Chapter 5

Construct and concurrent validity of a Nintendo Wii video game made for training basic laparoscopic skills

M.B. Jalink, J. Goris, E. Heineman, J.P.E.N. Pierie, H.O. ten Cate Hoedemaker


Abstract

Background: Virtual reality laparoscopic simulators have been around for more than ten years and have proven to be cost- and time-effective in laparoscopic skills training. However, most simulators are, in our experience, considered less interesting by residents and are often poorly accessible. Consequently, these devices are rarely used in actual training. In an effort to make a low cost and more attractive simulator, a custom-made Nintendo Wii game was developed. Ultimately, this game could be used to train the same basic skills as virtual reality laparoscopic simulators ought to do. Before such a video game can be implemented into a surgical training program, it has to be validated according to international standards.

Methods: The main goal of this study was to test construct and concurrent validity of the controls of a prototype of the game. In this study, the basic laparoscopic skills of experts (surgeons, urologists and gynecologists, n=15) were compared to those of complete novices (internists, n=15), using the Wii Laparoscopy (construct validity). Scores were also compared to the FLS Peg Transfer test, an already established assessment method for measuring basic laparoscopic skills (concurrent validity).

Results: Results showed that experts were 111% faster (P=0.001) on the Wii Laparoscopy task than novices. Also, scores of the FLS Peg Transfer test and the Wii Laparoscopy showed a significant, high correlation (r=0.812, P<0.001).

Conclusions: The prototype setup of the Wii Laparoscopy possesses solid construct and concurrent validity.
Introduction

Virtual reality (VR) laparoscopic simulators have been around for more than ten years\(^1\) and have proven to be cost- and time-effective in laparoscopic skills training\(^2\)\(^-\)\(^4\). Nowadays, simulators are widely accepted as tools for training these skills\(^4\) and are incorporated into several surgical training programmes. However, it is the trainers’ experience that VR simulators are not as often used as they should be. Training is advised and in some cases compulsory, but it is often seen that residents only use simulators to reach the fixed training goals, but avoid systematic practice afterwards. This way, the evidence based benefits can’t come to fruition\(^5\).

It has been suggested that a custom-made video game can be used to train basic laparoscopic skills in surgical novices as well, creating a fun, easily accessible, and low cost alternative to existing simulators\(^6\)\(^-\)\(^10\). A collaboration between two hospitals from the Netherlands and a game developer among others, started the development of a custom-made video game for Nintendo’s Wii video game console that could be used to train basic laparoscopic skills. The goal of this co-operation is to create and sell a video game that solely focuses on training these basic skills, creating a fun training method that will eventually be put on the market for an acceptable price, even for the general consumer.

The Wii (Nintendo Co., Ltd., Kyoto, Japan) is a popular, low cost video game console that is controlled by motion-sensing, remote like controllers. To mimic the movements made in laparoscopic surgery, a custom-made Wii Remote add-on was developed (figure 1). Using two of these controllers and specially written game software, novices can train basic laparoscopic skills, such as inverted movements, eye hand coordination, depth perception, and ambidexterity. The development of a traditional simulator was knowingly avoided. Although similar setups have been made with existing Wii games\(^9\), this is the first time that video game hard- and software were specifically designed for training basic laparoscopic skills.

There is general consensus that laparoscopic simulators have to be validated before they can be implemented into a surgical training program. The demonstration of construct and concurrent validity are essential parts of this process\(^11\)\(^-\)\(^13\). Therefore, the construct and concurrent validity of the current prototype was tested according to international consensus described by the Work Group for Evaluation and Implementation of Simulators and Skills Training Programmes of the European Association for Endoscopic Surgery (EAES)\(^11\).
Goals and hypotheses

The main goal of this study was to acquire construct and concurrent validity for the controls of the Wii Laparoscopy, which are in short two of the four validation tests a fully developed simulator has to endure before it can be deployed as an actual, evidence-based laparoscopic simulator. The other two validation tests are face validity, which will be tested separately in a larger group of experts, and content validity, which does not apply for a simulator that knowingly lacks anatomy and real procedures. In this study it was hypothesized that the custom-made Wii Laparoscopy video game could differentiate between novices and laparoscopic experts and therefore had good construct validity, and that performance scores from the Wii Laparoscopy would correlate with a similar assessment method used as a control, proving its concurrent validity.

Methods

Participants

In this study, the basic laparoscopic skills of experts were compared to those of complete novices, using the Wii Laparoscopy game and the Fundamentals of Laparoscopy (FLS) Peg Transfer test. The expert group consisted of surgeons, gynecologists and one urologist with extensive laparoscopic expertise (minimally 100 laparoscopic interventions in the last five years). In contrast to other studies in the field, medical interns were not used as laparoscopic novices, because they are likely to be more adapted to modern technology.
(games, blind typing, smartphone usage) than current-generation surgeons\textsuperscript{16}. Therefore, other, non-surgical doctors of the same generation were chosen to form a group with similar baseline characteristics such as gender, age, educational level, and video game experience\textsuperscript{17,18}. Hence, internists with no experience with laparoscopy or endoscopy were chosen as novices. Participants were recruited throughout the Netherlands. No participants suffered from a physical disability that prevented them from successfully finishing one of the tasks and none had a medical reason to avoid video games, such as epilepsy. Surgeons and gynecologists that had already partaken in our pilot studies were excluded.

