Validation of a video game made for training laparoscopic skills
Jalink, Maarten

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2014

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 1

Introduction and outline of this thesis

About (serious) games
When thinking of a person who plays video games, a certain stereotype comes to mind. It is a lonely, obese man, around the age of 20-30, reeking of sweat. He avoids sunlight by living in his parents’ basement, and only comes outside to wait in line in front of a video game store to pick the newest expansion of the online fantasy game World of Warcraft. Although this is how the stereotype gamer is depicted in most popular media\(^1\), it does not apply for the large majority of gamers nowadays, if it ever was applicable. In the past decades, video game systems have evolved from arcade cabinets, stuffed in dark halls, into widely accepted and accessible television consoles, used by the complete family. Consequently, the video game industry has transformed from a niche business into an ever-growing, multi-billion dollar market\(^2\).

In its latest, annual report on sales, demographic, and usage data on video games, the Entertainment Software Association showed that 58% of all Americans play video games, and that half of all United States households owns a dedicated video game console\(^2\). The average age of gamers is 30 years, and about 45% of all gamers are female, a number that has risen from 38% in 2006\(^3\). In the last years, portable devices, such as smartphones and tablets, are becoming more popular gaming platforms.

As said, the video game industry is a big one. In 2012, an estimated $20.77 billion was spent by consumers on video games, gaming systems, and accessories in the United States alone\(^2\). In comparison, the money consumers spent on movies throughout all of North America in 2012 was $25.58 billion\(^4\). Recently, a new video game console, the PlayStation 4 (Sony Corporation, Tokyo, Japan), was sold 1 million times in only 24 hours after its release\(^5\). Similarly, the popular video game Grand Theft Auto V (Rockstar Games, New York, NY, United States), which had cost $200 million to develop, had a turn-over of a whopping $1 billion in its first 72 hours\(^6\). These facts and numbers show that video games have slowly become part of Western culture, making its appearance in the entertainment industry as common as books, radio, and movies.
Just like those media, video games can be used for educational purposes as well. These games are categorized under the oxymoron “serious games”. The category itself has various, sometimes conflicting definitions, such as “games that do not have entertainment, enjoyment or fun as their primary purpose” \(^7\). And, they should not be confused with simulators, although there is a large, grey area. Though the first serious game was developed more than 50 years ago \(^8\), its market only started expanding about a decade ago \(^7\). In 2010, the serious games market was estimated to be worth $1.5 billion worldwide \(^9\).

The first (digital) serious games were developed during the Cold War, mainly for military purposes \(^7\). A good example is the 1955 strategy game HUTSPIEL, developed by the Operations Research Office, a civilian military research center run by the John Hopkins University in Baltimore \(^8\). The game was played by two persons, who were either the commander of a red (USSR) or blue (NATO) team, that fought a war along the Rhine river. The goal of the game was to learn and research the possible outcomes in case the incident actually happened.

Of course, HUTSPIEL and similar games were not available for the general public. Gaming arrived in the living room in 1972, when the German-American engineer Ralph Baer released the first commercially available video game console; the Magnavox Odyssey (Magnavox, Napa, CA, United States). The console could only display simple, monochrome graphics, lacked the ability to produce sound, and only had 27 different games in total. Although most games were purely aimed at entertainment, some games were meant for education as well. In the game “States”, one would learn the topography of the United States. A light-gun peripheral and game were aimed at firearms training \(^7\).

In subsequent years, different video game consoles and personal computer systems surfaced, all spawning various kinds of educational games. And in most cases, the validity or effects were not evaluated. One of the first serious games aimed at healthcare was Captain Novolin (Raya Systems, Mountain View, CA, United States), a 1992 Super Nintendo game in which the player controlled a diabetic superhero \(^7\). The goal of this game was to teach diabetes self-management, and demonstrated the relationship between food, insulin, and blood glucose levels. An evaluation amongst 23 patients showed that the game was entertaining, educating, and helped patients to explain their disease to their friends \(^10\).
Although digital games made for training and education have been around for decades, the term “serious games” and its complete field of interest, both commercially and scientifically, started flourishing around 2002. This turning point is said to be marked by the release of America’s Army, a realistic first person shooter made by the United States Army, which was successfully used as a recruitment tool. Another example is Re-Mission (HopeLab, Redwood City, CA, United States), a 3D game aimed at children with cancer. It is remarkable due to a big, randomized controlled trial, that showed that playing the game raised self-efficacy and knowledge, and that patients had greater medication adherence, which can be a problem in pediatric patients. The best examples of commercially successful serious games are Brain Training and Wii Fit (Nintendo, Kyoto, Japan), games respectively aimed at improving concentration and physical condition. Nintendo, however, did not research the effects of their games, but external scientific papers on the positive effects of Wii Fit are surging (see chapter 7).

