Serious gaming as a training tool for medical students to improve patient safety

How effective is it ?

Universiteit Utrecht

Author: Olivier Richters
Master Thesis
August 2014, Utrecht

Supervisors: dr. R.S. Batenburg, dr. M.R. Spruit and Drs. Mary Dankbaar

Master Program: Business Informatics

University of Utrecht, Department of Information and Computing Sciences, of the Faculty of Science of Utrecht University, in the Netherlands

[Faculty of Science]
Information and
Computing Sciences

Abstract

Patient safety education for medicine students is needed to decrease preventable adverse events in Health Care. In 2011, the University Medical Center Utrecht (The Netherlands), developed the Serious Game ''Air Medic Sky-11'' (AMS-1). The Erasmus Medical Centrum Rotterdam is interested in whether or not they can use this Serious Game as a training tool for their fifth year medical students to improve their patient safety skills. To answer this question the game AMS-1 and an E-module that was created for this purpose were studied for their effectiveness in the following items: (1) knowledge, (2) self-reported skill and (3) self-reported behavior in stress and awareness on patient safety related issues. Analysis on this subject show an significant increase in knowledge, but not in self-reported skill or in self-reported behavior. Therefore it is very questionable that AMS-1 should be used as a training tool for their fifth year medical students, also since a simpler and cheaper E-module had the same effect on these items. This study does not deny that AMS-1 does not work effectively for its intended purpose, but rather that this might not be the best target group for this game. A summary of the underlying findings and reasons can be found in the discussion/conclusions section.

Acknowledgements

I started the master in Business Informatics program, at Utrecht University, in September 2012. Now, after so many months of studying I can present the results of my work.
When I was still a child, I used to play a lot of videogames and I still do in my spare time. I was always interested in the medical industry, but unfortunately I was unlucky at the numerus fixus. Now, I can say that both my interests, games and medicine, come together.

During the period of writing my thesis, many people helped me in various ways and I feel obliged to thank them. I would like to thank Supervisors: dr. R.S. Batenburg and Drs. Mary Dankbaar, for offering me the opportunity to research this topic and for advising me throughout the process of writing my thesis. My family and friends supported me in all possible ways, and I would like to thank them for being there for me.
This thesis is the outcome of almost nine months of hard work, of constructive and creative thoughts, efforts to wake up in the morning, to organize myself and to set my own deadlines. I believe that my efforts flourished, I can present a satisfactory result and I hope that the reader is going to find it as interesting as I did.

Utrecht, March 2013
Olivier Richters

Content

1 Introduction

1.1 Problem Statement

Patient Safety is concerned with the prevention of errors and adverse events/effects in hospitals; (a negative outcome resulting from a medical intervention that is not due to the underlying condition of the patient [13]) to patients associated with health care [3]. Patient safety is ensured in: ''a situation in which the patient will not suffer, or has only a slight risk of suffering any damage caused by health care professionals who are not acting in accordance with the professional standards or by failures in the care system" [83]. In the year 1999, the American Institute of Medicine (IoM), who advise the nation on how to improve health, released the report "*To Err is Human*" [1]. This influential report, over 8000 citations in twelve years, changed the view on health care quality by focusing on patient safety. The report was based upon the analysis of multiples studies by a variety of organizations. It concluded that between 44.000 to 98.000 people in America die each year as a result of preventable medical errors. For comparison, this is more than deaths by motor vehicle, AIDS or breast cancer. In 2001 another report named ''*Crossing the quality chasm*'' [2] followed "*To Err is Human*" and urgently called for change to the health care system processes to improve the quality of health to reduce preventable medical errors. Since that moment, several other studies documented the extent of medical errors and underlined the need for patient safety education [15-17]. With the considerable numbers that these documents showed, they were instrumental in raising the topic Patient Safety to a major concern in health care and among policymakers.

In the Netherlands, the topic Patient Safety has and still is getting a lot of attention, both in practice and scientifically. The numbers from a report in 2004 are staggering [4]. The report shows that each year 1735 (mean) preventable deaths occur, from the total hospital admissions 5,7% experiences an adverse event and out of the 3% death rate under hospital admissions 10,3% died from adverse events which between 3,80% (1482) and 5,23% (2032) were preventable. The study also indicated that the hospital stays are 9,1 days (mean) longer as a result of adverse events. These are numbers that get refreshed each year by the Dutch media and in other literature.

The authors of *"To Err is Human''* proposed to develop patient safety curricula in medical schools and recently the World Health Organisation published The Patient Safety Curriculum Guide [14], a valuable aid in achieving this goal (see appendix A). This guide helps translating patient safety topics into curricular models. It identifies topics that a medical curriculum should contain and provides an overview of patient safety competencies to aim at. A systematic literature review [5] shows that patient safety education into the medical school curriculum is most commonly implemented in medical schools in developed countries such as the United States of America and the United Kingdom. Most of these courses are optional courses or are integrated into clinical internship or skill courses, and have not been formally included in the undergraduate medical education system. Also, this literature review shows that there are great difference in course design and contents, who is taught, the teaching resources, faculty development and outcome evaluation. Other studies about patient safety education for medical students include investigations of the effectiveness of classical lectures and other types of education methods like case-based discussions, seminars and small group discussions [6–12]. All reports claim to be effective, but there is little consensus on the best teaching approach and the strength of the underlying evidence is limited. It is questionable whether the WHO recommendations can be easily achieved with traditional methods of teaching. An approach that contains active participation of students is likely to be the most

successful [5].

1.2 Air Medic Sky-1

One novel approach that seems promising is serious gaming. Serious Educational Games (SEGs), games with educational purposes beyond entertainment, have been shown to engage the learner through exploration, experimentation and stimulate learning through increased visualization, creativity and interactive engagement. Educational games and simulation programs fit into the new challenge-based, active forms of learning in education. In the last decade, research and interest in the potential of gaming on learning has exploded, as has the diversity of games themselves [21].

At University Medical Center Utrecht's Patient Safety Center, an international team [20] was established to develop a SEG for teaching and training in patient safety issues. This resulted in *Air Medic Sky One* (AMS-1), a SEG in patient safety, specifically designed for final year medical students and junior residents, completed in 2011. Young doctors must learn to deal with the stress of being responsible for many sick patients simultaneously, while effectively communicating with superiors and nurses, sometimes in matters of life and death and under considerable time pressure. Taken together, these issues can be a serious patient safety risk. AMS-1 can't change these contexts with a game. However, it could change the way doctors manage within these contexts. AMS-1 fills a void in current medical training by teaching doctors about non-medical factors which can influence their performance. AMS-1 improves doctors' ability to recognize problems before they occur, and so give them more time to adjust and prevent harm to their patients. AMS-1 offers a safe place to practice decision making under extreme pressure, so doctors can improve their style and become comfortable in their ability to cope. AMS-1 uses biofeedback to let doctors experience their own stress regulation and become capable of consciously influencing their own stress levels, independent of external stressors.

AMS-1 is not about teaching medicine. There are plenty of tools that apply to that need. AMS-1 is about helping doctors use medical knowledge in the real world. Medical curricula teach doctors about diseases and illnesses, but insufficiently equip them to cope with the demanding and rapidly evolving situations in which they have to apply that knowledge. This is something doctors typically have to figure out for themselves, on the job, sometimes at the cost of patient harm. AMS-1 could help; it provides safe opportunity for doctors to enhance their performance in patient safety topics so they can remain stable in an unstable environment so they, as individuals and as part of their team, can always deliver to the best of their abilities, independent of the external strains that they are subject to. Thus, AMS-1 is about improving healthcare outcomes without improving medical knowledge.

AMS-1 takes around 3-4 hours to finish. The game consist out of three parts and starts with the Biofeedback part. The game comes with a device to learn to control physiology via simple breathing exercises. This device is called the biofeedback device and uses three finger sensors to measure heart rate variability and skin conductance. The exercises learn the player how to use breathing exercises to focus, deal with stress and recognize and deal with the signs of sleep deprivation. For each breathing exercise the player gets points which are shown at the bottom of the user interface.

The second part is the called the Biodome. In the Biodome players can follow lectures about patient safety on different topics, taught by experts in specific areas. These topics are patient safety, communication, focus under stress, teamwork, sleep deprecation and depression.

Completing a lecture will earn the player different points than for the breathing exercises which is also shown at the bottom of the user interface.

When the player has gathered enough points from the Biofeedback exercises and Biodome lectures, he will be asked to help on different missions around the world. These missions are the third and last part of the game. For instance, the first mission is an earthquake in Italy with many causalities where patients can be diagnosed and treated. There are a total of 25 missions which translate into 25 different patients. These patients have been cases in real life and are therefore realistic. The player will need to apply what he has learned from the Biofeedback exercises and Biodome lectures to complete these missions. The missions are divided in several levels: beginner, resident, chief officer, etc. and thus the longer your play, the harder the missions get. A Manual has been created by Olivier Richters for the participants in this research who will play AMS-1. This manual can be found in Appendix E.

In short, AMS-1 offers lessons about patient safety and self-management in a game format. The three main learning goals of AMS-1 are: knowledge, self-management

1) Patient safety knowledge and skills: Learn about the different topics of patient safety trough video lectures: patient safety, communication, focus under stress, teamwork, sleep deprivation, depression, pre-briefing, handovers, challenges to patient safety, effective teamwork and communication.
2) Stress: Learn how to perform effective self-management (know how to use breathing exercises to focus or to deal with stress)
3) Awareness and: skill 'read-back' and 'hear-back', recognize 'red flags' in clinical situations, 'speak up' when recognizing one or more red flags.

1.3 e-module

For the benefit of the research study, the Erasmus University Medical Center has developed an e-learning module on patient safety. It was created in 2013 by Mary Dankbaar who is an e-learning program manager at the Erasmus University Rotterdam and Olivier Richters who is a last year master student of the study Business Informatics at the Utrecht University. This e-module has been validated by Dr. Kalkman at the UMC. The e-module is based on the game AMS-1. Transcripts have been made of all lectures and extra information in AMS-1 on the following topics:

a) Patient Safety: Introduction to Patient Safety, The 4 challenges of Patient Safety, 9 Red Flags
b) Communication: Pre-briefing, Assertive Statement, Call Back & Read Back, Handovers
c) Focus under Stress: Focus on your Patient, Focus on the Now through centered breathing
d) Teamwork: Shared mental model, Teamwork: Ken Catchpole –teamwork to control stress level Using Resources
e) Sleep Deprivation: Signs of Sleep Deprivation
f) Depression: Signs of Depression, Depression in Teammates, Depression & Drug Abuse, Making Mistakes when depressed
g) Self-management: breathing exercises.

The Erasmus MC uses the program called CEL to create e-modules. The transcripts and extra information from AMS-1 have been used in CEL. The e-module takes around 1-2 hours to complete. The end results can be found here:
http://erasmusmcelearning.nl/cel/clients/erasmusmc/publish/erasmusmc-44167/44167/

1.4 Research design, research methods and research questions

In the Netherlands, medicine students have to do their first clerkship of ten weeks in their fifth year. This is also called the clinical phase, because the medicine students train their clinical skills. At the Erasmus University Rotterdam it is standard procedure that all students have an introduction week about Patient Safety before their ten weeks of clerkship starts. This is called the Regular Patient Safety Education. Educational games fit in to the new challenge-based, active forms of learning. And because serious gaming may be at least as effective for patient safety training as any traditional mode of transferring medical knowledge [18,19]. The Erasmus University wants to know if they can use AMS-1 as a training tool for their medical students to improve patient safety skills in this introduction week or if they should stick to a cheaper way of educating on patient safety with the use of the e-module that was created by themselves. This results into the following main question of this research:

> **Main question***: How effective is serious gaming, used as a training tool, at improving medical student patient safety skills and what are the probable causes for this result?*

In order to answer the main question, the research is divided into three main parts with different sub-questions. The first part is a literature review, the second part is an effectiveness study and the third part is an evaluation study. In Figure 1 on page 12, the main parts of the research project are visualized in a Process-Deliverable Diagram (PDD). This PDD, developed by Brinkkemper University Utrecht [0], the activities and deliverables of the research project are shown, with varying levels of detail and showing dependencies. On the left side of a PDD the activities are shown and the objects on the right side of the diagram represent main deliverables. The arrows between activities and deliverables suggest that the activity results in a deliverable. For each of these three parts we will now discuss the used research method (briefly, and more elaborate at the chapter itself), and give the sub-questions that will that will be studied to ultimately lead to the answer to the main question.

Part 1: Literature Study
The database Pubmed will be used as data sources for the literature study. For each subject different search terms are used (table 1). Due to the scope of this research this will not be a systematic literature. Paper selection is based on search keywords, the title of the paper and the contents of the abstract. The goal is of the literature study is to get a better look at the current situation and knowledge of two subjects that are related to this research. The first subject is Patient safety which is followed by the second subject Serious Gaming. These two subjects are divided in sub-questions to get a better view of the current situation and knowledge. In this way, our results can be placed in a better position and can be compared to other researchers/literature.

<u>Patient safety</u>

> **Sub-question 1.1**: *Patient safety, why is this such an important subject?*
> **Sub-question 1.2**: *Patient safety in the Netherlands, what is the current situation?*
> **Sub-question 1.3**: *Patient education, what is the current situation in patient safety education into the medical school curriculum in different countries.*

Serious gaming

Sub-question 1.4: *Is there empirical evidence about the positive impacts and outcomes of computer games ad serious games with respect to learning in the medical domain?*

Part 2: Effectiveness

To measure the effectiveness of AMS-1, a quasi-experiment will be conducted. This is done with a control group and two intervention groups. The control group will follow the Regular Patient Safety education. The two intervention groups will follow the Regular Patient Safety education with either the e-module (first intervention group) or the game AMS-1 (second intervention group) on Patient safety. The main goal of the E-module and AMS-1 is to enhance the patient safety skills and knowledge of fifth year medical students. Thus, the effectiveness of reaching that goal has to be measured and be compared between the control, E-module and AMS-1 group in order to decide which one can and should be used as best choice. To measure the effectiveness of the e-learning module and AMS-1, level two of Kirkpatrick's framework will be used [49]. Kirkpatrick's framework is perhaps the best known evaluation methodology for judging learning processes. Level two of the framework states that if you want to measure the effect of an educational intervention, then the difference in knowledge, behavior and skill needs to be measured. It fits perfectly with the concept of AMS-1 as can be seen on page 8: the three learning goals of AMS-1. This results into the following sub-questions:

Sub-question 2.1: *Do students show enhanced knowledge in patient safety related issues during clinical training when they have played the AMS-1 game, compared to students who have done the e-module?*

Sub-question 2.2: *Do students show enhanced self-rated skill in patient safety related issues during clinical training when they have played the AMS-1 game, compared to students who have done the e-module?*

Sub-question 2.3: *Do students show enhanced self-reported behavior in Stress and Awareness in patient safety related issues during clinical training when they have played the AMS-1 game, compared to students who have done the e-module?*

.