In total 15 experts (11 surgeons, 3 gynecologists and 1 urologist) and 15 novices were included in the study. Experts had performed an average of 272 (± 123) laparoscopies in the last five years; novices had no laparoscopic experience whatsoever. Demographic data is shown in table 1, demonstrating that both groups are equal for sex, mean age, hand dominance, and game experience on every age.

<table>
<thead>
<tr>
<th>Table 1: Basic demographic data</th>
<th>Novices (n=15)</th>
<th>Experts (n=15)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male)</td>
<td>12</td>
<td>10</td>
<td>1.000 \textsuperscript{F}</td>
</tr>
<tr>
<td>Mean age in years (SD)</td>
<td>44 (±10)</td>
<td>42 (±8)</td>
<td>0.629 \textsuperscript{T}</td>
</tr>
<tr>
<td>Hand dominance (right)</td>
<td>13</td>
<td>15</td>
<td>0.483 \textsuperscript{F}</td>
</tr>
<tr>
<td>Mean GE as a child (scale 1-7)</td>
<td>1.67</td>
<td>2.47</td>
<td>0.706 \textsuperscript{M}</td>
</tr>
<tr>
<td>Mean GE as an adolescent (scale 1-7)</td>
<td>1.93</td>
<td>2.87</td>
<td>0.581 \textsuperscript{M}</td>
</tr>
<tr>
<td>Mean GE as a student (scale 1-7)</td>
<td>1.80</td>
<td>2.80</td>
<td>0.145 \textsuperscript{M}</td>
</tr>
<tr>
<td>Mean GE as a resident (scale 1-7)</td>
<td>1.73</td>
<td>2.80</td>
<td>0.103 \textsuperscript{M}</td>
</tr>
<tr>
<td>Mean GE as an attending (scale 1-7)</td>
<td>2.27</td>
<td>1.93</td>
<td>0.932 \textsuperscript{M}</td>
</tr>
</tbody>
</table>

\* = significant difference (\(\alpha = 0.05\)), GE = game experience

\textsuperscript{T} = independent samples Student’s T-test, \textsuperscript{F} = Fisher’s exact test, \textsuperscript{M} = Mann-Whitney U test (all are 2-tailed)
Apparatus and tasks

The complete Wii Laparoscopy prototype consists of a 15 inch (16:9) flat-screen TV, a Nintendo Wii, two Nintendo Wii Remote controllers in custom-made laparoscopic tool shells, software, four infra-red (IR) Light Emitting Diodes (LEDs) placed on a base plate, and two small oarlocks on poles to resemble trocar sites. Because participants were tested in different hospitals throughout the North of Holland, the prototype setup was mounted into a trolley briefcase (figure 2). The Wii Remotes use a built-in IR camera to see the IR LEDs on the bottom of the briefcase. This way, their position can be determined and linked to the tools shown on the screen. The handles trigger the joysticks mechanically, so that the graspers on screen can be opened in a natural way. A button on the back of the joystick is used to activate a driller tool.

Players use the controllers to play a game that is based on the movements made during laparoscopic surgery. In contrast to other simulators, the Wii Laparoscopy does not contain actual medical content, but comprises a fictional mine-world in which players have to drill pieces of scrap of a wall, using two large, rigid robotic arms. Participants then had to drop the ten pieces of scrap metal into a virtual melting pot. A driller was placed in the dominant hand and a set of graspers was placed in the non-dominant hand. Scrap had to be hold firmly with the graspers before it could be drilled loose. To evade testing properties as intelligence or vision, a simple and clear level design was chosen.

A sum score was computed to determine participants’ performance. Time in seconds was measured between the first grab and the final drop. The maximum time limit was set at 300 seconds and remaining scrap was counted as a penalty. Scrap that was dropped outside of the melting pot was counted as an error. Errors and penalties counted as 17 seconds and were added to the total score. This was chosen in accordance to the FLS Peg Transfer test. Wii Laparoscopy scores were compared to the sum time score of the FLS Peg Transfer test\textsuperscript{19}, a test that is extensively validated\textsuperscript{20-23} to measure eye hand coordination, inverted movements, depth perception, and ambidexterity. In this test participants had to transport six beads that are placed on a board with twelve pegs from the left to the right pegs and back with two laparoscopic graspers while transferring the beads in mid-air.

A camera was placed between the trocar poles and a plate was placed above the operating field to prevent peeking. Tests were performed using the official FLS pegboard and beads.
Assessment
Testing took place in a quiet, separate room in different hospitals. Before the start, participants received a brief introduction. All participants filled in a questionnaire for the assessment of demographic data. Video game experience on different ages was measured using 7-point Likert scales, as described by Schlickum et al. Afterwards, the FLS Peg Transfer test and the Wii Laparoscopy assignment were both performed twice. The first attempts were considered introductory rounds; the second attempts were used for analyses. Participants received no compensation for their participation.

