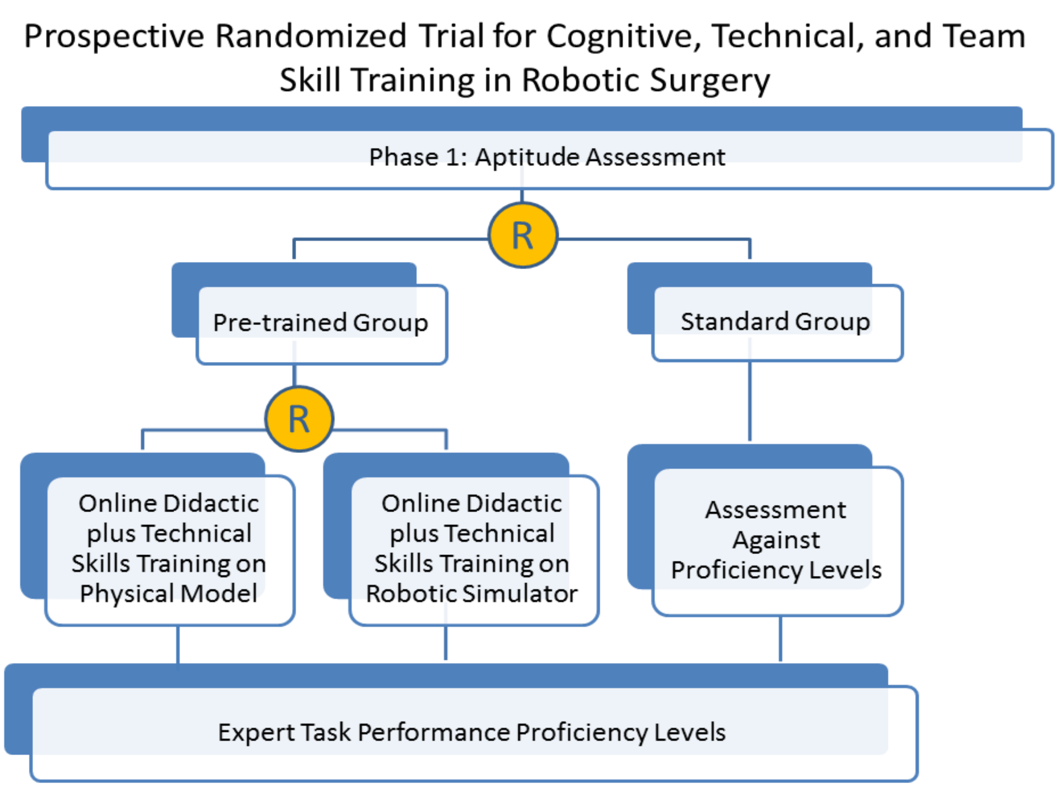
- Write questions that directly address whether each task accomplished the learning objectives that were defined
- Sanket Chauhan volunteered to develop the demographics questions
- Cyril Perrenot and Sara Kim volunteered to develop the face/content validity questions

## Phase 4a Study

### Concurrent study:

- The purpose is to calibrate the definition of expertise and support construct validity/linkage to Global Evaluation and Assessment of Robotic surgery Skills (GEARS)
- Include a specific number of experienced robotic surgeons (from previous phase)
  - They submit video (short segment).
  - They complete the questionnaire (aptitude test)
  - They perform tasks on simulator (dome)
  - The video is graded on GEARS
- Determine the correlation between simulator performance and GEARS score
- The actual procedure doesn't matter, but needs to have a "suturing" element
  - A 5 minute clip must be submitted
- Raters
  - Non-surgeon
  - Review two throws
  - There is a distinct starting and stopping point
  - Use GEARS evaluation

### Study design review




### Description of component of the study design

1) Phase 1: Aptitude test baseline assessment (all novices)
2) Randomize (R) main group

    a. Pre-trained group (see below)
    b. Standard group (same training as they would normally get today)
3) Randomize (R) Pre-trained group
    a. Online Didactic (Cognitive, Team Training) plus Technical Skills Training on physical model (dome)
    b. Online Didactic (Cognitive, Team Training) plus Technical Skills Training on robotic simulator
4) All three groups (2 pre-trained and one standard) must demonstrate performance proficiency levels

### Criteria for institutions participating in Phase 4 studies

1) ACS/AEI accredited Institutes
2) Access to large number of subjects
3) Access to surgeons who are novice robotic surgeons with an interested in learning robotic surgery
    a. Minimum of seven robotic surgeons from at least 3 specialties involved
4) Support staff
    a. Staff familiar with behavior data collection and study design management
    b. Dedicated fellow/coordinator
    c. Staff acquainted with simulator and have on the spot Tech support
    d. Administrative support to help with IRB requirements (at least average IRB turn-around time)
5) Access to Si robot
6) Access to robotic simulator
7) Proven academic track record in surgical education record

# OPEN FORUM NOTES

Following the main meeting, the floor was opened to everyone including industry to provide input and ask any questions. This was not part of the main meeting and had nothing to do with curriculum development, so there is no real or perceived conflict (bias) from industry.