Part 3: evaluation

Even if the e-learning module and the serious game AMS-1 both point out to have a positive effect in changing the knowledge, self-rated skill and self-rated behavior of young medical student, we do not know the underlying reasons. By making use of an evaluation questionnaire for both the E-module group and AMS-1 and interviews with students who played AMS-1, the results of the effectiveness study will be explained. For the questionnaire new instrument was created: an extension on the Technology Acceptance Model (TAM) [43]. The TAM Model is extended with the USE questionnaire [45] (which has and overlap with TAM), and the MSLQ (on motivation for learning) questionnaire. To see what aspects of the game AMS-1 students appreciate and dislike, semi-structured interviews have been conducted. The interview consisted out of three parts and each part had a few predetermined questions. Each questions unfolds in a conversational manner offering participants the chance to explore issues they feel are important, instead of answering yes or no.

Sub-question 3.1: *Do they evaluate the game better than the E-module and in what respect?*

Sub-question 3.2: *What aspects of the game do students appreciate and dislike?*

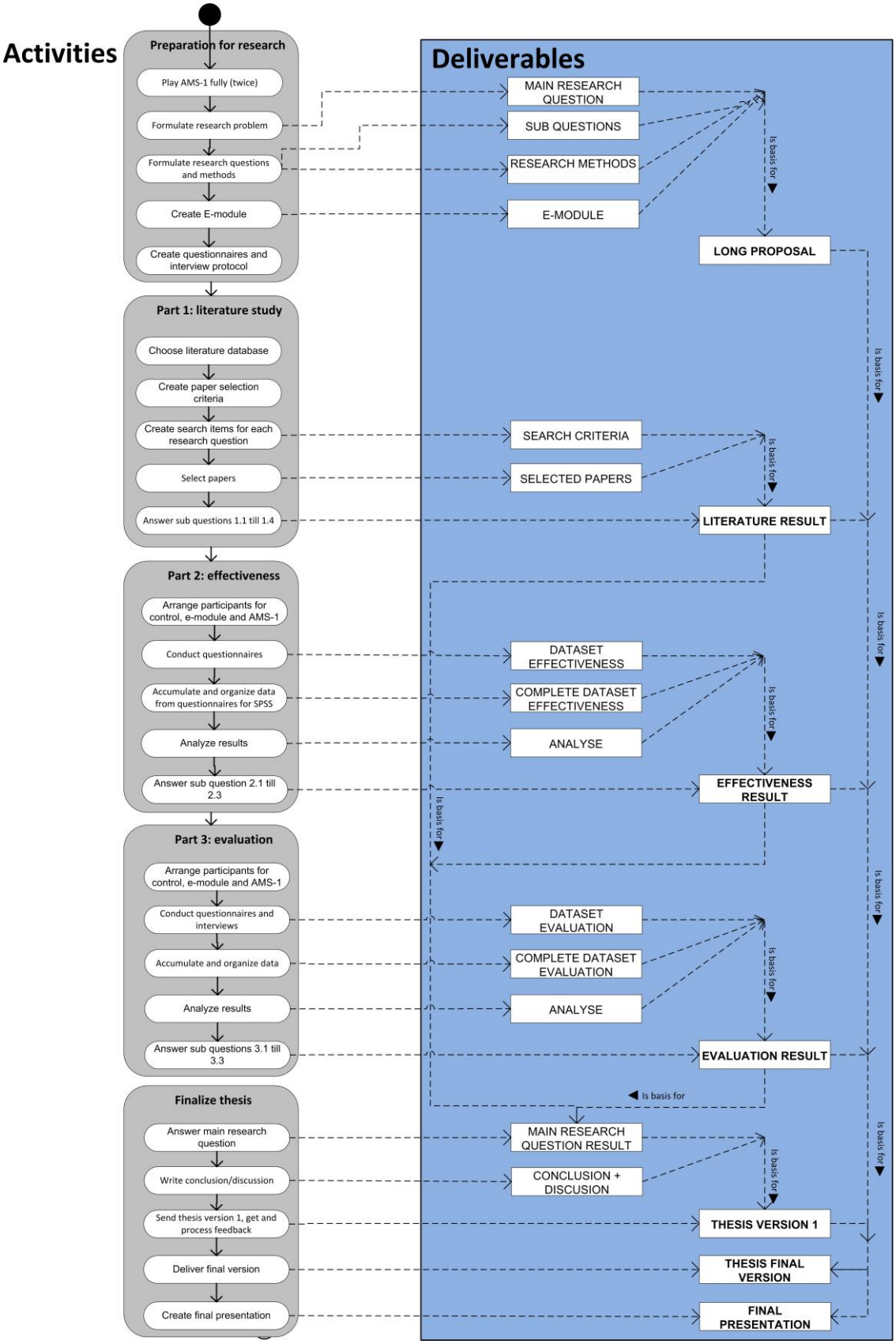


Figure 1, PDD of the research project

1.5 Contributions

The importance of this research lies at different point of views in respect to social and scientific relevance. These different views are discussed below.

Patients - The staggering numbers in the literature (see introduction) showed that a lot of patients admitted to a hospital suffer unintended harm. This has to do with the topic Patient Safety, which is concerned with prevention of errors and adverse events associated with health care. It raised the topic Patient Safety to a major concern in health care and among policymakers. The relevance of this research to the patient is to see if we can improve the quality patient safety by raising Patient Safety skills of doctors with the use of the e-module and/or the serious game AMS-1. Consequentially, this will lower preventable errors and adverse events to patients associated with health care.

Medicine students - Young doctors must learn to deal with the stress of being responsible for many sick patients simultaneously, while effectively communicating with superiors and nurses, sometimes in matters of life and death under considerable time pressure. Taken together, these issues can be a serious patient safety risk. Medical curricula teach doctors about diseases and illnesses, but insufficiently quip them to cope with the demanding and rapidly evolving situations in which they have to apply that knowledge. If this research proves that the serious game AMS-1 can help to fill this void, then AMS-1 can be used as a training tool for medical students to improve patient safety. It could offer a safe place to practice decision making under extreme pressure, so young doctors can improve their style and become comfortable in their ability to cope.

Medical Centers – In the introduction it has become clear that there is an urgent call to improve the patient safety skills of doctors. This begins at the education of medical students in their fifth year before their first clerkship starts. Commissioned by the University medical Center hospital the game serious game AMS-1 has been created. The end-goal of the UMC is to distribute AMS-1 as a commercialized product to other medical centers to improve patient safety skills. Validation of the game is needed before this is even remotely possible. This research will could provide validation on effectiveness of the game (hospital point of view) and validation on the user experience (medical student point of view).
Literature [5] shows that patient safety education into the medical school curriculum is most commonly implemented in medical schools in developed countries such as the United States of America and the United Kingdom, that there is little consensus on the best teaching approach and the strength of the underlying evidence is limited. Not only does this tend to reveal the e-learning module and the serious game AMS-1 is a good training tool with underlying evidence, but it can also provide a validated tool for improving patient safety in medical centers where patient safety is not yet in the medical school curriculum. For instance, the Erasmus University wants to know if they can use AMS-1 as a training tool for their medical students to improve patient safety skills in this introduction week or if they should stick to a cheaper way of educating on patient safety with the use of the e-module that was created by themselves.
Interesting for the Medical Centers is to see the difference in effect of a e-module compared to the serious game AMS-1. This is because there is a big difference in costs. Creating a serious game is by far not cheap, whereas a e-module is. If AMS-1 doesn't prove to be more effective than a e-module then costs will have a big impact on which learning method will be chosen.

Serious gaming research – The topic of Serious Gaming in de medical domain is still young and immature. A lot of literature can be found on the potential and optimism games can have for

learning, but several authors have noted that there has been a gap of high quality empirical evidence to support these claims [22,23]. Also, no systematic literature that examines the literature on serious games in the medical domain can be found yet. It is interesting to examine serious games in the medical domain in regards to the potential positive impact of gaming, with respect to learning, skill enhancement and engagement. The combination of these factors will make the systematic literature in this research interesting for science. More and more games for the medical domain are created, so doctors can improve themselves without harming patients. In the last decade, research and interest in the potential of gaming on learning has exploded, as has the diversity of games themselves [21]. Creating such a serious game for the medical domains is not an easy tasks as many parties involved and a lot of factors need to be taken into account to make the game effective and create a good user experience. This research looks at serious gaming attributes that are important for making the game effective and giving a good the user a good experience. It is interesting to look at what exactly makes a game appropriate for learning by examining the game attributes. Without an accurate understanding of the root factors influencing the effectiveness of serious games, these training tools cannot be harnessed to their fullest potential. The importance for science lies in the fact that these game attributes can then be taken into account in the development of other serious games in the medical domain.

2 Part 1: Literature Review

In this chapter we'll look into five subjects which concern this research. The goal is to get a better understanding of the current situation and knowledge of two subjects that are related to this research. The first subject is Patient safety, followed by the second subject Serious Gaming. These two subjects are divided in sub-questions to get a more detailed view of the matter. In this way, our results can more easily be compared to other researchers/literature in the discussion.

2.1 Research approach

Pubmed will be used as the source for our data. The intention was to use two databases, but after finding enough relevant papers on Pubmed to get a clear view of the current situation, it was decided that Pubmed alone was sufficient. For each subject different search terms are used (table 1). Due to the scope of this research this will not be a systematic literature. Paper selection was based on the search keywords (Table 1), the title of the paper and the contents of the abstract.

Patient safety

Sub-question 1.1: *Patient safety, why is this such an important subject?*

Sub-question 1.2: *Patient safety in the Netherlands, what is the current situation?*

Sub-question 1.3: *Patient education, what is the current situation in patient safety education regarding the medical school curriculum in different countries.*

Serious gaming

Sub-question 1.4: *Is there empirical evidence about the positive impacts and outcomes of computer games ad serious games with respect to learning in the medical domain?*

Topic	1.1 Patient Safety	1.2 Patient Safety in the Netherlands	1.3 Patient safety education	1.4 Evidence serious games in medical domain
Keywords for PubMed	No search terms were used. This part will mostly be based on the report: To err human 1999; building a better healthcare system. The rest of the articles were found during the writing of this Thesis.	"patient safety" AND "Netherlands" AND "medical errors"	medical errors AND patient safety AND medical education AND curriculum AND teach AND medical student AND undergraduate	games AND learning AND skill AND medical Also a few papers that were used in the papers found with these search terms.
results		54	23	12
Selected	7	5	8	9

Table 1: Key words used in specific topic searches using PubMed.

2.2 Results

2.1.1 Patient safety

Patient Safety is concerned with the prevention of medical errors and adverse events in hospitals to patients associated with health care [22]. Medical errors can be defined as the failure of a planned action to be completed as intended or the use of a wrong plan to achieve a goal [23]. An adverse event is defined as a negative outcome resulting from a medical intervention that is not due to the underlying condition of the patient [24]. Thus, patient safety is ensured in a situation in which the patient will not suffer, or has only a slight risk of suffering any damage caused by health care professionals who are not acting in accordance with the professional standards or by failures in the care system.

Patient safety has always been a known subject in the medical domain, but when did patient safety get the increase in attention and why? It began in the year 1999 when the American Institute of Medicine (IoM) released the report *''To Err human''* [23]. The Institute of Medicine (IOM) is an independent, nonprofit organization that works outside of government to provide unbiased and authoritative advice to decision makers and the public [25]. This influential report had over 8000 citations in twelve years and changed the view on health care quality by focusing on patient safety. The report was based upon the analysis of multiples studies by a variety of organizations. It concluded that the healthcare in the United States is not as safe as it should be and can be. More than 1 million preventable adverse events occur each year in United States, and between 44.000 to 98.000 people die each year due to these preventable medical errors. In comparison, this is a higher death toll then caused by motor vehicle, AIDS or breast cancer. Next to the cost in human lives there is another significant toll. The report shows that as a result of medical errors and adverse events in between 17 billion and 29 billion dollars per year are wasted in hospitals nationwide, caused by the need for additional care necessitated by the errors, lost income and disability. These errors and adverse events also influence psychological factors in the medical domain; professionals pay with the loss of morale and frustration at not being able to provide the best care possible and patients lose trust and get diminished satisfaction from their health care [23]. *To Err Human [26]* lays out a comprehensive strategy by which government, health care providers, industry and consumers can improve to reduce preventable errors. One of the four strategies that is given is about developing methods to educate about patient safety. This is where the method serious gaming can play a role, in our case specifically Air Medic Sky-1.

In 2001 another report from the IoM named ''*Crossing the quality chasm*'' [26] followed "*To Err is Human*". The report urgently called for change to the health care system processes to improve the quality of health to reduce preventable medical errors. With the considerable numbers that these two documents showed, they were instrumental in raising the topic Patient Safety to a major concern in health care and among policymakers. Since that moment, several other studies documented the extent of medical errors and underlined the need for patient safety education.

This is not just a problem in the United States, but worldwide. It is estimated that in EU Member States between 8% and 12% of patients admitted to hospitals suffer from adverse events whilst receiving Healthcare [27]. The European Centre for Disease prevention and control (ECDC) estimates that healthcare associated preventable infections occur in 5% of hospitalized patients. This would equate to 4.1 million patients a year in the EU. They also estimated that about 37.000 preventable deaths occur [27]. Since it is such a serious concern, the Council of the European Union adopted recommendations on patient safety in 2010. One

of the recommendations is about education and training of healthcare workers, focusing on patient safety.

As has been indicated by the To Err human report, these staggering patient safety numbers also effect the patients perception of health care. The so called Eurobarometer survey has been conducted in 2009 with the main objective of exploring Europeans' perceptions regarding patient safety and their attitudes toward the quality of healthcare in their country and cross-border [28]. This was done among 27 European countries with 1000participents per country. The main results were that nearly half of the respondents feel they could be harmed their health care, 25% claim that they or a member of their family have experienced an adverse event with healthcare and respondents from 16 EU countries consider the quality of healthcare in their country as worse than in other countries. When judging healthcare quality, most participants thought the most important criterium is a well-trained medical staff. Again this is where the method serious gaming can play a role, in our case specifically Air Medic Sky-1.

2.1.2 Patient safety in the Netherlands

In the Netherlands, the topic Patient Safety has and still is getting a lot of attention, both in practice and scientifically. This section will look at the numbers surrounding patient safety in The Netherlands and what important actions have been taken from year 2004 until 2012.

In the year 2005 and 2006, 21 hospitals participated in a study where patient record from 2004 were read thoroughly [29]. The total amount of patient records read was 7926. These were the results. In 2004, 1.3million people were hospitalized and with 5.7% of these patient an adverse event occurred. Although this is less than the estimate of Europa as total (8-12%) [27], it is still very high: 76.000 patients in total. Of these patients, 5% (around 4.000) ended up with permanent damage and 8% (around 6.000) died due to adverse events. 10.7% of the total deaths in 2004 that occurred in Dutch hospitals was due to an adverse events of which 4.1% could have been prevented. The study also showed that hospital stays were on average 10.3 days longer as a results of adverse events and on average 9.1 days longer when these adverse events were not preventable. The average stay in 2004 was 7.3, which is significantly lower. Each day extra in a hospital adds a lot of costs. The direct average extra costs was 4761 euro per preventable adverse event and 3844 euro per not preventable adverse event which includes around 1000 euro for extra actions performed and compensation to the patient.

