Testing the Construct Validity of A Virtual Reality Hip Arthroscopy Simulator

Vikas Khanduja MA, MSc, FRCS (Orth)¹, John E Lawrence MA MB BChir MRCS¹, Emmanuel Audenaert MD PhD²

1 Department of Trauma & Orthopaedics, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 0QQ, UK.
2 Department of Orthopaedic Surgery and Traumatology, Ghent University Hospital, De Pintelaan 185, 9000 Gent, Belgium.

Corresponding Author:
Mr Vikas Khanduja MA, MSc, FRCS (Orth),
Consultant Orthopaedic Surgeon & Associate Lecturer,
Department of Trauma & Orthopaedics,
Addenbrooke’s – Cambridge University Hospital,
Hills Road, Cambridge CB2 0QQ, UK.
Email: vk279@cam.ac.uk
Tel: +44 1223 257093

Running Title: Validating a hip arthroscopy simulator
Abstract

Purpose

The purpose of this study was to test the construct validity of the hip diagnostics module of a virtual reality hip arthroscopy simulator.

Methods

Nineteen orthopaedic surgeons performed a simulated arthroscopic examination of a healthy hip joint using a 70 degree arthroscope in the supine position. Surgeons were categorized as either expert (those who had performed 250 hip arthroscopies or more) or novice (those who had performed fewer than this). Twenty-one specific targets were visualized within the central and peripheral compartments; nine via the anterior portal, nine via the anterolateral portal and three via the posterolateral portal. This was immediately followed by a task testing basic probe examination of the joint in which a series of eight targets were probed via the anterolateral portal. During the tasks, the surgeon’s performance was evaluated by the simulator using a set of pre-defined metrics including task duration, number of soft tissue & bone collisions, and distance travelled by instruments. No repeat attempts at the tasks were permitted. Construct validity was then evaluated by comparing novice and expert group performance metrics over the two tasks using the Mann–Whitney test, with a p value of less than 0.05 considered significant.

Results
On the visualization task, the expert group outperformed the novice group on time taken (P=0.0003), number of collisions with soft tissue (P=0.001), number of collisions with bone (P=0.002) and distance travelled by the arthroscope (P=0.02). On the probe examination, the two groups differed only in the time taken to complete the task (P=0.025) with no significant difference in other metrics.

**Conclusions**

Increased experience in hip arthroscopy was reflected by significantly better performance on the VR simulator across two tasks, supporting its construct validity.

**Clinical Relevance**

This study validates a virtual reality hip arthroscopy simulator and supports its potential for developing basic arthroscopic skills.

**Introduction**

With an ever-increasing number of diagnostic and therapeutic indications, hip arthroscopy is one of the most rapidly expanding areas of modern surgery [1-4]. The ball and socket nature of the joint, together with the thickness of the capsule and soft tissue envelope make it a technically demanding field with a steep learning curve [2]. An estimated thirty training cases are required before a reduction in complications is
observed [5]. Such complications include labral or chondral injury from instrumentation,
bleeding or nerve injury from portal insertion and traction-related neurapraxia [6]. It is
therefore vital that orthopaedic surgeons gain exposure to arthroscopic hip surgery early
in their training and quickly develop the basic skill set necessary to progress towards
competency.

Although the traditional model of the apprenticeship in surgical training still applies
today, the evolution of more technically demanding surgical techniques combined with a
reduction in the trainee's caseload has led to steep learning curves in the modern era [7-
11]. Several studies have highlighted the length of time taken to become a safe and
effective arthroscopic surgeon, demonstrating both increased operative time and a
higher incidence of complication when an inexperienced clinician is performing
arthroscopic hip surgery [12-15].

Multiple studies have also suggested a role for virtual reality (VR) simulation in
overcoming this steep learning curve by equipping the orthopaedic trainee with the basic
skill set required to safely perform arthroscopic procedures [16-18]. In the context of hip
arthroscopy, improving skills with simulation prior to performing a procedure in the
operating room has the potential to reduce those complications more closely associated
with surgical technique, such as labral and chondral injury. Validation of a virtual reality
hip arthroscopy simulator as an evaluation tool is required before it can be used to this
end.
The purpose of this study was to test the construct validity of the hip diagnostics module of a virtual reality hip arthroscopy simulator by comparing novice and expert performance over two tasks of increasing difficulty. The hypothesis was that expert arthroscopic hip surgeons would achieve superior performance on the simulator when compared to those with little experience, supporting the simulator’s construct validity.