### General comments/questions included:

- A pilot study will be needed for the simulator too
- A simulator cannot do everything all at once as prescribed by the physical dome model
    - There are some limitations due to processing power
    - Tasks on a simulator, however, can be done individually and then the trainee would move on to the next task
- Need to collect some data for simulation in addition to the model
- There were doubts about the cost of the physical dome model being below $500
    - Need to determine what instruments and how many instruments are going to be used to determine costs
    - Remember, the main purpose of the physical dome model is for assessment, not training
    - The robotic simulators will probably be the most likely training model, along with lower fidelity physical models
- Materials on the currently conceived physical dome model can't simulate live tissues exactly, but can come close
- The goal of FRS is to be open source, but this might have issues from one model to another between companies. It is important to have tight standards to ensure validity.
- There may not be a need for a physical model in the future. The simulator is very robust.
- Are the simulation companies going to work in parallel to build simulated based training for high stakes examinations?

# APPENDIX A: PARTICIPANT DIRECTORY


### Arnold Advincula, MD, FACOG, FACS
*Professor of Obstetrics & Gynecology,*
*University of Central Florida College of Medicine*
Florida Hospital Celebration Health
400 Celebration Place
Celebration, FL 34747
USA
407-303-4939
arnold.advincula.md@flhosp.org

### Rajesh Aggarwal, MD, PhD, MA, FRCS
*Instructor in Surgery*
University of Pennsylvania
3400 Spruce Street
Philadelphia, PA 19104
USA
215-385-6549
rajesh.aggarwal@uphs.upenn.edu

### Garth H. Ballantyne, MD, MBA
*Chief of Surgery*
Charlie Norwood VA Medical Center
2926 Bransford Road
Augusta, GA 30904
USA
706-830-2025
garth@lapsurgery.com

### Timothy C. Brand, MD, FACS
*Chair, Robotics Steering Committee*
Madigan
9040 A Reid St
Tacoma, WA 98431
USA
253-330-1718
timothy.c.brand@us.army.mil

### Janis Cannon-Bowers, PhD
*Director of Research*
Center for Advanced Medical Learning & Simulation
USF Health
124 S Franklin Street
Tampa, FL 33602
USA
321-217-0986
jancb16@gmail.com, jcannonb@health.usf.edu

### Sanket Chauhan, MD
*Simulation Fellow*
SimPORTAL
420 Delaware st SE
Minneapolis, MN 55455
USA
612-624-SIMS
sanket@umn.edu

**Rodney L. Dockter**
*Graduate Student*
University of Minnesota
Dept of Mechanical Engineering
309 8th Street, SE
Minneapolis, MN 55455
USA
651-587-1557
dockt036@umn.edu

**Brian J. Dunkin, MD, FACS**
*Professor of Clinical Surgery, Weill Cornell Medical College; John F., Jr. and Carolyn Bookout Chair in Surgical Innovation and Technology; Head, Section of Endoscopic Surgery; Medical Director, MITIESM*
The Methodist Hospital, MITIE
6550 Fannin Street, Suite 1661A
Houston, TX 77030
USA
713-441-6382
bjdunkin@tmhs.org

**Anthony G. Gallagher, PhD, DSc**
*Professor*
University College Cork
School of Medicine
Brookfield Health Sciences Complex
College Road
Cork, Cork
Ireland
353877932465
anthonyg.ggallagher@btinternet.com

**Alvin C. Goh, MD**
*Director of Advanced Laparoscopic & Robotic Urologic Programs*
The Methodist Hospital, MITIE
Department of Urology
6560 Fannin Street, Suite 2100
Houston, TX 77030
USA
713-441-6455
agoh@tmhs.org

**Jacques Hubert, MD**
*Professor*
Urology Dept & Robotic School
CHU Nancy Brabois
Allee du Morvan
Vandoeuvre les Nancy, Lorraine 54511
France
+33 (0)3 83 15 36 94
j.hubert@chu-nancy.fr

**Wallace Judd, PhD**
*President*
Authentic Testing Corp.
932 Edwards Ferry Rd NE #106
Leesburg, VA 20176
USA

530-559-1478
WJudd@AuthenticTesting.com

**Sara Kim, PhD**
*Director of Educational Innovations and Strategic Programs*
University of Washington
Box 356410 1959 NE Pacific, Surgery Pavilion, Room SP-1120
School of Medicine, University of Washington
Seattle, WA 98195
US
310-428-9634
sarakim@uw.edu

**Thomas S. Lendvay, MD, FACS**
*Associate Professor; Co-Director, Seattle Children's Robotics Surgery Center*
University of Washington Seattle Children's Hospital
Seattle Children's Hospital
4800 Sand Point Way, NE
Seattle, WA 98102
USA
206-987-1461
thomas.lendvay@seattlechildrens.org