In response to this research in 2004, doctors, nurses and hospitals took the initiative to create a nationwide patient safety program (www.vmszorg.nl). This program, called ''Prevent harm, work safely'', was launched officially in 2008 and has a durations of five years. This program started in 2008 and included an ambitious plan to improve on the basis of ten substantive themes for the establishment of a safety management system. The main goal of the program is implementation and compliance with these ten themes to reduce preventable adverse events with 50% within the next 5 years, from 2008 until 2013. All Dutch hospitals committed themselves to the program. This has been agreed with the starters of the vmszorg.nl, the healthcare inspectorate IGZ and the Ministry of Health.

In 2008, The NIVEL research center, who examine the effectiveness and quality of healthcare in the Netherlands, has conducted an similar study as the one in 2004 [30]. As such, the present measurements of 2008 can provide insight into the situation at the beginning of the safety program (goal reducing 50% of the preventable adverse events) and be compared to

the results 5 year later in 2013. The study from 2008 shows that the percentage of adverse events has increased by 2.3% (5.7% to 8.0%) compared to 2004 which is significant, but the percentage of adverse events that was preventable did not change significantly (2.3% to 2.9%). Neither did the preventable deaths increase significantly compared to 2004, from 1.735 to 1.960 (4.1% to 5.5%). A few explanations were given for the numbers not significantly decreasing. First is an increase in complex patients and new technologies that have been used. Second is the fact that since the report of 2004 medical staff see more adverse events because they have a more critical view on the matter. The second explanation would be a positive thing, because it is a first step to be able to learn from mistakes and make changes.

In 2013 the NIVEL research center published an evaluation report in the Dutch hospitals about the VMS program "Prevent harm, work safely'' [31]. The evaluation shows that there is a broad movement in the field of patient safety since it is put in motion. Safety has improved on all of the ten themes, but not all them equally in each hospital. None of the themes have been fully implemented in any hospital, but some come very close. The report concludes that this is why the implementation of the themes is feasible, but more time is needed. Most importantly: was the goal main goal to reduce preventable adverse events with 50% within the next 5 years reached? In 2012, five after the start of the VMS program (2008), damage from preventable adverse events was decreased by 45% and preventable deaths was decreased by 53% [32]. These percentages come from the report published by NIVEL in the year 2011/2012 [33]. This report did a similar study as the one from 2004 [29] and 2008 [30]. In 2011/2012 7.1% was confronted with an adverse event, which is significantly more than 2004 and comparable with 2008. With 1.6% of all patients (2011/2012) that were hospitalized a preventable adverse event occurred, which is 30% less than 2004, and 45% less than 2008. Preventable deaths decreased by 37% compared to 2004 and 53% compared to 2008. In absolute numbers, the deaths that were cause by a preventable medical error decreased from 1960 (2008) to 970 (2011/2012) [32]. The report leader, Cordula Wagner, who works at NIVEL and is professor in patient safety at the Free University Medical Centre, said that the findings showed room for further improvement, like periodic training of health professionals in the application of medical equipment and devices. Also, the national safety campaign needed to "develop continuously in order to anticipate new risks,".

2.1.3 Patient safety education

At the Erasmus University Rotterdam it is standard procedure that all students have an introduction week about Patient Safety one week before their ten weeks of clerkship starts. In this week, students get lessons about the human factors such as surroundings, equipment and education levels in teamwork. The fact that it is never one person that made the mistake, but rather the system that failed is also addressed. In the begin of the year 2014 the subject from human factor changed to self-reflection; taking a critical look at yourself about how you perform during your clerkship. This is called the Regular Patient Safety Education, but there is no real patient safety curriculum. This is also why we research AMS-1 to see if it can be added to this week to extend patient safety education. In this section we are interested in what other countries are doing in the educational domain of patient safety.

It is not a question whether or not patient safety education is needed; it is important for medical students to learn about patient safety topics. From a patients perspective the need for improvement, which starts with education, is high. The numbers for unintended harm to patients speak for themselves in the previous pages. From the medical students perspective, the attitude towards more or improved patients safety education shows positives signs. Firstly this is shown in our own research where in the evaluation questionnaire out of the 62 5th year

medical students 58% strongly agrees and 39% agrees on the question whether it is important for medical students to learn about patient safety topics. In a German report in 2014 [34], 64% of the 167 medical students wished for more education on patient safety issues. No negative attitudes towards learning about patient safety were found with the used search terms.

Traditionally, curriculums for medical students have focused on three major competencies: `medical knowledge`, `technical skills and judgment` and `clinical decision making`. The non-technical and professional competencies such as situational awareness, teamwork and leadership, communication and collaboration, risk management and human factors are not usually explicitly taught or assessed [35]. Recently (2011) the World Health Organization published The Patient Safety Curriculum Guide [35]; a valuable aid in achieving these non-technical skills for patient safety (see appendix A). This guide helps translating patient safety topics into curricular models. It identifies topics that a medical curriculum should contain and provides an overview of patient safety competencies to aim at. Patient Safety Curriculum Guide is used for the education of undergraduate and postgraduate health-care professionals in patient safety. In the lasts years, the guide has been foundational to patient safety education and countries all over the world (Western Pacific, Eastern Mediterranean, Mercosur and central American countries) started to having meetings and launches about this patient safety curriculum guide [36]. This guide will enable and encourage medical schools to include patient safety education. It was also used as a basis for gaining patient safety knowledge in AMS-1.

To date, a lot of research has been done on different patient safety education topics. These studies are mainly to see if a developed method or course on a specific topic, which have not been implemented in a curriculum, improves patient safety skills. Interesting to note is that most of these studies are applicable to one of the topics in the WHO curriculum guide. A few examples and results will be given. A study from 2011 [37] states that prescribing writing errors are well documented medical literature, but have remained an issue in patient safety for many years. Yet education in safe prescribing seems to be lacking. The research looks into 'Check and Correct' which has been developed by the Hospital in West Sussex developed, reaffirming the assertion that further education is required in safe prescribing and concluding that learning methods as 'Check and Correct' works.
In 2012 the University of Miami developed a mandatory week-long patient safety course for incoming third-year medical students [38]. The goals of this course, scheduled right before the start of clinical rotations, are to introduce students to the subject of patient safety, focus their attention on the role of teamwork and communication in providing quality care, and create an awareness of safe practices. A total of 122 students completed the course and 93% responded that the course improved their patient safety knowledge and skills.
In 2013 a research [18] designed a non-technical skills training (a half –or full day intervention) to see if it can enhance patient safety for a number of mixed groups of undergraduate medical students and doctors in postgraduate training. The researched showed that satisfaction was high and patient safety attitudes have improved post-intervention.

As of yet, there are relatively few countries and medical colleges that have followed the trend by implementing patient safety education. The aim of a study from 2009 was to describe the current patient safety curricula at U.S. and Canadian medical schools in the year 2006 [40]. A survey was mailed to institutional members of the Clerkship Directors in Internal Medicine at U.S. and Canada. With a response rate of 76%, only 25% (83/110) of the institutional members reported that their schools had explicit patient safety curricula. No relationship was found between the presence of a curriculum and demographic factors. These curricula all use

lectures and small-group instructions. Topics included analysis of medical errors, reporting adverse incidents, analysis of medical errors, physician order writing, and national patent safety goals. Of the 25%, 72% agreed that patient safety instruction should occur during medical school. The research concluded that few schools have implemented specific patient safety curricula. A more recent research could not be found for the U.S. and Canada on the curricula situation.

Although a of these pilot studies have been published on the education of patient safety, only one systematic review has been done on the success of implemented patient safety education curriculum for undergraduate. This systematic review is from 2011 and searched a total of seven databases for literature [41]. A total of 1481 studies were identified of which a total of seven studies were selected with the use of 13 item quality criteria.
Most studies were from the USA and the patient safety courses were implemented in the third year. The teachers are a mix of interdisciplinary professionals, including clinicians, ethicists and medical education experts. The courses were not intergraded into the formal undergraduate medical education system, rather a selective course or incorporated into a clinical rotation of internship training like Erasmus University Rotterdam. With an average of nine and a half hours, the duration of the courses varied a lot from 4 hours to 30 even hours. There were a total of eight different teaching formats: interactive lectures/discussions, recommended texts, case-based discussions, seminars, small group discussions, role play, interdisciplinary team work and videotaped simulation with a standardized patient.
All seven studies reported the effect of the course on knowledge, skill and attitude, but there was no uniform criteria to evaluate the effectiveness of teaching. In the studies, all questionnaires were self-made. The knowledge lectures were about medical errors, rated and types of adverse events, error classification, contributing factors to medical errors and mechanisms for learning from errors. Out of the seven courses, six reported to have achieved an increase in knowledge. Patient safety skill training lectures were about recognition of error, dealing with error, reporting and learning from error and supporting others involved in error. Only four out of the seven studies showed an increase in patient safety skills. Lastly the attitude towards patient safety was explored, focusing on an understanding of a just culture, willingness to learn from mistakes, being prepared to acknowledge and deal with error, being prepared to reflect on practice, and aspects of trust and respect. Only three studies out of seven evaluated that the attitude towards patient safety was indeed improved.

2.1.4 Evidence serious games in the medical domain

In the past decades, simulations have been used as a method for education. In the medical field simulations are done with the help of human actors, dolls and animal cadavers. But with the ever-increasing computational power, simulation options have increased to serious gaming. Serious gaming is a new field in medical education, which has the potential to become an important tool for healthcare professionals for learning a wide range of clinical skills. Serious gaming is defined as games that are designed to entertain players as they educate, train, or change behaviour [42]. Most often this primary purpose is learning. It is used in industries like defence, education, health care, engineering and scientific exploration. The first serious game ever made was Army Battlezone in 1980, designed for military training. Since that moment serious games have gotten more attention over the last decades, but more importantly they have become more realistic due to technology (graphics, new controls etc.). In this section we will look at papers where in the medical domain serious game is the main subject.

The attitude of medical student towards using video games for medical education is rather positive. A study in 2010 was conducted on 217 medical students attitudes towards

videogames for the medical domain [43]. About half were female and half considered themself an experienced gamer. 98% Of the participants liked the idea of using games for education and 96% felt that their education should make more use of it. About 80% believed that video games can add educational value. Almost all participants (97%) felt that whether or not it was fun and developed their medical skills (90%) were the most important factors for using the game in their own free time. Also in this AMS-1 study participants show a positive attitude towards using a serious gaming to increase medical skills. From the 31 participants, 87% replied that they like the idea of using a serious game to learn more about patient safety.

Research shows that in the medical domain, games can help to develop a range of skills: surgical procedures, strategic planning, team collaboration, interactive communication and negotiating skills [44] [45]. In 2011 a systematic literature review was conducted to evaluate the current status of serious gaming in medicine [46]. With the use of the database PubMed, 515 articles were found of which 45 were relevant. The systematic literature review showed the following. Studies show that doctors who are experienced in gaming from childhood games can translate those skills to medical simulations; they are associated with greater surgical skill, especially for laparoscopy. Most participants enjoy serious games as a learning tool, but do find it stressful in some situations. Two skills that have been proven to increase with medical serious gaming are: those in laparoscopic virtual reality and better decision-making. Studies that have looked into the difference between traditional learning methods and serious gaming are scarce. One study shows that serious gaming was more effective in triage training than traditional training. The studies that were found in this systematic literature suggest that serious gaming is likely to be an effective training method. However, there is a paucity of studies showing the conclusive clinical benefit of serious gaming.

Although this is beyond the scope, it is worth mentioning that serious games in the medical domain is not only for people who work in the medical domain. For example, two studies have demonstrated that Wii-based (Nintendo console) serious gaming is an effective alternative for rehabilitation therapy. It promotes motor recovery after stroke. Another example is research in patient oriented disease management and prevention [47]. The research shows successful use of a serious game to prevent obesity and type 2 diabetes.

Some criticism can be found for serious gaming as a learning tool for training medical skills [48]. The most commonly used criticism is that the colourful, dynamic world of a serious game will distract the player from the learning process and that the generation gap in gaming experience makes it a difficult learning method for some. Whatever the criticism might be, it is clear that everything has to be taken in consideration when creating a game; lives are at stake when we are talking about education in the medical sector and the need to prove a game's effectiveness first is clear.

2.3 Basis of this research

These finding form the basis for this research and answers the sub questions 1.1 till 1.4. We will keep these finding in mind as we proceed further into this research. If a finding from this literature review can be related to findings of part 2: the effectiveness research and/or part 3: evaluation, then this will be written down in the conclusion of these parts.

3 Part 2: effectiveness

In the previous chapter we have seen why patient safety is such an important subject all over the world, what the current situation of patient safety education in medical schools is and what empirical evidence there is of serous gaming in the medical domain. In this chapter we will measure the effectiveness of AMS-1 on *knowledge, self-rated skill and self-reported behavior*. First we will look at the research approach; the procedure, the research population and the research design. After this the statistical program SPSS will be used to execute the statistical analysis; the results will be shown and from these results conclusions will be drawn.

3.1 Research approach

3.1.1 Procedure

To measure the effectiveness of AMS-1, a quasi-experiment will be conducted. This is done with a control group and two intervention groups. The control group will follow the Regular Patient Safety education. The two intervention groups will follow the Regular Patient Safety education with either the e-module (first intervention group) or the game AMS-1 (second intervention group) on Patient safety.

To measure the effectiveness of the e-learning module and AMS-1 level two of Kirkpatrick's framework [24] will be used. Kirkpatrick's framework is perhaps the best known evaluation methodology for judging learning processes. Level two of the framework states that if you want to measure the effect of an educational intervention, then the difference in knowledge, behavior and skill needs to be measured. It fits perfectly with the concept of AMS-1 as can be seen on page 8: the three learning goals of AMS-1. All three groups will undergo the same pre, during and post questionnaires.

Knowledge
Knowledge on patient safety will be measured with a written test, which is administered before and after the clerkship to the three different groups. The knowledge test consists of 126 multiple choice questions; 100 true/ false questions, 11 three options MC questions, 10 four options MC questions, 3 five options MC questions and 2 matching questions. Patient safety is a big topic which can be divided into many sub-topics. The Erasmus Medical Centrum Rotterdam has the following patient curriculum topics:

a) Patient Safety: Introduction to Patient Safety, The 4 challenges of Patient Safety, 9 Red Flags
b) Communication: Pre-briefing, Assertive Statement, Call Back & Read Back, Handovers
c) Focus under Stress: Focus on your Patient, Focus on the Now-through centered breathing,
d) Teamwork: Shared mental model, Teamwork: Ken Catchpole –teamwork to control stress level Using Resources
e) Sleep Deprivation: Signs of Sleep Deprivation
f) Depression: Signs of Depression, Depression in Teammates, Depression & Drug Abuse, Making Mistakes when depressed

The regular patient safety education, e-module and AMS-1 are designed to teach the participants these topics and thus, the knowledge questionnaire contains questions about each one. The full patient safety knowledge test can be found in Appendix B.