**Materials and Methods**

For this study, orthopaedic surgeons were voluntarily recruited following one day of a cadaveric training course in arthroscopic hip surgery for orthopaedic residents in December 2015. The study was open to all surgeons participating in the course. This included the surgical residents (surgeons at various stages of a six year speciality training programme in orthopaedic surgery) as well as course faculty members (practicing orthopaedic surgeons who have successfully completed the six year training programme and gained certification as orthopaedic specialists). The first day of the course covered the equipment used in arthroscopy of the hip joint (including the 70 degree arthroscope and probe set), portal placement in the supine and lateral positions, applied anatomy of the hip joint and finally visualization of the central and peripheral compartments in cadaveric specimens. Basic demographic information was provided by each volunteer including gender, age, height and glove size as well as detailed information on their surgical experience and any previous use of VR simulation (Table 1). A cut-off of 250 independent hip arthroscopies was utilized to classify participants as novice (less than 250 arthroscopies) or expert (250 arthroscopies or more). This number was taken from...
The competency criteria published by the German Speaking Society of Arthroscopy; the only formal published criteria for proficiency in arthroscopy [19].

The Simbionix Arthro Mentor™ (Simbionix Ltd, Airport City, Israel) virtual reality simulator was utilized for this study. This simulator consists of a mannequin with fixed 5mm portals at the anterior, anterolateral and posterolateral sites, a 70 degree arthroscope, a rounded 5mm blunt tip probe, a computer and a monitor which produces three-dimensional images in response to the actions of the operator. The simulator provides haptic feedback for the operator via motors connected to the instruments (Figure 1). Each participant was asked to perform two consecutive simulated procedures in the supine position; one basic visualization task and one basic probe examination of the joint (video 1). Each task began with identical written instructions on-screen, with the procedure time starting on insertion of the arthroscope into any of the three portals (anterior, anterolateral and posterolateral). The visualization task involved locating a series of twenty-one targets using three arthroscopy portals. Nine targets were visualized via the anterolateral portal, nine via the anterior and three via the posterolateral (Table 2). Targets were located on specified areas of a healthy hip joint, and participants were required to place them in clear view on the center of the monitor for three seconds before the simulator selected the next target. The order in which targets appeared remained the same for each participant.

The basic probe examination involved visualizing a series of eight targets via the anterolateral portal and achieving sustained contact with them for three seconds using a
probe inserted through the portal of the participant’s choice (Figure 2). As with the visualization task the order of targets remained the same for each participant (see table 3). During the tasks, the surgeon’s performance was evaluated by the simulator using a set of pre-defined metrics: procedure duration, time elapsed between targets, number of soft tissue & bone collisions, total length of any femoral head scratches performed by the instruments, camera-tissue contact time and distance travelled by instruments.

Statistical analysis was performed using version 3.2 of R (Foundation for Statistical Computing, Vienna, Austria). Differences in the performance metrics between the expert and novice groups were compared using the Mann–Whitney test, with a p value of less than 0.05 considered significant. Normality of distribution was tested with the Shapiro-Wilk test. Correlation between participant variables (years of surgical experience and number of hip arthroscopies performed) and outcome measures was tested using Pearson product-moment correlation coefficient.

**Results**

Ten surgical residents (surgeons at various stages of a six year speciality training programme in orthopaedic surgery) and nine faculty members (practicing orthopaedic surgeons who have successfully completed the six year training programme and gained certification as orthopaedic specialists) volunteered for the study. All nine faculty members had performed 250 hip arthroscopies or more, forming the expert group, with all ten surgical residents performing less than this number, forming the novice group. All
but one of the expert group had performed over 500 hip arthroscopies, in contrast to the
novice group in which only one surgeon had performed ten or more hip arthroscopies.
None of the participants had previous experience of using a hip arthroscopy simulator.
There were no exclusions from analysis.

On the visualization task, the expert group significantly outperformed the novice group
as measured by the time taken to complete the task (p=0.0003), number collisions with
soft tissue (p=0.001), number of collisions with bone (p=0.002), total contact time
between the arthroscope and tissues (p=0.001) and distance travelled by the
arthroscope (p=0.023) (Table 4). There was no significant difference in total length of
femoral head scratches (p=0.10) or camera steadiness (p=0.45). On the basic probe
examination, the two groups differed only in time taken to complete the task (p=0.025)
and total contact time between the arthroscope and tissues (p=0.014), with no
difference noted in any other outcome measures (Table 5).