Nowadays, serious games are serious business. In the Netherlands alone, there are more than ten independent game studios that specialize in the development of serious games, such as Grendel Games, MAD Multimedia, Ijsfontein, and Little Chicken. The Delft University of Technology offers a master’s degree in the development of serious games, and the Noordelijke Hogeschool Leeuwarden (NHL) has a specific serious games lectorate. Other academic institutions have also incorporated subjects on the development of serious games in their curricula. Consequently, research on serious games is evolving, shown by the emergence of dedicated, peer-reviewed journals for video games and healthcare. It is hard to predict the future of serious games, but based on the progression of the last ten years, the concept of using video games for serious purposes will undoubtedly become more common.

**Laparoscopy, training, and simulators**

Laparoscopy, derived from the ancient Greek words λαπάρα (lapara), meaning “flank”, and σκοπέω (skopeó), meaning “to see”, is a form of minimal invasive surgery that arose in the beginning of the 20th century, and started to gain popularity in the ‘70s. In a laparoscopic intervention, the abdominal (or peritoneal) cavity is inflated with carbon dioxide gas. A long scope (nowadays with an added video camera) and special instruments, such as graspers, scissors, dissectors, or electrosurgical probes, are then inserted into the abdominal cavity through so-called trocars, special ports that penetrate the abdominal wall via small incisions. This way, surgeons, urologists, and gynaecologists can perform a wide, and growing array of
operations, while leaving minimal visible marks on the body of the patient. Best-known laparoscopic operations are the cholecystectomy and appendectomy, respectively the removal of the gall bladder and the appendix. Although the promoted benefits, such as shorter hospital stay, of laparoscopic surgery are not always evident in comparison to conventional (open) surgery\textsuperscript{17}, this form of minimal-invasive surgery has become a standard for many operations.

In the last twenty years, laparoscopy has become a standard part of the surgical traineeship. Therefore, surgeons in training are required to learn and refine cognitive and psychomotor skills\textsuperscript{18}, such as inverted movements, eye hand coordination, depth perception, ambidexterity, and the fulcrum effect. At most universities in the Netherlands, the basics of laparoscopy are taught through special courses, in which box trainers and animal models are used. Besides these training programs, trainees have the possibility to practice their skills on their own. Traditional box trainers and animal models, however, require human monitoring, and are thus not always available. This makes traditional methods subjective, expensive, and time consuming\textsuperscript{19}.

Therefore, virtual reality simulators are becoming more important, especially in the last decade\textsuperscript{20}. Virtual reality simulators consist of a computer and a specially designed input device in the shape of laparoscopic instruments, which can be used to maneuver on-screen, instruments in a realistic, computer simulated, three-dimensional abdomen. By performing several, pre-programmed operations, one cannot only train basic laparoscopic skills, but also learn properties such as anatomy, and procedural skills and knowledge. Some examples of simulators are the LAP Mentor II (Simbionix Ltd., Cleveland, OH, United States) (figure 1), and the SIMENDO (DeltaTech, Delft, The Netherlands). Most virtual reality simulators have been validated according to an international consensus\textsuperscript{21,22}, have proven to be cost- and time-effective in laparoscopic skills training\textsuperscript{23-26}, and their taught skill is shown to be transferable to the operative setting\textsuperscript{27}. 

At the University Medical Center Groningen (UMCG, Groningen, The Netherlands), however, it is the trainer’s experience that surgical trainees do not use the simulators as they’re supposed to. The Wenckebach Skills Center, the clinical training centre of the UMCG, has a large collection of simulators, and among them are several laparoscopic simulators. Although training on these advanced simulators is advised, and in some cases obligatory, it is often seen that residents only use them to reach fixed training goals, and avoid systematic practice afterwards. After the basic laparoscopy course, most trainees only practice a few times or completely skip the virtual reality simulators, and start their learning curve on actual patients. This way, the proved benefits of the simulators cannot come to fruition. Most residents consider the procedures “repetitive” and “boring” on the long run, and complain about the distance between their workplace and the simulators. Since most virtual reality simulators are expensive, they are safely locked away in secured rooms. Another problem is the fact that some simulators, such is the LAP Mentor II, have a fragile control mechanism, which requires a lot of maintenance. Hence, the threshold for frequent simulator practice is high. To increase voluntary basic laparoscopic skills training in surgical trainees, a serious game was developed.
Underground: a serious game made for training basic laparoscopic skills

The project that is evaluated in this thesis is a serious game, developed for Nintendo’s Wii U video game console (Nintendo Co., Ltd., Kyoto, Japan), that is made for training and maintaining basic laparoscopic skills. Its goal is not to replace existing simulators, but rather to be a cheap, and fun addition to the curriculum, which can be utilized voluntarily. The game, called Underground, will be commercially available, but is primarily aimed at surgical trainees. In this paragraph, the history, development, technical details, and validation of the game and hardware will be presented.