Self-reported skill
To measure the effectiveness of the subject skill regarding patient safety, the self-efficacy will be tested before and after the clerkship for the three different groups. Self-efficacy is defined as a person's judgment of their capability to successfully perform a specific task [50]. To assess self-reported skill regarding patient safety, a 12-item questionnaire was designed by the UMC Utrecht. Students can rate their self-confidence related to specific patient safety tasks and behavior, by giving a score from 0-100 in an internet application (0= I cannot do this, 100= I can do this perfectly). The 12-item self-efficacy questionnaire can be found in Appendix C.

Self-reported behavior
Changing the behavior towards stress of students is one of the goals of AMS-1. This is done by teaching them how to perform effective self-management, for instance knowing how to use breathing exercises to deal with stress. We want to measure the effectiveness of AMS-1 compared to the e-learning module and the control group in regard to lowering the stress that students experience during clinical training. The *Perceived Stress Scale* (PSS) will be used [51]. It is the most widely used psychological instrument for measuring the perception of stress. It is a measure of the degree to which situations in one's life are appraised as stressful. Evidence for validity was gained through studies such as failure to quit smoking, failure among diabetics to control blood sugar levels and greater vulnerability to stressful life-even-elicited depressive symptoms.
Items were designed to tap how unpredictable, uncontrollable, and overloaded respondents find their lives. The scale also includes a number of direct queries about current levels of experienced stress. The items are easy to understand, and the response alternatives are simple to grasp. Moreover, the questions are of a general nature and are hence relatively free of content specific to any subpopulation group. The questions in the PSS ask about feelings and thoughts during the last month. In each case, respondents are asked how often they felt a certain way.
PSS scores are obtained by reversing responses (e.g., 0 = 4, 1 = 3, 2 = 2, 3 = 1 & 4 = 0) to the four positively stated items (items 4, 5, 7, & 8) and then summing across all scale items. A short 4 item scale can be made from questions 2, 4, 5 and 10 of the PSS 10 item scale.
In this research we will use question 3, 6 and 8 to the students online once a week during their clerkship. High scores indicate a high level of perceived stress. Two questions on the awareness of adverse events were included; high scores indicate a high level of awareness. The PSS questionnaire can be found in Appendix D.

3.1.2 Research population

The research population consists of fifth year medical students at the Erasmus University Medical Center, doing their preparatory education and 10-week Internal Medicine Clerkship (first clerkship). In general, a group of 13 students start their clerkships internal medicine at Erasmus MC or regional hospitals every 2 weeks. The group of 5th year medical students is typically around 23/24 years old, 60-70% female. The fieldwork took around 12 months. Data-collection for the control group was from April 2013 until September 2013.

Students from these groups were asked to participate in the study. For the control group, a total of 4 groups were asked, with 13 students in each group (52 students). The bruto response rate for the control group was 75% (39 students) and netto response rate 65% (34

students). For the game and e-module group (intervention groups), students were also asked to participate in the study. Data-collection for the game and e-module group was from August 2013 until Feb 2014. When they consented they were ad random admitted to either the game or the e-module group and received the game or the link to the module on the first day of the preparatory education week (Monday). We promised to raffle an i-pad among participating students. For the interventions groups (were asked together and are randomized), a total of 12 groups, with 13 students in each group (156students) were asked to participate. The bruto response rate was 58% (90 students) and the netto response rate was 42% (65 students).

Knowledge
For the knowledge test, our control group has 34 participants (44% male, 56% female), the E-module group 32 participants (41% male, 59% female) and the AMS-1 group has 33 participants (30% male, 70% female). The average age is unknown, but the probability of it being variant from 23/24 years old is very low. The total research population for the knowledge test is n= 99. The fieldwork took around 12 months.

Self-rated skill
For the self-reported skill, our control group has 39 participants (44% male, 56% female), the E-module group 27 participants (41% male, 59% female) and the AMS-1 group has 30 participants (30% male, 70% female). The average age is unknown, but the probability of it being variant from 23/24 years old is very low. The total research population for the self-rated skill test is n= 96. The fieldwork took around 12 months.

Self-reported behavior
For the attitude, our control group has 27 participants (48% male, 52% female) who continued to give their input for an average of 7.75 weeks. The E-module group has 21 participants (33% male, 67% female) who continued to give their input for an average of 6.38 weeks. The AMS-1 group has 20 participants (10% male, 90% female) who continued to give their input for an average of 6.67 weeks. The average age is unknown, but the probability of it being variant from 23/24 years old is very low. The total research population for the self-reported behavior test is n= 68. The fieldwork took around 12 months.

3.1.3 Research design

The procedure for measuring the changes in the three different groups regarding knowledge, self-reported skill and self-reported are summarized in Table 2 below. On Monday the E-module or AMS-1 will be given to the participants. This implies that they have 4 days to play AMS-1 3-4 hours or read the e-module 1-2hours. The knowledge and Self-rated skill test will be taken on Friday. The attitude test is taken in the following clerkship internal medicine for 10 weeks which starts directly the next Monday. At the end of each week they have from Friday night till Sunday night to fill in the questionnaire.

	Introduction Course before clerkship (1 week)					Clerkship Internal medicine (10 weeks)									
	Monday	Tuesd	Wedne	Thursd	Frid	1	2	3	4	5	6	7	8	9	10
Group 1: control (n=34)															
-Knowledge test					x										
-Self- reported skill test					x										
-Self-reported behavior						x	x	x	x	x	x	x	x	x	x
Group 2: e-module(n=33)															
-Knowledge test					x										
-Self- reported skill test					x										
-Self-reported behavior					x										
-Post: Evaluation quest.						x	x	x	x	x	x	x	x	x	x
Group 3: AMS-1 (n=32)															
-Knowledge test					x										
-Self-reported skill test					x										
-Self-reported behavior					x										
-Post: Evaluation quest.						x	x	x	x	x	x	x	x	x	x

Table 2: Research design

3.2 Results

The software IBM SPSS STATISTICS 22 will be used for every statistical test. For each test (knowledge, self-rated skill, and attitude) the hypothesis will be stated, assumptions will be checked, the right statistical test will be chosen plus executed, the hypothesis will be either accepted or rejected and a conclusion will be drawn.

3.2.1 Knowledge test

Hypothesis

The research questions that will be answered is whether or not the knowledge scores from the three groups are significantly different, and this is the case, which group is/are responsible?

The research question will be answered by analyzing the following hypothesizes

- H_0: There are no differences between the knowledge score means of the samples. Median: group1=group2=group3.
- H_1: There are differences between the knowledge score means of the samples. Median: group1≠group2≠group3.

Raw data and assumptions
Part of the process involves checking the raw data to make sure it can actually be used for ANOVA. It requires to 'pass' six assumptions to give a valid result. Each of the six assumptions will be checked below.

<u>Assumption 1</u>: the dependent variable should be measured on a continuous scale.
This assumption is valid for the data. The score that the participants can get is between 0-97.

<u>Assumption 2</u>: The independent variable should consist of categorical, independent groups.
This assumption is valid for the data. There are three independent categorical groups: Control, E-Module and AMS-1.

<u>Assumption 3</u>: The data should have independence of observation.
This assumption is valid for the data. There are participants in each group with no participants being in more than one group.

<u>Assumption 4</u>: There should be no significant outliers.
This assumption is valid for the data. Outliers have been checked with the use of SPSS version 22. Extreme values and boxplots are not significant.

<u>Assumption 5</u>: The dependent variable should be approximately normally distributed for each group of the independent variable.
This assumption is valid for the data. To test normality the Shapiro-Wilk test of normality has

been used.

Tests of Normality

	Groep	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Kennistoets Bonus	1.0	.108	34	.200*	.963	34	.295
	2.0	.090	33	.200*	.984	33	.884
	3.0	.086	32	.200*	.984	32	.915

The null-hypothesis of this test is that the population is normally distributed. Since p-value is not less than the chosen alpha level, the null-hypothesis is accepted.

Assumption 6: There is homogeneity of variances.
This assumption is valid for the data. To test homogeneity of variances, the Levene's test for homogeneity is used in SPSS. It tests the null hypothesis that the population variances are equal. The P-value is not less than the critical value, thus the null hypothesis is accepted.

Test of Homogeneity of Variances

Kennistoets Bonus

Levene Statistic	df1	df2	Sig.
.837	2	96	.436

Analysis of variance (ANOVA)
For testing the hypothesis ANOVA is used with a 95% confidence interval and all assumptions are assumed. ANOVA will be used to determine if the means are statistically different. Post-Hoc Scheffe's test will be used to see which groups are significantly different from each other. The full results can be found in the Appendix and will be explained further below. We will now discuss the results of these tests:

Descriptives

Knowledge test

	N	Mean	Std. Deviation	Std. Error
Control	34	54.5	7.4	1.3
E-Module	33	61.4	7.0	1.2
Game	32	62.6	6.2	1.1
Total	99	59.4	7.8	.8

ANOVA

Knowledge test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1300.1	2	650.1	13.6	.000
Within Groups	4587.8	96	47.8		
Total	5887.9	98			

multiple Comparisons

Dependent Variable: Knowledge test

Scheffe

(I) Groep	(J) Groep	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Control	E-Module	-7.0*	1.7	.000	-11.2	-2.8
	Game	-8.2	1.7	.000	-12.4	-3.9
E-Module	Control	7.0*	1.7	.000	2.8	11.2
	Game	-1.2	1.7	.791	-5.4	3.1
Game	Control	8.2*	1.7	.000	3.9	12.4
	E-Module	1.2	1.7	.791	-3.1	5.4

*. The mean difference is significant at the 0.05 level.

The mean for the AMS-1 group (62.62, 6.2SD) is higher than the mean of the E-module group (61.45, 7.0SD). The Control group has the lowest mean (54.46, 7.4SD). The ANOVA test shows a significance of p=0.000 (F=13.6; df=96). Hypothesis H_0 is rejected and H_1 is accepted. We see that there is a statistically significant difference between the groups that are compared. The differences between the group Means are not likely due to chance and are probably due to the interventions.

The Post-Hoc Scheffe's test shows that there is a significant difference between the control group versus both the AMS-1 and E-module group (p= 0.000). There is no significant difference between the AMS-1 and E-module group.

Interesting to see is if time learned influences this result. Therefor the same test has been conducted, but only for students who played AMS-1 for at least 2.5hours and students who read the e-module for at least 45 minutes. This gave the same result. The scores also did not significantly differ from the previous test. There was also no correlation between time played and score.

We can conclude that medical students who use the E-module or AMS-1 improve patient safety knowledge, but there is not a significant difference between using the E-module or AMS-1 to improve patient safety.

3.2.2 Self-reported skill test

The second purpose of the e-learning module and AMS-1 is to improve the self-reported skill of fifth year medical students in patient safety related issues. Thus, the effectiveness of reaching that goal has to be measured and be compared between the control, e-learning module and AMS-1 group in order to decide which one can and should be used.

The research questions that will be answered is whether or not the self-reported skill scores from the three groups are significantly different, and this is the case, which group is/are responsible?

The research question will be answered by analyzing the following hypothesizes

- H_0: There are no differences between the self-reported skill score means of the samples. Median: group1 =group2=group3.
- H_1: There are differences between the self-reported skill score means of the samples. Median: group1 ≠group2≠group3.

Raw data and assumptions

Part of the process involves checking the raw data to make sure it can actually be used for the Independent sample T-test. The independent t-test requires to 'pass' seven assumptions to give a valid result. Each of the seven assumptions will be checked below.

<u>Assumption 1:</u> the dependent variable should be measured on a continuous scale.
This assumption is valid for the data. The score that the participants can get is between 0-100.

<u>Assumption 2:</u> The independent variable should consist of categorical, independent groups.
This assumption is valid for the data. There are three independent categorical groups: Control, E-Module and AMS-1.

<u>Assumption 3:</u> The data should have independence of observation.
This assumption is valid for the data. There are participants in each group with no participants being in more than one group.

<u>Assumption 4:</u> There should be no significant outliers.
This assumption is valid for the data. Outliers have been checked with the use of SPSS version 22. Extreme values and boxplots are not significant.

<u>Assumption 5:</u> The dependent variable should be approximately normally distributed for each group of the independent variable.
The data was tested for normality (Shapiro-Wilk test); for 2 of 3 groups data was normally distributed (Control and E-module group). We will take this assumption as valid.

Assumption 6: There is homogeneity of variances.
This assumption is *not valid* for the data. To test homogeneity of variances, the Levene's test for homogeneity is used in SPSS. It tests the null hypothesis that the population variances are equal. The P-value is less than the critical value, thus the null hypothesis is rejected.

Test of Homogeneity of Variances

S_Total

Levene Statistic	df1	df2	Sig.
5.501	2	1241	.004

Assumption 6: The scales are reliable
The reliability of the self-rated skill scale was good (Cronbach's alpha=0.80).

Analysis of variance (ANOVA)
For testing the hypothesis ANOVA is used with a 95% confidence interval and all assumptions are assumed. ANOVA will be used to determine if the means are statistically different in the mean of all scales together. We will now discuss the results of these tests:

Descriptive
Self-reported skill Mean

	N	Mean	Std. Deviation	Std. Error
Control	32	67.8	8.9	1.6
E-Module	34	65.0	10.7	1.8
AMS-1	31	67.8	7.7	1.4
Total	97	66.8	9.2	.9

ANOVA
Self-rated skill Mean

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	177.539	2	88.770	1.046	.356
Within Groups	7981.108	94	84.905		
Total	8158.647	96			

The mean of the control group is 67.8 (SD 8.9), 65.0 (SD=10.7) for the E-module group and 67.8 (SD=7.67) for the AMS-1 group. The ANOVA test shows a significance of 0.356 (F=1.05; df=96). Hypothesis H_0 is accepted and H_1 is rejected. We can conclude that there is no statistically significant difference between the groups compared.
Interesting to see is if time learned influences this result. Therefor the same test has been conducted, but only for students who played AMS-1 for at least 2.5hours and students who read the e-module for at least 45 minutes. This gave the same result. The scores also did not significantly differ from the previous test. There was also no correlation between time played and score.

3.3.3 Self-reported behavior

Hypothesis
The behavior of the students can change in two ways. First by the possibility that the interventions can reduce stress during the clerkships of the participants. Second by the possibility that the interventions can make students more aware of patient safety related issues, thus change their attitude towards these issues during their clerkship.