**Jeffrey S. Levy, MD**
*CEO*
CaseNetwork
7 Ashbrook Lane
Newtown Square, PA 19073
USA
215-356-2174
jlevy@casenetwork.com

**Michael Marohn, MD**
*Associate Professor-FAC*
John Hopkins University
School of Medicine
Department SOM Sur Ravitch Service, Halsted 608
Baltimore, MD 21205
USA
410-502-0723
mmarohn1@jhmi.edu

**Martin A. Martino, MD, FACOG**
*Medical Director, Minimally Invasive Robotic Surgery*
Lehigh Valley Health Network
1240 S Cedar Crest Blvd, Suite 403
Allentown, PA 18103
USA
610-402-3670
Medmartino@gmail.com

**Paul C. Neary, MD, FRCSI**
*Vice Dean Surgcial Training RCSI; Senior Lecturer RCSI and TCD; Head of Colorectal Surgery AMNCH, Dublin, Ireland*
Tallaght Hospital

Tallaght
Dublin, Dublin D24
Ireland
011 353 41412212
paulcneary@msn.com

### Perez Manuela, MD, PhD
*Doctor*
University Hospital Nancy
Avenue de Strasbourg
Nancy, 54000
France
+33 (0)3 83 85 98 59
m.perez@chu-nancy.fr

### Cyril Perrenot, MD
*Resident in General Surgery in charge of Robotic Simulation Training at School of Surgery*
School of Surgery of Nancy
Faculté de Médecine de Nancy - Bât D - 2ème étage
9, Avenue de la Forêt de Haye
Vandoeuvre-les-Nancy, 54505
France
+33 (0)6 89 20 94 67
perrenotcyril@gmail.com

### Judith S. Riess, PhD
*CEO, Education & Training Solutions*
Education & Training Solutions
P.O. Box 720656
Orlando, FL 32872-0656
USA
703-395-6608
judithriess@education-trainingsolutions.com

### Col. Robert Rush, MD
*Department of Surgery Chief*
Madigan Healthcare System
Madigan Army Medical Center
& Army Central Simulation Committee
Tacoma, WA 98431
USA
360-556-2838
robert.rush1@us.army.mil

### Richard Satava, MD, FACS
**Principle Investigator**
*Professor of Surgery*
University of Washington
Dept of Surgery, Mailstop 356410
1959 NE Pacific St, 9040 Fitzsimmons Dr.
Seattle, WA 98195
USA
206-616-2250
rsatava@u.washington.edu

### Brendan G. Sayers, PA-C
*Robotic Surgery Physician Assistant*
UT Southwestern Medical Center

5323 Harry Hines Blvd.
Southwestern Center for Minimally Invasive Surgery
L1.100
Dallas, TX 75390-9092
USA
214-648-2186
Brendan.Sayers@UTSouthwestern.edu

## Nazema Y. Siddiqui, MD, MHSc
*Assistant Professor, Urogynecology & Reconstructive Pelvic Surgery*
Duke University Medical Center
DUMC 3192
Durham, NC 27710
USA
919-401-1006
nazema.siddiqui@duke.edu

## Mika N. Sinanan, MD, PhD
*Professor of Surgery President, University of Washington Physicians*
University of Washington
BB430, 1959 N.E. Pacific St.
Seattle, WA 98195
USA
206-543-5511
mssurg@uw.edu

## Roger D. Smith, PhD
**Principle Investigator**
*Chief Technology Officer*
Florida Hospital Nicholson Center
404 Celebration Place
Celeration, FL 34747
USA
407-303-4353
roger.smith@flhosp.org

## Dimitrios Stefanidis, MD, PhD, FACS
*Medical Director, Carolinas Simulation Center*
Carolinas Healthcare System ALSO representing Association for Surgical Education
1025 Morehead Medical Dr Suite 300
Charlotte, NC 28204
USA
704-355-5114
dimitrios.stefanidis@carolinas.org

## Robert M. Sweet, MD, FACS
*Associate Professor Urology*
*Joseph Sorkness Endowed Professor*
*Director of Medical School Simulation Programs*
*Director of the Kidney Stone Clinic*
University of Minnesota
420 Delaware St SE
MMC 394, A590-2
Minneapolis, MN 55455
USA
952-913-0770
rsweet@umn.edu

**Victor T. Wilcox, MD**
*Research Fellow*
The Methodist Hospital, MITIE
6560 Fannin Street, Suite 2100
Houston, TX 77030
USA
713-441-6455
vwilcox@tmhs.org

## MIRA/FRS STAFF

**Adeena N. Bleich**
*Program Manager*
MIRA/FRS
11300 W. Olympic Blvd, Suite 600
Los Angeles, CA 90064
USA
310-437-0555, ext. 141
adeena@roboticfundamentals.org

**Erin Drummond**
*Administrative Assistant*
MIRA/FRS
11300 W. Olympic Blvd, Suite 600
Los Angeles, CA 90064
USA
310-437-0555, x.144
erind@roboticfundamentals.org