History and development

The initial idea for the development of a game for training laparoscopic skills was conceived by abdominal surgeon Henk ten Cate Hoedemaker and educational consultant Jetse Goris in 2008. Both noticed that the laparoscopic simulators at the Wenckebach Skills Center were hardly ever used, while the majority of surgical residents did have the time to play games on their smartphones while staying in the hospital. This seemed paradoxical, since simulators and video games have a lot in common. Why would residents, who are interested in surgery, rather play a casual game like Angry Birds (Rovio Entertainment Ltd., Espoo, Finland) on their phones, than practice their skills by playing on a far more advanced and dedicated simulator? The answer to this question seems to lie in the inferior entertainment value of simulators. Simulators are purely aimed at skills training, and this can only be achieved by strenuous repetition of predictable exercises. Video games, on the other hand, are designed to entertain people. And, if a game is challenging, unpredictable, and immersive, people will continue to play until they are experts in the thing they are doing. However, making people proficient at a certain game (and gaming mechanism) is not the primary goal. For example, Nintendo’s famous game developer Shigeru Miyamoto did not design the Super Mario games to make consumers good at pressing buttons to make a man jump on fictional enemies. But eventually, upon completion of one of the games, players are an expert at it. In a well developed serious game, this interesting side-effect can be used to teach players a certain skill, while they may not be aware of the fact they’re actually practicing it.

To explore the possibilities of a serious game for training basic laparoscopic skills, Goris and Ten Cate Hoedemaker contacted game development studio Grendel Games (Leeuwarden, The Netherlands). After gaining their interest, a prototype of a Nintendo Wii game was developed
and presented at the Game Developers Conference 2010 (San Francisco, CA, United States). This version consisted of a mechanical and rather cumbersome setup, which used two Wii Remote controllers with Wii Motion Plus add-ons, which adds motion sensing through gyroscopes, attached to it (figure 2, upper left photo). Using these gyroscopes, the player could maneuver two on-screen laparoscopic graspers in a three-dimensional, fictional robot world. It was found out that this contraption was not precise enough to simulate the movements made during laparoscopy, mainly because the Motion Plus gyroscopes require frequent calibration. Also, the hardware was too large and expensive to make a useful and affordable product. These ideas were scrapped and the team went back to the drawing board.

Figure 2: various prototypes of the Underground hardware, including the first version (top left) and a 3D printed, near-final version (bottom right)

Through an iterative process, which involved various other parties, such as master student biomedical engineering Ivar Bosma 29 and product innovation company Pezy (Groningen, The Netherlands), a new concept was developed. The development process was based on Scrum, a product development framework mainly used in software development 30. It is characterized by splitting larger projects into smaller, more manageable pieces that are
evaluated every other few weeks or months. The theory behind this framework is that it promotes creativity and makes quick changes to the project possible through frequent feedback. This way, the course of development can be changed along the way, making room for changes in both hard- and software of the game.

In this stage of the development process of Underground, several “sprint meetings” were held every other week. During every session, new work was presented and evaluated at the Grendel Games office in Leeuwarden. There, the next steps in the process were discussed and the goals for the next session were arranged. The following two weeks each part of the team (e.g. research, hardware or software development, or financial management) then worked on their part of the project, which was evaluated in the upcoming sprint meeting. Grendel Games, for example, developed the video game itself, starting with the underlying code for the control mechanism, slowly working towards basic, but playable levels and finally a full video game, complete with a storyboard, computer-generated animations, and orchestrated music. The first version of the hardware was built using simple materials from do-it-yourself hardware stores (figure 2, upper right photo). Later iterations, which were rapidly made for the sprint meetings, were made using rapid prototyping by means of a 3D printer. After evaluation by the team, the various handles were then evaluated by abdominal surgeons at the UMCG and a definitive design was chosen. Eventually, a precise motion sensing system, which utilizes the infrared camera of the Wii Remote, was developed. After careful consideration of the production costs, molds for injection molding were created in China by a third party (Fluctus Industrial Design & Engineering, Groningen, The Netherlands). This way, the hardware design was finalized and prepared for mass production. Technical details are discussed later in this chapter.