The research questions that will be answered:

- The research questions that will be answered is whether or not the *stress scores* from the three groups are significantly different, and this is the case, which group is/are responsible?
- The research questions that will be answered is whether or not the *awareness scores* from the three groups are significantly different, and this is the case, which group is/are responsible?

The research question will be answered by analyzing the following hypothesizes

- H_0: There are no differences between the stress score means of the samples. Median: group1 =group2=group3.
- H_1: There are differences between the stress score means of the samples. Median: group1 ≠group2≠group3.

- H_2: There are no differences between the awareness score means of the samples. Median: group1 =group2=group3.
- H_3: There are differences between the awareness score means of the samples. Median: group1 ≠group2≠group3.

Raw data and assumptions
<u>Assumption 1</u>: the dependent variable should be measured on a continuous scale.
This assumption is valid for the data. The score that the participants can get is between 0-5.

<u>Assumption 2</u>: The independent variable should consist of categorical, independent groups.
This assumption is valid for the data. There are three independent categorical groups: Control, E-Module and AMS-1.

<u>Assumption 3</u>: The data should have independence of observation.
This assumption is valid for the data. There are participants in each group with no participants being in more than one group.

<u>Assumption 4</u>: There should be no significant outliers.
This assumption is valid for the data. Outliers have been checked with the use of SPSS version 22. Extreme values and boxplots are not significant.

Assumption 5: The dependent variable should be approximately normally distributed for each group of the independent variable.
This assumption is not valid for the data. To test normality the Shapiro-Wilk test of normality has been used.

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Groep	Statistic	df	Sig.	Statistic	df	Sig.
GemStress	1	.211	26	.004	.912	26	.030
	2	.146	21	.200*	.943	21	.251
	3	.104	20	.200*	.977	20	.894
GemSafety	1	.133	26	.200*	.932	26	.086
	2	.121	21	.200*	.939	21	.212
	3	.150	20	.200*	.847	20	.005

*. This is a lower bound of the true significance.

The null-hypothesis of this test is that the population is normally distributed. Since p-value is less than the chosen alpha level, the null-hypothesis only not accepted for group 1 of the stress test.

Assumption 6: There is homogeneity of variances.
This assumption is valid for the data. To test homogeneity of variances, the Levene's test for homogeneity is used in SPSS. It tests the null hypothesis that the population variances are equal. The P-value is not less than the critical value, thus the null hypothesis is accepted.

Test of Homogeneity of Variances

GemStress

Levene Statistic	df1	df2	Sig.
.405	2	64	.669

Test of Homogeneity of Variances

GemSafety

Levene Statistic	df1	df2	Sig.
.567	2	65	.570

Analysis of variance (ANOVA)
For testing the hypothesis ANOVA is used with a 95% confidence interval and all assumptions are assumed. ANOVA will be used to determine if the means are statistically different. Post-Hoc Scheffe's test will be used to see which groups are significantly different from each other. The full results can be found in the Appendix and will be explained further below. We will now discuss the results of these tests:

Stress

Descriptive

Average stress

	N	Mean	Std. Deviation	Std. Error
Control	27	2.22	.55	.10
E-Module	21	2.38	.40	.09
AMS-1	20	2.13	.41	.09
Total	68	2.24	.47	.06

ANOVA

Average stress

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.66	2	.33	1.52	.23
Within Groups	14.10	65	.22		
Total	14.77	67			

The mean for the control group (2.22, 0.55SD, N=27) is lower than the mean of the E-module group (2.38, 0.40SD, N=21) and the AMS-1 group (2.13, 0.41SD, N=20). The ANOVA test shows a significance of p=0.23. Hypothesis H0 is accepted and H1 is rejected. We can conclude that there is not a statistically significant difference between the groups compared. Also the Post Hoc shows no significant values. The differences between the group Means are not likely due to the interventions.

Interesting to see is if time learned influences this result. Therefor the same test has been conducted, but only for students who played AMS-1 for at least 2.5hours and students who read the e-module for at least 45 minutes. This gave the same result. The scores also did not significantly differ from the previous test. There was also no correlation between time played and score.

We can conclude that medical students who use the E-module or AMS-1 do not have reduced stress level during their clerkship. There is also not a significant difference between using the E-module or AMS-1 to improve the stress level.

Awareness

Descriptive

Awareness

	N	Mean	Std. Deviation	Std. Error
Control	27	1.68	.49	.093
E-Module	21	1.56	.44	.095
Game	20	1.60	.57	.13
Total	68	1.62	.50	.06

ANOVA

Awareness

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.201	2	.101	.402	.671
Within Groups	16.268	65	.250		
Total	16.469	67			

The mean for the control group (1.68, 0.49SD, N=27) is higher than the mean of the E-module group (1.56, 0.44SD, N=21) and the AMS-1 group (1.60, 0.57SD, N=20). The ANOVA test shows a significance of p=0.671. Hypothesis H0 is accepted and H1 is rejected. We can conclude that there is not a statistically significant difference between the groups compared. Also the Post Hoc shows no significant values. The differences between the group Means are not likely due to the interventions.

Interesting to see is if time learned influences this result. Therefor the same test has been conducted, but only for students who played AMS-1 for at least 2.5hours and students who read the e-module for at least 45 minutes. This gave the same result. The scores also did not significantly differ from the previous test. There was also no correlation between time played and score.

We can conclude that medical students who use the E-module or AMS-1 do not make students more aware of patient safety related issues in their clerkship. There is also not a significant difference between using the E-module or AMS-1 to improve awareness.

3.3.4 Conclusion effectiveness results

The knowledge test consisted out of 126 multiple choice questions. The mean score of the AMS-1 group was the highest, followed by the E-module. The control group had the lowest mean. Analysis showed that AMS-1 and E-module had a significant better score than the control group, but there was no significant difference between AMS-1 and the E-module. Thus the conclusion is that medical students who use the E-module or AMS-1 will have improved patient safety knowledge.

The self-reported skill test consisted out of 12-tiem questionnaire. Students can rate their self-confidence related to specific patient safety tasks and behavior, by giving a score from 0-100 in an internet application (0= I cannot do this, 100= I can do this perfectly). The score of the control and E-module group was the same and the AMS-1 group had the lowest score. The difference in score was not significant, concluding that medical students who use the E-module or AMS-1 do not improve their self-reported skill related to patient safety topics.

The self-reported behavior test is split into two subject. The first one is Stress where three questions from the validated Perceived Stress Scale were asked, measuring their stress level on a scale of 1 to 5 during clerkship. The second subject is Awareness, where we use two questions to see whether or not their behavior towards patient safety has changed; do the participants see/experience more events when they have undergone the intervention. For both subjects, no significant differences were found. We can conclude that medical students who use the E-module or AMS-1 do not have a reduced stress level and their awareness of adverse events does not change at their clerkship.

Interesting to see was if time learned influences these result. Therefor the same tests has been conducted, but only for students who played AMS-1 for at least 2.5hours and students who read the e-module for at least 45 minutes. This gave the same results. The scores also did not significantly differ from the previous tests. There was also no correlation between time played and score.

A summery is of the results is given in table 3 below.

	CONTROL (1)	**E-MODULE (2)**	**AMS-1 (3)**
Knowledge n=34			
Mean (max 125)	54.5	61.4	62.6
Significant	No	yes (1vs2)	yes (1vs3)
Self-reported skill n=39, 27, 31			
Mean (max 100)	67.8	64.8	67.9
Significant	No	No	No
Self-reported behavior n=27, 20, 21			
Stress			
Mean (1 to 5 likert)	2.19	2.38	2.13
Significant	No	No	No
Awareness			
Mean (1 to 5 likert)	1.7	1.56	1.6
Significant	No	No	No

Table 3. Summary effectiveness results.

4 Part 3: Evaluation

In the previous chapter we have seen that the interventions only positively (significant) influence the patient safety knowledge of the participants. The interventions had no significant (positive or negative) effect on self-rated skill and self-rated behavior. In this chapter we will look at an user questionnaire that was given after the intervention and eight interviews with participants who played AMS-1. The evaluation gives the ability to look further than the statistical analysis, giving some clarification on the effectiveness research. We will start with the user questionnaire and follow up with the interviews. For each part the research approach will first be given.

4.1 User questionnaire

4.1.1 Research approach

An user questionnaire is a well-used, standard scientific way of getting end user reactions to a piece of software. For evaluating users after they have used AMS-1 or the E-Module, we can use existing questionnaires, long as they are adequate to our specific case. If this is not the case, a new instrument has to be created to fit in. There are a lot questionnaires available online which have been validated many times. Search terms, “Evaluation Questionnaire” and “Usability Questionnaire” have been used to find these questionnaires. The search terms found 19 questionnaires, 10 were selected under the following criteria:

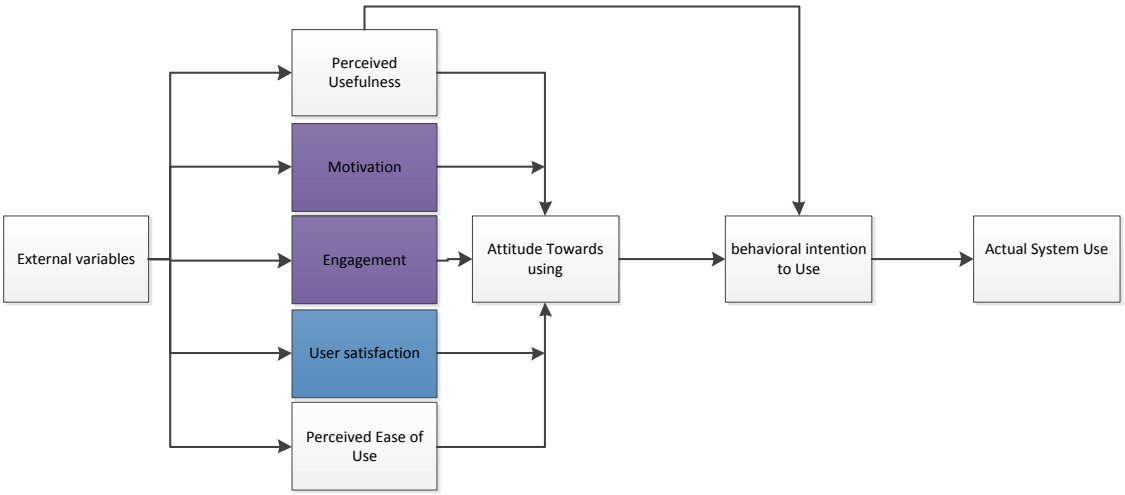
- The had to be scientifically validated questionnaire are scientifically validated
- Clear and understandable questions, applicable for our case.
- Questionnaires are widely used in papers and cited.

These 10 questionnaires are summarized in table 2.

Shortcut name	*Full name*	*Reference*	*# of questions*
QUIS	Questionnaire for Use Interface Satisfaction	Chin et al, 1988 [52]	27
PUEU	Perceived Usefulness and Ease of Use	Davis, 1989 [53]	12
NAU	Nielsen’s Heuristic Evaluation	Nielsen, 1993 [54]	5
NHE	Nielsen’s Heuristic Evaluation	Nielsen, 1993 [55]	10
CSUQ	Computer System Usability Questionnaire	Lewis, 1995 [56]	19
PHUE	Practical Heuristics for Usability Evaluation	Perlman, 1997 [57]	13
ASQ	After Scenario Questionnaire	Lewis, 1995 [58]	3
PUTQ	Purdue Usability Testing Questionnaire	Lin et al, 1997 [59]	100
TAM	Technology Acceptance Model	Davis, F.D., 1989 [60]	30
SUMI	Software usability Measurement Inventory	Jurek et al. [61]	50
USE	Usefulness, Satisfaction, and Ease of Use	Lund, A.M. 2001 [62]	30
MSLQ	Motivated Strategies for Learning Questionnaire	R. Pintrich et al. [63]	81

Table 1. all questionnaires which have been looked into.

After reviewing the questionnaires it became clear none were fully applicable to our specific case: evaluating the game AMS-1 and compare it with the e-module. A new instrument was created: an extension on the Technology Acceptance Model (TAM) [43]. The TAM Model is extended with the USE questionnaire [45] (which has and overlap with TAM), and the MSLQ (on motivation for learning) questionnaire. The final instrument is shown below:


Perceived Usefulness
Motivation
External variables
Engagement
Attitude Towards using
behavioral intention to Use
Actual System Use
User satisfaction
Perceived Ease of Use

4.1.2 Research results

External Variables

In this research external variables are defined as factors that are out of our reach to control and that influence the results. In this research we have three external variables. The first one is gaming experience. For example, if we have a group that for example that is very inexperienced with gaming, then this might be an explanation why some scores of the effectiveness research are low. The second external variable is time played/learned with the intervention. In this research, the participants were told to play AMS-1 for 3-4hours and read the E-module for around 1 hour in the next 4 days. This is considered the average time that needs to be spent to research the effectiveness of the interventions. It was the participants own responsibility to adhere this request and thus out of our control. The third external variable is allowing access to the other group. Clearly the results are influenced when the two intervention groups exchange their learning material. Questions have been put in the evaluation to investigate these external variables and the results can be found below.

<u>Gaming experience</u>

We are dealing with a group that is not unknown to gaming even though the study Medicine is very time consuming. Out of the all 31 participants AMS-1 participants, only 9,68% has no gaming experience. Most participants have basic experience in gaming 38,71% (1-2hours per week), followed by a moderate experience of 25,81% (2-4hours a week), an intermediate experience of 19,35% (4-6hours a week) and as last an expert gaming experience of 6,45% (6+ hours a week).

<u>Timed played (ams-1) and learned (e-module)</u>

Ams-1

Most participants did not play until the end. Out of the 31 participants, only 12,90% finished all aspects of the game for 100%. This is not a strange result. The game has a total of 25 missions which each take around 10mintues to finish, apart from the other two parts of the game. To The game takes around 6-7hours to finish we asked the participants to play 3-4hours including the other parts of the game. Looking at the data the average play time of the students is exactly 3 hours, at the boundary of what we asked. No one played less than one hour or more than 5 hours.