On the visualization task, there was moderate inverse correlation between years of
surgical experience and the number of tissue collisions ($r^2=-0.68$), and moderate
correlation between the number of previous hip arthroscopies performed and time taken
to complete the task ($r^2=-0.62$). On the basic probe examination, there was no
correlation between participant variables, including number of previous hip
arthroscopies performed and number of hip arthroscopies performed per week, and
outcome measures.
Discussion

These results demonstrate a statistically significant difference between the performance of experts and novices on a virtual reality hip arthroscopy simulator, with experts performing superiorly across two consecutive tasks. This implies that skill in the operating room is reflected on the simulator, supporting the simulator’s construct validity. These results mimic those of similar studies relating to the use of virtual reality simulation in orthopaedic training, and support the hip arthroscopy simulator’s use in helping trainees gain the basic skill set necessary to train in hip arthroscopy [20-25].

For simulation to have maximal benefit in the clinical setting, it should achieve not only a technical familiarity with hip arthroscopy, but also a familiarity with basic procedures. In this regard, evidence supporting VR simulation in training orthopaedic surgeons is still minimal, with few studies of arthroscopy simulation showing the benefits to the trainee in the operating room. A study by Howells et al. in 2008 showed trainees randomised to receive simulator training in addition to their standard surgical curriculum achieved significantly improved competence scores when performing a diagnostic knee arthroscopy in the operating room [26]. A subsequent study by Cannon et al. in 2014 had similar findings, with trainees randomised to receive simulator training showing superior probing skills when performing diagnostic knee arthroscopy in the operating room (OR), although procedure duration and visualization skills did not differ between the two groups [27]. A 2015 study by Angelo et al. showed increased likelihood of proficiency in performing an arthroscopic Bankart repair at the completion of training when trainees
were randomised to receive simulation training in addition to the standard curriculum [28]. Although these studies are encouraging, this relative lack of evidence contrasts with other surgical specialties such as general surgery, in which a much larger number of studies have demonstrated the positive effects of simulator training on operating room performance [29-35]. If the orthopaedic community is to commit to the use of virtual reality training, clinicians must continue to work closely with technology companies to both develop VR simulation and investigate the effects of simulator use on trainee performance across a broad range of clinical settings, including arthroscopic hip surgery.

One aspect of VR simulation that may limit its application is the cost of purchasing and maintaining a simulator. This cost may be prohibitive for some surgical units and prevent simulation from entering the curriculum on a national level. However, health-economic evaluations of simulation as a training method have shown them to be cost-effective when compared with traditional methods of training [36-38]. This suggests that VR simulators that have construct validity and a proven benefit to the trainee in the clinical environment represent a sensible investment for the training of surgeons.

Limitations

One limitation in this study is that it was performed over the course of a single day. A more precise analysis of the benefits in terms of skill acquisition could be achieved by repeating a single task multiple times with a fixed interval between attempts. This would
provide more information on skill retention and learning curves over time; an important factor when considering the utility of VR simulator training [39].

Unlike some previous studies of simulation, the number of previous procedures and years of experience showed only moderate correlation with simulator performance metrics [40]. This is in keeping with suggestions from previous studies that beyond the initial learning curve of arthroscopy, a ‘plateau’ phase is reached where performance is more susceptible to natural variation, preventing any obvious trends from being observed [11]. However, this study is limited in that sample size was dictated by the number of surgeons attending the course and their willingness to participate, rather than by performing A priori power calculations. Furthermore, the novice group had almost no experience of arthroscopy. Studies involving larger numbers of participants with a wide spread of experience would be useful in exploring these trends further.

When designing the study, we opted to use the competency criteria published by the German Speaking Society of Arthroscopy (250 independent arthroscopies) as a cut-off value that defined a surgeon as an expert. However, these criteria are currently used with reference to knee arthroscopy only, and cannot be assumed to apply to arthroscopy of the hip joint. Unfortunately there are no formal criteria for defining an expert in hip arthroscopy.
While we are able to objectively support the use of virtual reality simulation to develop the basic skill set necessary for hip arthroscopy, it is more challenging to evaluate the role of the simulator in familiarizing the trainee with particular procedures. A limitation of this study in this respect is that the tasks performed did not fully replicate the basic arthroscopic examination. Although the targets for visualization and probing were located in clearly defined and appropriate anatomical areas, the fixed order in which they appeared was not in keeping with the sequence of an arthroscopic examination. Furthermore, the data analysis performed by the simulator is limited in that it does not isolate particular parts of the task a participant found challenging. For example, it would be valuable for a trainee to discover in a simulated setting that they were most likely to cause tissue damage whilst visualising the posterior labrum.