Power analysis
Power analysis for a Pearson product-moment correlation coefficient was conducted in Power Analysis and Sample Size (PASS) for Windows, version 11 (NCSS, Kaysville, UT, United States of America) to determine a sufficient sample size using an alpha of 0.05, a power of 0.80, a large effect size \( r=0.65 \), a null hypothesis of 0.2, and two tails. The expected effect size was based on an unpublished pilot study using a less advanced prototype \( r=0.55 \) and a similar study in which the FLS Peg Transfer test was compared to the path length of a virtual reality version of the same test (novices: \( r=0.78 \), experts: \( r=0.86 \)). Based on the aforementioned assumptions, the desired sample size for concurrent validity was 27. Sample
sizes for the construct validity were based on the pilot study and were set at ≥11 participants per group.

**Evaluation**

Data were analyzed using the Statistical Package for the Social Sciences (SPSS) for Mac, version 19 (IBM, Armonk, NY, United States of America). Alpha was set at 0.05. Since a small data set was used, normality was tested with a Shapiro-Wilk test. Homogeneity of variance was checked using Levene’s test.  

**Results**

*Construct validity*

Experts scored significantly better on both the Wii Laparoscopy and the FLS Peg Transfer test, as shown in table 2. Since variance in data was not equal, a Welch’s T-test was performed to show the difference in means. This showed that experts performed 111% significantly faster on the Wii Laparoscopy than novices.

**Table 2: Construct validity data, showing mean scores and their standard deviations in seconds**

<table>
<thead>
<tr>
<th></th>
<th>Novices (n=15)</th>
<th>Experts (n=15)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLS Peg Transfer test total score</td>
<td>200 (19)</td>
<td>87 (5)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Wii Laparoscopy total score</td>
<td>209 (95)</td>
<td>99 (31)</td>
<td>0.001*</td>
</tr>
</tbody>
</table>

* = significant difference (α = 0.05)

*Concurrent validity*

Because experts were in general faster on both tests, combined total scores of novices and experts for both the FLS Peg Transfer test and the Wii Laparoscopy were positively skewed. After Log transformation both were normalized. The correlation between the FLS Peg Transfer test and the Wii Laparoscopy total scores was analyzed using Pearson’s product-moment correlation coefficient. This correlation was significant, r=0.812, P (2-tailed) <0.001. The linear regression line is shown in figure 3.
Discussion

There is general consensus that laparoscopic simulators must be validated before they can be implemented into a surgical training program. The demonstration of construct and concurrent validity are essential parts of this process \(^ {11-13}\). This study demonstrated that the tested prototype setup of the Wii Laparoscopy simulator is able to distinguish between two different levels of experience, in this case expert laparoscopic surgeons (>100 laparoscopic interventions performed in the last five years) and internists without any experience with laparoscopy and/or endoscopy. More importantly \(^ {12}\), a significant high correlation between the Wii Laparoscopy and the FLS Peg Transfer test scores was found. This indicates that our video game-driven simulator has a solid concurrent validity as well, and is thus able to mimic basic laparoscopic training exercises \(^ {11}\). This proves that a custom-made video game is able to distinguish between different levels of experience and, even more important, is highly correlated to an already established instrument \(^ {12}\).
Although the controls of the prototype have proven their construct and concurrent validity, it should be noted that an eventual final version will be different from the current construction. The hardware will be modified into a more user-friendly design and the software will be changed into a game in which the player has to solve more kinds of puzzles, using various techniques. Therefore, this study does not qualify as a validation of the final product. In the future, all parts of the final version will have to undergo a similar experiment to finally prove their face, construct and concurrent validity, before this video game can be used to actually train basic laparoscopic skills.

**Limitations of the study**

In this study only one parameter, a cumulative time and error score, was used. Although a time score is used more often \(^\text{13,15,23}\), one can not say which parameter is the best indication of laparoscopic dexterity \(^\text{27}\). Theoretically it is possible to measure other parameters, such as path length or economy of movement, with the Nintendo Wii. But because the FLS test is one of the most validated and accepted scoring methods to date \(^\text{20-23}\), we decided to compare this gold standard to our model. The Peg Transfer test was chosen in particular because of the similarities in tested basic laparoscopic skills, such eye hand coordination, inverted movements, depth perception, and ambidexterity.

Because there is currently no validated scale for measuring video game experience, it is not possible to say whether or not both groups really match. To analyze the experience with video games in different age categories and to create uniformity, the method described by Schlickum et al. was used \(^\text{6}\).

This study demonstrates the construct and concurrent validity of a custom-made video game for Nintendo Wii, indicating that it is possible to use such a construction to train basic laparoscopic skills. In the future, the Wii Laparoscopy setup will be developed further into a full, commercially available video game, containing various levels for training a wider array of basic laparoscopic skills.

**Acknowledgements**

The authors would like to thank H. Groen, Department of Epidemiology, University of Groningen, University Medical Center Groningen, for checking our statistical analyses.
References