E-module

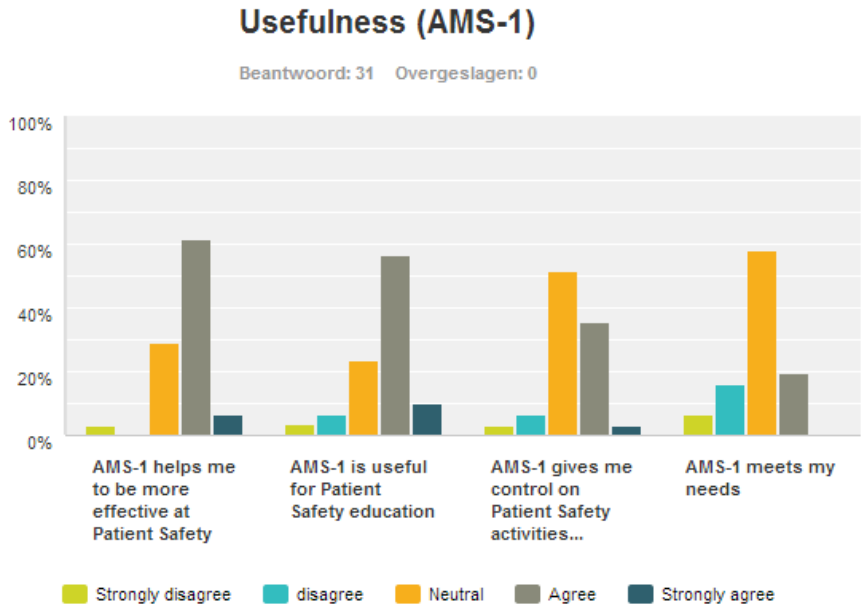
Most of the 31 participants did read until the end, but a big part (22,58%) out of the 31 did not. The reasons were the following: tired, lack of time, too much text and the lack of battery. Still, those who did not read everything read more than 75% which became clear when after asking personally. The average reading time was 54minutes. We estimated that the E-module would take around 60minutes to read.

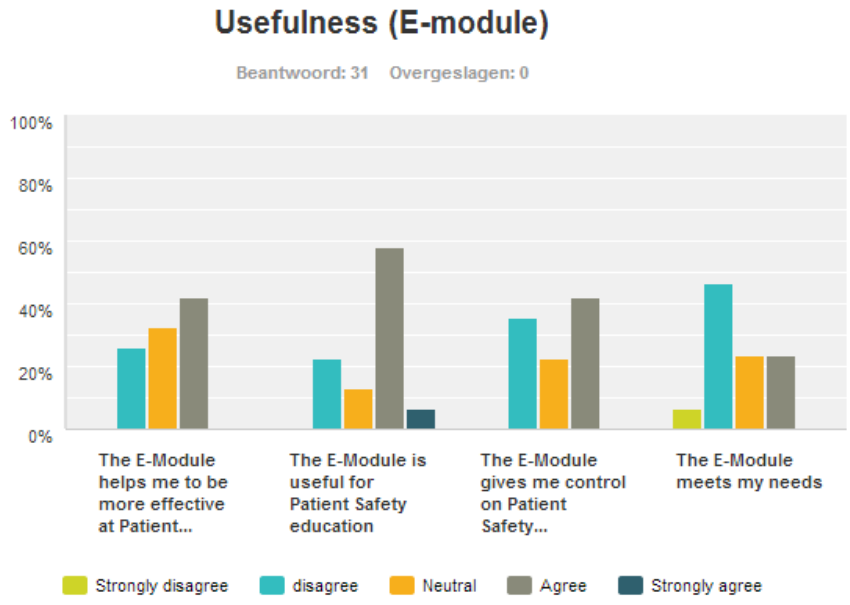
<u>Allowing access to the other group</u>

No one in the intervention week allowed access to their e-module nor AMS-1 to the other group.

Perceived Usefulness

Overall, Usefulness scored high with a mean of 3.38 for AMS-1 and 3,29 for the E-module out of 5.00 where 1.00 stands for strongly disagree and 5 is strongly agree. This average was calculated out of four sub-questions with a good reliability score (Cronbach's alpha = 0.891). The independent sample t-test shows that the difference in score between AMS-1 and the E-module group is not significant (0.19). The sub-questions are summarized in Graph below.

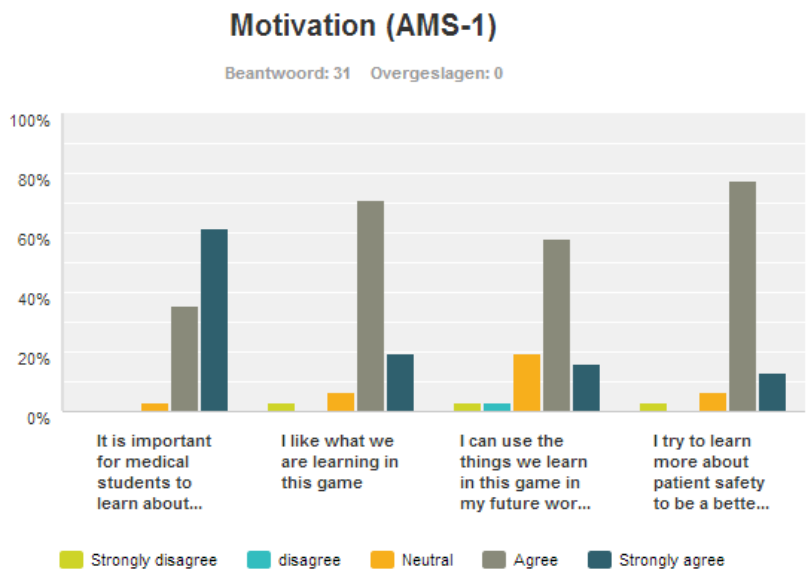


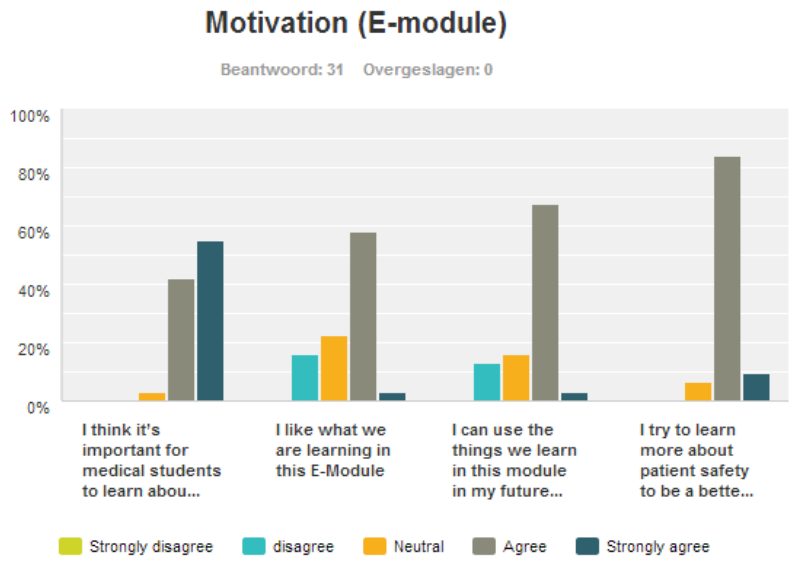


Notable is the biggest score difference is in the first question (mean: AMS-1 = 3.29 , E-module = 3,68). This difference is significant (sig. 0.13). The opinion of the AMS-1 group significantly goes towards AMS-1 making them more effective at Patient Safety than the E-module.

Motivation

Motivation scored high with a mean of 4.07 for AMS-1 and 3,91 for the E-module out of 5.00 where 1.00 stands for strongly disagree and 5 is strongly agree. These are also the highest mean scores given of the whole evaluation. This mean was calculated out of four sub-questions with a reliability score (Cronbach's alpha = 0.71). The independent sample t-test shows that the difference in score between AMS-1 and the E-module group is not significant (2.18). The sub-questions are summarized in Graph 2 below.

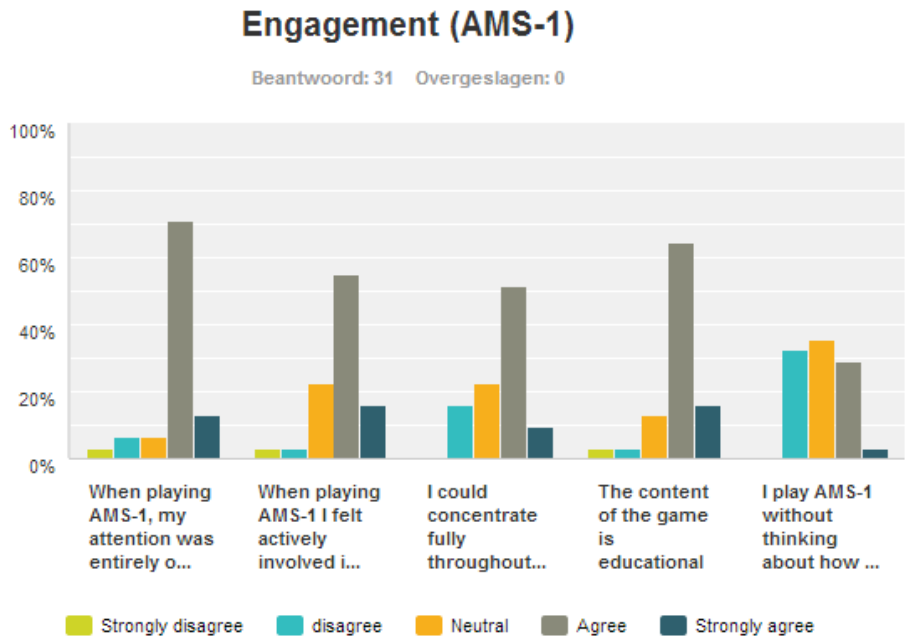


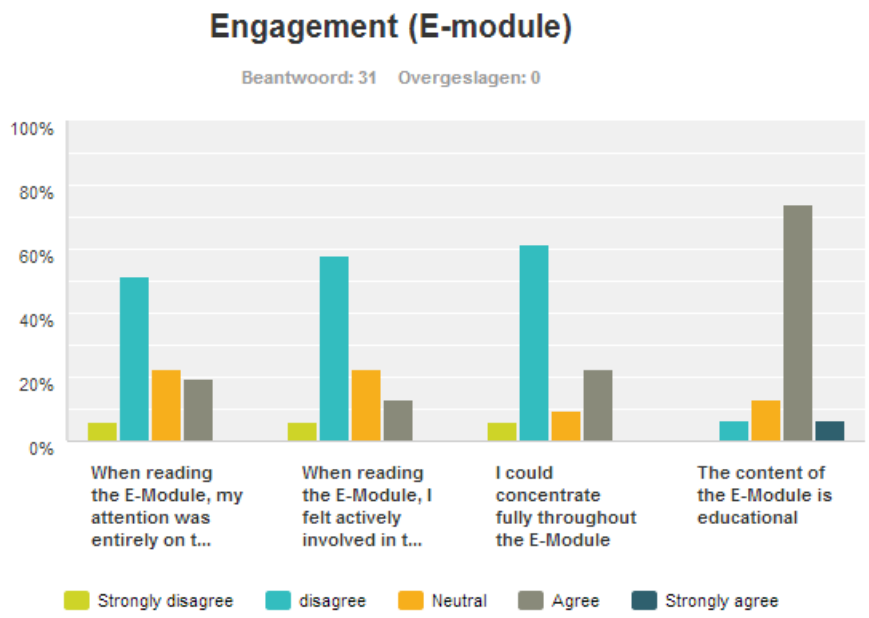


Notable is the high score for ''agree'' and ''Strongly agree'' for all questions. The opinion of both groups goes towards finding it important to learn about Patient Safety subjects, liking what they are learning, able to use what they have learned and trying to learn more about patient safety.

Engagement

Engagement scored high with a mean of 3.8 for AMS-1 and low with 2.8 for the E-module out of 5.00 where 1.00 stands for strongly disagree and 5 is strongly agree. This average was calculated out of four sub-questions with a good reliability score (Cronbach's alpha = 0.807). The independent sample t-test shows that the difference in score between AMS-1 and the E-module group is significant (0.00). The sub-questions are summarized in Graph 3 below.

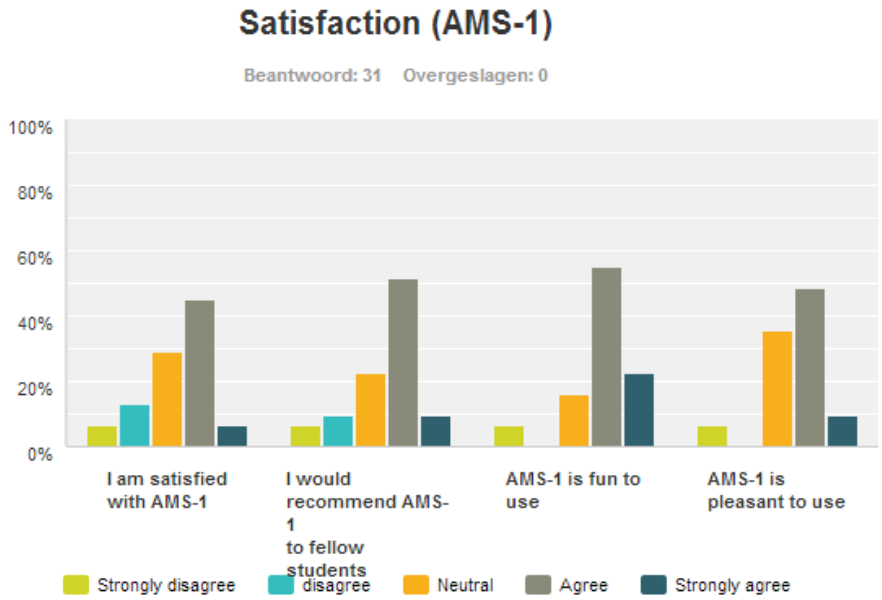


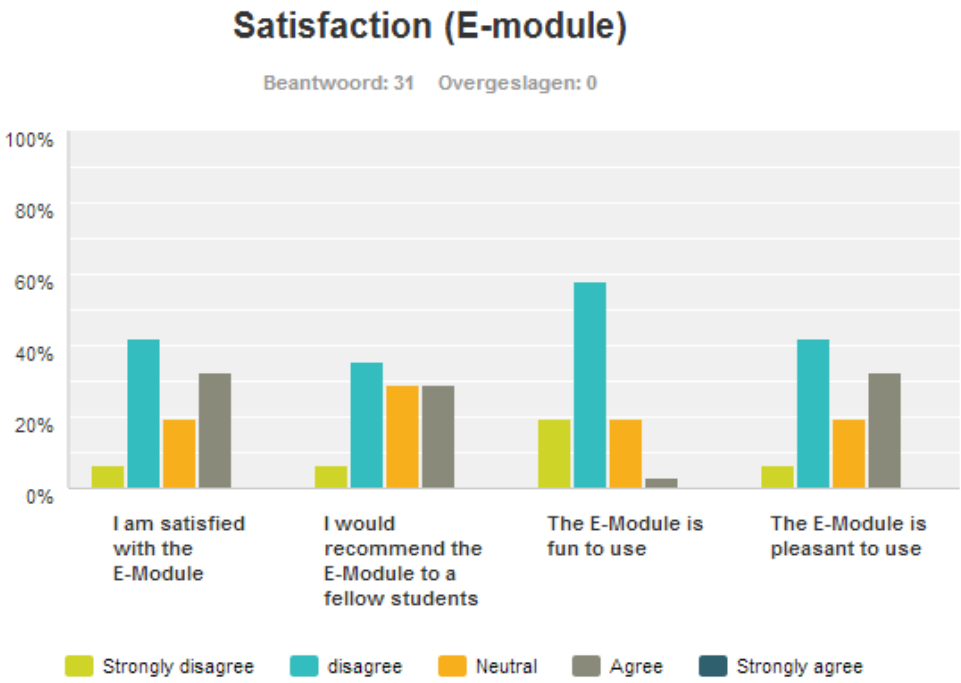


Notable is that the first three questions score significantly lower at the E-module group compared to the AMS-1 group ($p>0.5$). The opinion of the AMS-1 group significantly goes towards that AMS-1is making them have more attention, feel more actively involved and being more concentrated during the intervention.