Whilst the images and movements produced by the simulator were highly realistic, the haptic feedback was less so. This was reflected in the high number of tissue and bone collisions recorded during the tasks, which were far greater than one would expect during arthroscopy in the clinical setting. Although these observations mean this simulator is likely to lack face validity, we were unable to gain sufficient feedback from participants to draw conclusions to this end.

**Conclusion**

Increased experience in hip arthroscopy was reflected by significantly better performance on the VR simulator across two tasks, supporting its construct validity.
References


Figure Legends:

Figure 1. User interface on the Simbionix Arthro Mentor™

Figure 2. Screenshots of the visualization task (left) and the probe task (right).

Video 1 - Video clip of the visual examination and basic probe examination tasks.

Table 1. Summary of participant demographics

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean age (years)</th>
<th>Mean Height (cm)</th>
<th>Gender (male : female)</th>
<th>Total years surgical practice</th>
<th>Mean number of hip arthroscopies performed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced</td>
<td>46.7 ±5.7</td>
<td>180.5 ±7.7</td>
<td>9:0</td>
<td>20.1 ±4.7</td>
<td>911.1 ±492.9</td>
</tr>
<tr>
<td>Novice</td>
<td>38.5 ±6.8</td>
<td>177.7 ±6.2</td>
<td>10:0</td>
<td>8.7 ±6.0</td>
<td>9.1 ±26.9</td>
</tr>
<tr>
<td>Portal site</td>
<td>Targets to be visualized during task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior</td>
<td>Acetabular fossa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ligamentum Teres</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posterior Transverse Ligament</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anterior Transverse Ligament</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posterior Labrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anterior Labrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Superior Labrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anterior Acetabulum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posterior-Superior Capsule</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterolateral</td>
<td>Acetabular fossa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posterior Acetabulum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ligamentum Teres</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posterior Labrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anterior Labrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Superior Labrum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anterior Triangle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Posterior Capsule</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Femoral Head</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior</td>
<td>Weight-Bearing Acetabulum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2. List of targets to be visualized from each portal in the visualization task.

<table>
<thead>
<tr>
<th>Portal site</th>
<th>Targets to be probed during task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterolateral</td>
<td>Acetabular Fossa</td>
</tr>
<tr>
<td></td>
<td>Posterior Acetabulum</td>
</tr>
<tr>
<td></td>
<td>Ligamentum Teres</td>
</tr>
<tr>
<td></td>
<td>Anterior Triangle</td>
</tr>
<tr>
<td></td>
<td>Anterior Paralabral Sulcus</td>
</tr>
<tr>
<td></td>
<td>Anterior Labrum</td>
</tr>
<tr>
<td></td>
<td>Posterior Capsule</td>
</tr>
<tr>
<td></td>
<td>Femoral Head</td>
</tr>
</tbody>
</table>

### Table 3. Targets to be probed during the basic probe examination task.

<table>
<thead>
<tr>
<th>Time Taken (mins)</th>
<th>Soft tissue collisions (n)</th>
<th>Bone collisions (n)</th>
<th>Camera-tissue contact time (mins)</th>
<th>Distance travelled by arthroscope (cm)</th>
<th>Length of femoral head scratches (mm)</th>
</tr>
</thead>
</table>
### Table 4 – Results of the basic visualisation task

<table>
<thead>
<tr>
<th></th>
<th>Time Taken (mins)</th>
<th>Soft tissue collisions (n)</th>
<th>Bone collisions (n)</th>
<th>Camera-tissue contact time (mins)</th>
<th>Distance travelled by arthroscope (cm)</th>
<th>Length of femoral head scratches (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced</td>
<td>10.6±6.3</td>
<td>47.2±37.9</td>
<td>35.3±31.3</td>
<td>5.4±2.9</td>
<td>272.1±201.3</td>
<td>288.3±238.5</td>
</tr>
<tr>
<td>Novice</td>
<td>19.4±10.3</td>
<td>88.3±68.3</td>
<td>63.8±46.5</td>
<td>13.7±9.0</td>
<td>420.9±250.7</td>
<td>574.6±655.1</td>
</tr>
<tr>
<td>P Value</td>
<td>0.025</td>
<td>0.072</td>
<td>0.078</td>
<td>0.013</td>
<td>0.181</td>
<td>0.124</td>
</tr>
</tbody>
</table>

### Table 5 – Results of the basic probe examination task