To be able to realize this project, a dedicated company was erected in 2010; Cutting Edge B.V. (Groningen, The Netherlands). This company consists of three parties: the UMCG (through Stichting Triade), Grendel Games, and the Leeuwarden Institute of Minimally Invasive Surgery (LIMIS, Leeuwarden, The Netherlands). Cutting Edge is a commercial enterprise which has received several grants to be able to develop Underground. Jetse Goris and Henk ten Cate Hoedemaker, who both contributed to the research performed for this thesis, are intellectual property holders of the concept. Therefore it should be disclosed that they do have a financial interest in this project. A full disclaimer of the possible conflicts of interest of the author and all co-authors can be found in the appendix. To further maintain
scientific integrity, an independent academic professor was recruited as doctoral advisor for this thesis. Also, where possible, data were gathered anonymously and blindly analyzed.

Although Underground has not yet been released at the time of writing, it has received both national and international media attention\(^{31}\) and has won the Dutch Game Awards for Best Serious Game and Best Applied Game Design in 2013\(^ {32}\).

![Figure 3: a comic, based on Underground, that appeared in various Dutch newspapers\(^ {33}\)](image)

**Technical details**

The hardware of the game consists of various components. To start, the player needs to connect a Wii U video game console to his TV, and download the Underground game via the online Nintendo eShop\(^ {34}\). Next, two Wii Remotes and their Nunchuks need to be connected to the Wii U (figure 4). A Wii Remote is a wireless, remote like controller that is able to detect movement using an infra-red camera on top amongst others. The Nunchuk is a Wii Remote add-on with an analog joystick that connects via a cable to the bottom of the Wii Remote. Normally, a small “Sensor Bar” with four infra-red LEDs is placed underneath the TV\(^ {35}\). When pointing the Wii Remote at the TV, the infra-red camera sees the LEDs and calculates its position in comparison to the TV/Sensor Bar through triangulation. These data can be translated into on-screen actions, such as the display of a precise cursor or, for example, the movement of a gun. Using this mechanism, controllers do not require calibration.
To play Underground with its specially designed, laparoscopic style controls, both controllers need to be inserted into two, identical plastic shells (figure 5). These shells are developed to resemble laparoscopic instruments. The Wii Remote is placed on top, with the infra-red camera facing down when playing. The Nunchuk is placed inside the handle. When the handle is opened, the joystick on the Nunchuk moves downwards. This information is then sent to the game console, resulting in on-screen movement of instruments. The tool shells have a long rod attached to their handles, which are inserted into two small oarlocks on a base plate (figure 5). When holding the shells, the Wii Remotes point towards a designated area on the base plate, where they can “see” four tactically placed infra-red LEDs, similar to the Sensor Bar. This way, the controllers determine their position in the virtual world. Movements are translated into on-screen movement of two large, robotic arms, resembling laparoscopic instruments.
The game itself is nothing like a traditional, virtual reality laparoscopic simulator. Instead of working in an abdominal cavity, the player operates in a fictional mine, where it has to guide small robots to the exit. This is done by manipulating the environment through the means of two big robotic arms, which are controlled by moving the controllers. The concept of a mine was chosen because laparoscopic surgeons also work in a primarily dark area and have to break things (adhesiolysis, ligation of mesentery, resections), before they can start to rebuild (anastomoses, hernia repairs). The gameplay of Underground consists of a mixture of action and puzzles in a sandbox like environment. The main goal, escaping from the underground world, can be reached by solving puzzles that require basic laparoscopic skills, cognitive skills, and problem-solving skills. The levels of the game are set up in such a way that different laparoscopic skills are covered. The game has four distinct themes, each with five different levels and a final boss. To defeat the boss at the end of each world, the player has to combine all the skills that were learned during the previous levels. This way, the player practices its basic laparoscopic skills, while playing a video game. Originally, the game was developed for the Wii console, which was introduced in 2006. Because of the emergence of the Wii U in 2012, the game was ported to this more advanced successor of the Wii. The control mechanism remained unchanged, but due to the fact that the Wii U is graphically superior to the Wii, developer Grendel Games was able to make more complex levels and more detailed environments.
Validation

The main goal of this thesis is to validate the Underground game. Since the game is developed to more or less resemble a laparoscopic simulator, it was decided to validate it in the same manner as normal simulators are validated. It should be noted that “validation” is a very nonspecific term that consists of various concepts. For example, “experimental validity”, covers the actual effect of an intervention on its users. This is covered in chapter 6.

But when talking about validation in the field of laparoscopic simulators, one often means “test validity”. The Work Group for Evaluation and Implementation of Simulators and Skills Training Programmes of the European Association for Endoscopic Surgery (EAES) defines this type of validity as “the extent to which an assessment instrument measures what it was designed to measure”. In the EAES’s 2005 consensus guidelines for validation of virtual reality surgical simulators, four kinds of validity are described: face, content, construct, and concurrent validity.