User satisfaction

satisfaction scored high with mean of 3.5 for AMS-1 and low with 2.5 for the E-module out of 5.00 where 1.00 stands for strongly disagree and 5 is strongly agree. This average was calculated out of four sub-questions with a good reliability score (Cronbach's alpha = 0.878). The independent sample t-test shows that the difference in score between AMS-1 and the E-module group is significant (0.00). The sub-questions are summarized in Graph 4 below.

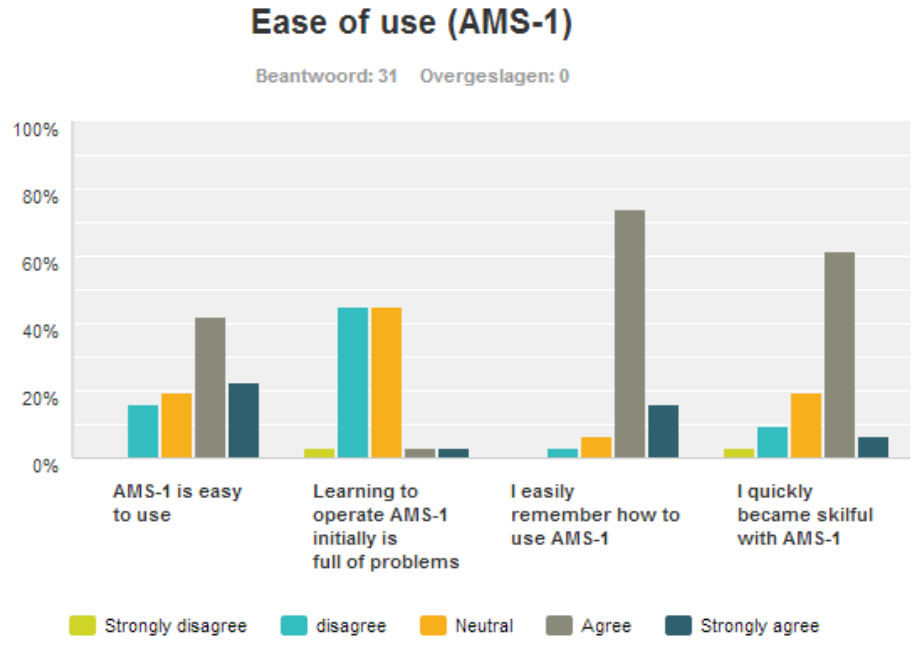


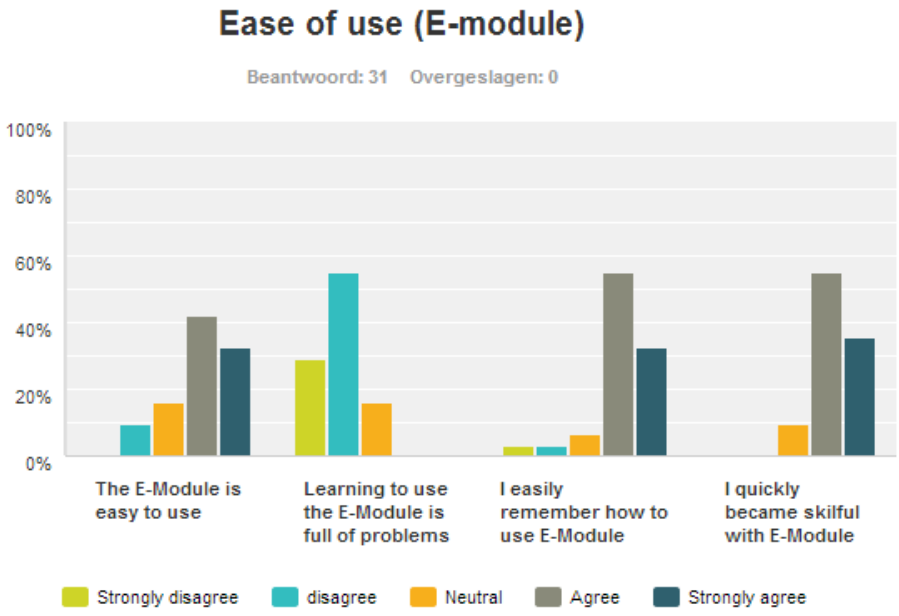


Notable is that the last three questions score significantly higher at the AMS-1 group compared to the E-module group ($p>0.5$). The opinion of the AMS-1 group significantly goes towards that AMS-1 would rather be recommended as a learning tool, is more fun and pleasant to use.

Ease of Use

Ease of use scored high with mean of 3.4 for AMS-1 and 3.5 for the E-module out of 5.00 where 1.00 stands for strongly disagree and 5 is strongly agree. This average was calculated out of four sub-questions with a low reliability score (Cronbach's alpha = 0.678) and does not get higher when items are deleted. The independent sample t-test shows that the difference in score between AMS-1 and the E-module group is not significant (2.12). The sub-questions are summarized in Graph 5 below.

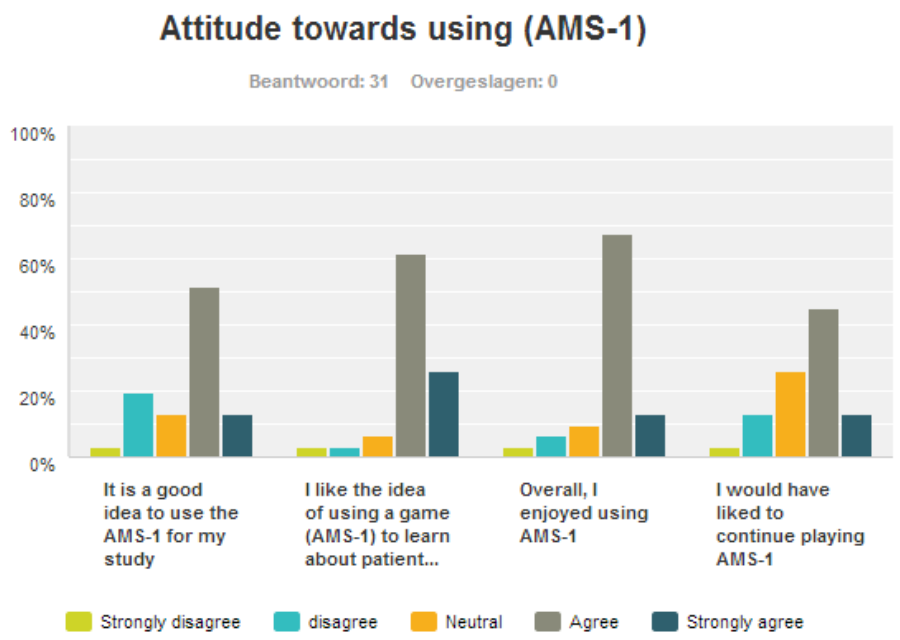


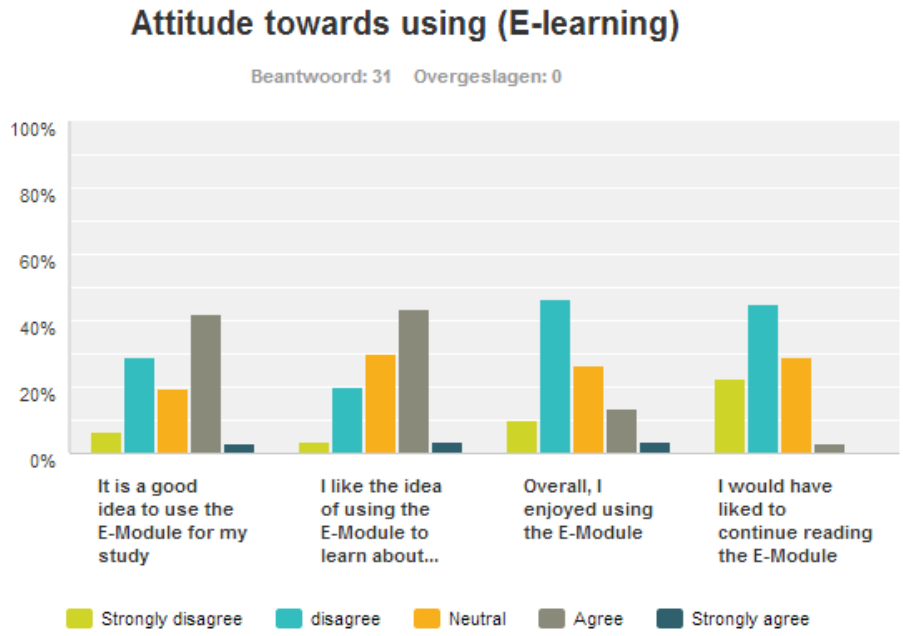


None of the questions are significantly different when the AMS-1 and E-module are compared ($p>.05$). The opinion of both groups significantly goes towards that they are easy to use, not full of problems, easily to remember how to use and easy to become quickly skilled.

Attitude towards using

Attitude scored high with mean of 3.7 for AMS-1 and low with 2.7 for the E-module out of 5.00 where 1.00 stands for strongly disagree and 5 is strongly agree. This average was calculated out of four sub-questions with a very good reliability score (Cronbach's alpha = 0.901). The independent sample t-test shows that the difference in score between AMS-1 and the E-module group is significant (0.00). The sub-questions are summarized in Graph 6 below.





Notable is that the last three questions score significantly higher at the AMS-1 group compared to the E-module group ($p>0.5$). The opinion of the AMS-1 group significantly goes towards that the participants would have liked to use AMS-1 longer, enjoyed played AMS-1 and liked the idea of using AMS-1 to learn more about patient safety.

Feedback

Ams-1

There were a total of 31 participants at AMS-1 and 27 at the E-module who responded to the positive and negative aspects. There was no limit to the amount of feedback that they were allowed to give. Most only gave one positive or negative aspects while others up to three. The feedback has been split up in elements, attributes and quantity. Elements are things the game consist of and attributes are characteristics of the elements. The quantity shows how many participants found it important to note the element. This is summarized in table 2 and 3.

It becomes clear that having fun, playing missions and the topics that are presented in AMS-1 are most appreciated. The breathing exercises were not appreciated. Over half of the feedback was the negative aspects of this element.

Participants found the e-module an easy way to learn, well organized and agreed that the information was important to read. One thing all 27 participants agreed on: reading the e-module was boring or not fun to read. This was likely due to too much text and not enough questions.

Table 2. AMS-1 Positives and negatives

Element	Positive attributes	Quantity	Negative attributes	Quantity
Fun	More fun than regular class, Way of learning, Time flies	9		0
Missions	Generating stress so you get more aware of what is coming, making decisions	9	Not enough challenge, graphics not realistic	2
Topics	Level of professionalism, Educational, Well developed, Based on reality, Interesting, Raising awareness about Patient Safety, Getting to know what to do when.	7	More a game than gaining knowledge, too much on breathing exercises, not enough focus on knowledge, not enough challenge, not in Dutch	7
Breathing exercises	It is like meditating which I love, the feeling of being involved in the game; engagement	3	Does not respond well; frustrating, do not match reality, if you are sweaty the finger clips do not respond well, too long, difficult, bugs, complicated, no measurement of heart rate, could not bring heart rate or breathing down, can't use this in real life; unrealistic	18
Feedback	Good instructions, triggers to asks these kinds of things in real clinic, learns to go to supervisor if there are doubts	2	Not the possibility to read the feedback back	1

Lectures		0	Bad English, unprofessional, skipped because you could go to the fun part (missions), boring but inevitable, no exercises in the lectures	6

Table 3. E-module positives and negatives

Element	Positive attributes	Quantity	Negative attributes	Quantity
Content	Important knowledge to know about, no extra information that is not interesting, written so easy to understand, educational, how to deal with sleep deprivation and stress, information makes sense, examples,	18	Boring, not fun to read, too much text on one page, more questions, language mistakes, hard to keep concentration, not interactive, dull, not in Dutch, no examples, questions don't let you think but repeat the theory.	28
Ease of use	Easy way to learn	6		0
Visual	Good overview of subjects, very organized	2		0

4.2 Interviews

4.2.1 Research approach

To see what aspects of the game AMS-1 students appreciate and dislike, semi-structured interviews have been conducted, sometimes referred to as informal, conversational or 'soft' interviews [64]. The interview consisted out of three parts and each part had a few predetermined questions. Each questions unfolds in a conversational manner offering participants the chance to explore issues they feel are important, instead of answering yes or no.

The first part was asking them openly what they thought of the game. The theoretical idea behind this is that the first subjects they name, are the most important for them, good and negative aspects. This unfolded every time in more questions, exploring these, the areas they found most important to mention.

The second part was about going through each part of the game. Starting with the biofeedback/breathing exercises, followed by the lectures, ending with the missions. Again, asking openly what they thought of this part of the game, letting them name the most important good and negative aspects and continue asking further on that.

The third part was prepared by looking at the evaluation questionnaire before the interview started. Each item in the evaluation questionnaire was discussed and any question that was answered out of the ordinary got the most attention.

In total, seven interviews were conducted, four female and 3 male. One interview was two participants at the same time Participants that were chosen are from the same population group as the effectiveness research and played AMS-1 for at least 2,5hours. The interviews

were conducted in person (3) and with the use of Skype (4). Every interview has been recorded with the phone Samsung Galaxy s4. The average time of the interviews was 20minutes per person. The transcription summary can be found in Appendix G.

4.2.2 Research results

Aspects of the game students appreciate

Each interview started by participants saying how much fun the game was to actually play. Five out of the seven interviews said they wanted to continue playing the game because of the fun they had. Most of this fun came from the mission; it was educational, being able to make decisions, realistic, and generating stress.

One part was explicitly mentioned in every interview: the red flag game. This was a single part of the game where a lecture, about the nine red flags, was put in a game format. Overall opinion was that this was a good example of how you should learn more about patient safety in a game format instead of lectures where you are only listening. Four interviews mentioned they would use what they had learned from the red flag game.

The score system was the most motivating game aspect of AMS-1 as mentioned in the interviews. It kept one participant motivated enough to watch all lectures.

The storyline was the second best mentioned game aspect, but no specific reason was mentioned.

Aspects of the game students dislike

The breathing exercises were most frequently mentioned as a negative aspect of the game in the evaluation questionnaire. This was the same for the interviews, it simply did not work most of the times. Reasons being: susceptible to distraction, prone to errors (sweaty hands and not able to reduce heart rate), realism issues (are you really able to take a step back in clerkship and do a breathing exercise) and taking too long to complete.