- Face validity is based upon the judgment of a defined group of subjects of the resemblance between the system under study and the real activity. This can be tested by interviewing surgeons about the hardware, such as the looks and feels of the laparoscopic tools in real life and its movement in virtual space. Although this is not an objective measure, inter-
subjectivity (multiple judgments combined) can be reached by interviewing a large group in a systematic manner. This way, one hopes to obtain a general expert opinion. Since face validity is a very subjective measure, it is the least important part of the validation study. It can be considered as an expert opinion (level 5 evidence).

- Content validity describes the level of which the simulator covers the subject matter of the real activity. In contrast to face validity, which discusses the looks and feels of the hardware of the simulator, content validity is concerned with the software, for example anatomy and correctness of the performed procedures. This can be obtained by a similar process of interviewing and is frequently combined with consensus in peer-reviewed literature and frequent meetings with experts (level 5 evidence).

- Construct (or contrast) validity refers to the ability of a simulator to distinguish between different levels of experience with the studied procedure to see whether it measures what it is supposed to measure. By letting novices and experts play on the simulator, a significant difference in score should be seen. If experts, who are supposed to be the best, do not get a high score, the apparatus does not simulate reality in a correct way (it does not measure what it is supposed to) and is therefore not a good simulator (level 3 evidence).

- Finally, concurrent (or predictive) validity is used to compare the outcome of the simulator to an established assessment method, designed to measure the same skills. Another, already validated simulator or skills test will differentiate between experts and novices, and the scores of the new simulator should correlate with the outcome of the chosen standard. Concurrent validity can be considered the most important part of test validation, because it objectifies one’s expertise and correlates it to the new test. If, for example, an experienced gamer gets a high score on Underground because it is a game, but does not show the same skills on a standard assessment method for basic laparoscopic skills, the new apparatus does not simulate the same aspects and is therefore, again, not a good simulator (level 2 evidence).

Face validity is covered in chapter 4. The construct and concurrent validity of a prototype of Underground is covered in chapter 5. Since the content of the game differs from real laparoscopic surgery on purpose – the player helps robots in a mine, instead of performing a virtual laparoscopy – content validity was deliberately let out of the picture. The game is not aimed at training anatomy or procedural tasks, and thus validating for such criteria makes no sense.
Aim of this thesis

The primary aim of this thesis is to create a thorough evaluation of Underground, by studying its idea, validity, and effect. A serious game can only be called “serious” if the project is evaluated with a scientific approach, proving its serious goal. Without it, it would be just “a normal game which could be serious”. The development of serious games is skyrocketing, and its research is growing 14, but due to the novelty of this field of development, validation of many projects is still lacking 37. Therefore, I sincerely hope that this thesis will be used as an inspiration and a stepping stone for serious game developers to back the claimed goals of their projects with a scientific validation.

The main hypothesis is that video games, with Underground in particular, can be used to improve basic laparoscopic skills. Sub hypotheses are that the Underground hardware possesses valid face, construct, and concurrent validity, and is thus able to be used to improve basic laparoscopic skills. The null hypotheses are that video games cannot be used to train laparoscopic skills, that Underground’s hardware does not mimic laparoscopy, and that it is not able to improve aforementioned skills in its players. The research presented tests the above formulated (sub) hypotheses.
Outline of this thesis
The next chapter discusses the effects of video games on laparoscopic skills in general. It is followed by four chapters that discuss the science, development, validation, and effects of Underground. The seventh chapter is a more clinical view on some negative side-effects of (Nintendo) games.

Chapter 2 is a literature review on the effects of video games on one’s basic laparoscopic skills.

Chapter 3 is a general description of the concept and development of the game.

Chapter 4 discusses the results of a face validity study, performed at the Chirurgendagen 2013 and the SAGES annual meeting 2014, using a near-final version of the game.

Chapter 5 describes the establishment of construct and concurrent validity for a prototype of the game.

Chapter 6 describes the short-term effect (warm-up) of the game on the basic laparoscopic skills of laparoscopic experts.

Chapter 7 is a literature review in which all known case reports and series of Nintendo-related injuries and other complaints are discussed.

Chapter 8 presents the discussion and conclusions, including the limitations, of this thesis.
References


38. Schijven MP, Jakimowicz JJ. Validation of virtual reality simulators: Key to the successful integration of a novel teaching technology into minimal access surgery. Minim Invasive Ther Allied Technol. 2005;14(4):244-6.