The lectures were second most mentioned. As mentioned in the positive aspects, students were positive about the red flag game. Reason being that the lectures were boring, hard to keep attention, low engagement, English not always at its best, . Because of this, not a lot of lectures were watched by these participants.

Even though that 'speak up' (red flag game, say something when you think something is going wrong) was mentioned as a positive aspects, it became clear that participants were divided whether they would use it in their clerkship. There is a culture where speaking up to your supervisor is not a daily practice. Reasons being that students are not confident enough yet in their own knowledge and find it hard to criticize someone who will give them a grade.

Opinions were divided if the game could really reduce stress. On the one hand the biofeedback exercises have potential to learn how to deal with stress. On the other hand participants did not find the missions realistic enough to generate stress. One person said there was not enough pressure, even when he was dealing with multiple patients, simply because you know it is a game, nobody is watching and you can fail without consequences. The suggestion was made that this game is better for AOIS (specialist in training) who have more responsibilities, thus more stress.

All participants agreed that they had too little time to play. This was due to the busy week they had which caused few possibilities to play in that week. All would have liked to continue playing the game. One participant filled in 'Neutral' at a lot of evaluation questions, because he felt that he was not in a position to judge with the hours that he played.

4.3 Conclusion user questionnaire and interviews

The evaluation, gives the ability to look further than the statistical analysis, giving some clarification on the effectiveness research. The evaluation was split up in a questionnaire and interviews. Each of the evaluation items will be discussed on the basis of the results. Results of the interview will be taken directly in these items, combining them. All scores are given are on a likert scale from 1 to 5 where 1 is to strongly disagree and 5 is to strongly agree.

External variables

In this study external variables are defined as factors that are out of our control and that influence the results. In this case we are dealing with three external variables. The first one is gaming experience. For example, if we have a group that is very inexperienced with gaming, then this might be an explanation why some scores of the effectiveness research are low. The second external variable is time played/learned with the intervention. In this study, the participants were told to play AMS-1 for 3-4 hours and read the E-module for around 1 hour in the next 4 days. This is considered the average time that needs to be spent to study the effectiveness of the interventions. It was the participants own responsibility to adhere this request and thus out of our control. The third external variable is allowing access to the other group. Clearly the results are influenced when the two intervention groups exchange their learning material. Questions have been put in the evaluation to investigate these external variables and the results can be found below.

Usefulness

With a mean score of 3.38 for AMS-1 and 3.29 for the E-module we see that students find the interventions useful for learning about patient safety. Most participants agreed that the knowledge they gained throughout the lectures in the interventions was fairly boring. Even though, the interventions score high on usefulness which might have motivated the participants enough to learn for the knowledge test resulting in a significantly better score for the intervention groups. Notable is the question whether or not the intervention makes them more effective at patient safety. Here, AMS-1 scored significant better. The question remains why, because this was not shown in the effectiveness study. The biggest difference between the E-module and AMS-1 is the missions part in AMS-1. Here student have an opportunity to put what they have learned to use on cases that have actually happened in the past. In the questionnaire and in the interviews it became clear that this was the part where participants were most excited about, so this might explain the better score for AMS-1.

Motivation

With a mean score of 4.07 for AMS-1 and 3.91 for the E-module, all questions on motivation were answered with a high score. We see that students find the subject patient safety very important. This is not a coincidence: in our literature study this was also found. A German report in 2014 showed that 64% of 167 medical students wished for more education on patient safety. In our research out of the 62 5th year medical students 58% strongly agrees and 39% agrees on the question whether it is important for medical students to learn about patient safety topics. These facts might show that the numbers that were given in the literature study about preventable adverse events are also known to students and therefore they find it an important subject, and are thus motivated to learn about it. Students like what they are learning in AMS-1 and E-module. This is remarkable for the E-module, because in the evaluation almost no one was positive about the content format. They said it was boring, not fun to read, too much text etc. You would think that this could undermine their motivated to learn, but that wasn't the case. Apparently the bigger picture is more important: gaining important knowledge. This is shown in the following questions that were answered positively: "I can use what we learn for my future work as a doctor" and "I try to learn more about

patient safety to be a better doctor".

Engagement
Engagement scored high on AMS-1 with a mean of 3.8 for and low on the E-module with 2.8. This score difference is significant, The opinion of the AMS-1 group significantly goes towards AMS-1 making them have more attention, feel more actively involved and being more concentrated during the intervention and the E-module is not. Students who have a higher engagement with their learning material are usually performing better, because they are more actively involved, focused and persistent on studying. Even though AMS-1 scores better in engagement, this did not influence the effectiveness tests. Here, the E-module scores similar to AMS-1. The probable causes of this lack of engagement for students of the E-module are summarized in the table Positives and negatives on page 44: it comes down to both having a more enjoyable experience than with regular study material (the mission part of the game played the biggest part in this) and people feeling more in control because they could make their own decisions with AMS-1.

Satisfaction
Satisfaction scored high with a mean score of 3.5 for AMS-1 and low with a mean of 2.5 for the E-module. The opinion of the AMS-1 group significantly goes towards recommending AMS-1 as a learning tool, because it is more fun and pleasant to use. The high score of 3.5 for AMS-1 is surprising: since it got a lot of complains in the evaluation. The breathing exercises were complained about the most for simply being ineffective. The reasons were: it was susceptible to distraction, it was prone to errors, it had realism issues and it took too long to complete. The lectures also got complaints for being boring and not always being explained in the best English. Because of this, not a lot of lectures were watched by these participants. The low score of 2.5 for the E-module is not surprising. It was considered boring, too long and very hard to keep concentrated on, and so on. Satisfaction influences potential leanings. Even though AMS-1 scored better in satisfaction, this did not influence the effectiveness tests. Here, the E-module scores just as well or bad as AMS-1.

Ease of use
Ease of use scored high with a mean score of 3.4 for AMS-1 and a mean of 3.5 for the E-module. The opinion about interventions significantly goes towards it being easy to use, having little problems, being easy to remember how to use and making it easy to quickly become skilled. This result was expected. Before each part of the game there is an extensive tutorial and there was not a single complaint in the evaluation about the ease of use. The participants were already familiar with E-modules, because this is used extensively in their education. Even though there was a high score for AMS-1 in ease of use, some answered very negative. This is probably caused by the problems that the breathing exercises caused. As stated before, it did not always work as intended.

Summary Evaluation of evaluation results game and e-module
The green items are significantly different.

Items	E-module (2)	AMS-1 (3)	p
Perceived Usefulness	3.29	3.38	P>0.5
Motivation	3.91	4.07	P>0.5
Engagement	2.8	3.8	0.000
User satisfaction	2.5	3.5	0.000
Ease of learning	3.5	3.4	P>0.5
Attitude towards using	2.7	3.7	0.000

5 Conclusions and discussion

Patient safety education for medicine students is needed to decrease preventable adverse events in Health Care. In 2011, the University Medical Center Utrecht (The Netherlands), developed the Serious Game ''Air Medic Sky-11'' (AMS-1). The Erasmus Medical Centrum Rotterdam is interested in whether or not they can use this Serious Game as a training tool for their fifth year medical students to improve their patient safety skills. To answer this main question the game AMS-1 and an E-module that was created for this purpose were studied for their effectiveness.

This study is split up into three parts: (1) literature study, (2) effectiveness study and (3) evaluation. Each part consists of multiple sub questions. These sub questions will be answered one by one in this section that will ultimately lead to the answer to the main question. The discussion will be given on the main part of this research: (2) effectiveness study with a discussion More elaborate conclusions with their discussion can be found in 2.3, 3.3.4 and 4.3.

1. **Literature study**

Sub-question 1.1: *Patient safety, why is this such an important subject?*
The first report used for this study was To Err Human (1999), stating that healthcare in the United States is not safe. Staggering numbers were given such as that between 44k-98k people die each year due to preventable medical errors. In the literature review we found that this is not just a problem in the United states, but worldwide. In the EU between 8-12% of patients admitted to hospitals suffer from adverse events.

Sub-question 1.2: *Patient safety in the Netherlands, what is the current situation?*
Zooming in, the Netherlands showed horrific numbers in 2004. The Netherlands took action from that point on with an ambitious plan to reduce preventable adverse events with 50% from 2008 until 2013. This goal was reached with 52% in 2012. A very positive sign, but in absolute numbers there are still around a 1000 preventable deaths a year. There is always room for further improvement and there needs to be continuously development in order to anticipate new risks.

Sub-question 1.3: *Patient education, what is the current situation in patient safety education into the medical school curriculum in different countries.*

This continuously development is also directed towards patient safety education. Luckily, medical students find this subject very important. For instance a German report in 2014 showed that 64% of 167 medical students would want more education on patient safety. This positive fact was also found in our research; out of the 62 5th year medical students 58% strongly agrees and 39% agrees on the question whether or not it is important for medical students to learn about patient safety topics. Literature shows that curriculums for medical students mainly focus on three major competencies (medical knowledge, technical skills and judgment), but the non-technical skills and professional competencies which are important for patient safety seem to be lacking. As of yet, there are relatively few countries and medical colleges that have followed the trend by implementing patient safety education. A systematic literature review shows that patient safety education in the medical school curriculum is most commonly implemented in medical schools in developed countries such as the United States of America and the United Kingdom. These curricula mainly use lectures and small-group instruction and educate on different topics such as analysis of medical errors and reporting

adverse events. A lot of research is being done on different patient safety education topics to see whether or not it improves patient safety skills. These researches show positives effects, but are yet to be fully implemented in curriculums. With the introduction of the WHO patient safety curriculum guide we hope this will change.

Sub-question 1.4: *Is there empirical evidence about the positive impacts and outcomes of computer games ad serious games with respect to learning in the medical domain?*

In our research we are trying to see if the method serious gaming, with the use of Air Medic Sky-1, can be used as a patient safety curriculum or as part of a patient safety curriculum. Research shows that students are positive towards using video games for medical education. This is also shown in our research where 87% replied that they like the idea of using a serious game to learn more about patient safety. A systematic literature from 2011 shows that there is a paucity of studies showing conclusive clinical benefit of serious gaming. This being said, studies show an increase in skill with laparoscopic virtual reality and better decision-making with the use of serious gaming. No other serious game were found with the focus on patient safety, which means our study is a first.

2. **Effectiveness**

Sub-question 2.1: *Do students show enhanced knowledge in patient safety related issues during clinical training when they have played the AMS-1 game, compared to students who have done the e-module?*

The mean score of the AMS-1 group was the highest, followed by the E-module. The control group had the lowest mean. Analysis showed that AMS-1 and E-module had a significant better score than the control group, but there was no significant difference between AMS-1 and the E-module. Thus the conclusion is that medical students who use the E-module or AMS-1 will have improved patient safety knowledge.

The average time spend was three hours on AMS-1 and one hour on the e-module. When one learns (AMS-1 or E-module) for a test, it is expected that he or she does better than the one (control group) that does not learn. The knowledge in AMS-1 was transcript for the E-module, thus the knowledge in the E-module was the exact same as that in AMS-1. Therefore it was also expected that there will be no difference between these two groups when both of them studies the material long enough.

The knowledge test was completed with the use of Blackboard. All 126 multiple choice questions were put in. After all participants filled in the test, results were checked. This is where it was discovered that Blackboard does not give share points. For instance, with the question *''of the twelve items are the seven red flags''*, you could only get one point if you got all seven right. It was concluded that this is too strict. All knowledge tests were then checked for the questions where you could get multiple points and these were allocated. Although this was done thoroughly, the problem here is that this was done by hand, which involves human error. There was no other solution.

Sub-question 2.2: *Do students show enhanced self-rated skill in patient safety related issues during clinical training when they have played the AMS-1 game, compared to students who have done the e-module?*

The score of the control and E-module group was the same and the AMS-1 group had the lowest score. The difference in score was not significant, so it can be concluded that medical students who use the E-module or AMS-1 do not improve their self-reported skill related to patient safety topics.

We expected to see the same improvements in the self-reported skill as we did at the knowledge scores. There are a few possible explanations for this: firstly, the time they had to play AMS-1 and read the E-module could be a factor, meaning that the same analysis has been done for participants score who played at least 2.5hours or read the E-module at least 45minuts. This gave the same result; no significant difference. Also no correlation between time played and score was found.
Secondly there is the measurement bias: using self-report is often biased by social desirability. Although participants were told that their scores were anonymous, some might felt pressure because they should be able to have these skills before they go into their clerkship or they were still afraid a supervisor would read their score. Another possibilities for bias is that people tent to rate themselves better when they have no knowledge about the subject, which would greatly affect the control group. For example, in a study comparing self-reported data of body height and weight with direct measured data, it was found that subjects tend to over-report their height, but under-report their weight in their own benefit. This would result in the control group rating themselves relatively high, because of their lack in experience in this field. The E-module and AMS-1 students would rate themselves relatively low, because they have seen how difficult the learning material can be, putting the scores of the control and intervention groups close together.
Thirdly, the game is not yet ready to improve self-reported skill. If this would be the case, the game needs to be further developed for this purpose.

> **Sub-question 2.3**: *Do students show enhanced <u>self-reported behavior</u> in <u>Stress</u> and <u>Awareness</u> in patient safety related issues during clinical training when they have played the AMS-1 game, compared to students who have done the e-module?*

The self-reported behavior test is split into two subject. The first one is Stress where three questions from the validated Perceived Stress Scale were asked, measuring their stress level on a scale of 1 to 5 during clerkship. The second subject is Awareness, where we use two questions to see whether or not their behavior towards patient safety has changed; do the participants see/experience more events when they have undergone the intervention. For both subjects, no significant differences were found. We can conclude that medical students who use the E-module or AMS-1 do not have a reduced stress level and their awareness of adverse events does not change at their clerkship.

AMS-1 consist out of three parts. The game goes over Stress management in the first part of the game (breathing exercises) and partly in the third part of the game (the missions). In the E-module, one chapter was available on stress management, but it was not as elaborate as the stress management in AMS-1. We therefore expected that AMS-1 would score significantly better on the stress test than the E-module and control group, but it did not.
During the evaluation interviews participants told us that their stress is low during their first clerkship, because they do not have a lot of responsibilities and they can always fall back on their supervisor. This is correct when we look at the stress level of the control group who had an average stress level of 2.22 where a score of 1 means *Never stressed* and 5 *Stressed very often* (5). One could wonder if it is even possible to lower a stress level significantly when stress is already fairly low. Originally the game is meant for AOIS, which are doctors in

specialist training. These are doctors who have a lot of responsibility and here it is a well-known fact that stress is a big issue. It is Interesting to see whether or not this study would have another outcome for AOIS. Another possibility is that in AMS-1 stress management needs some changes. As well in the interviews as in the evaluation questionnaire, it became clear that the breathing exercises do not work as intended most of the time. This was blocking the ability to learn from these exercises. Also the stress at the missions was not always high. Reasons given were that even though more patients were treated at the same time, it was still a game (not realistic enough to get stressful) and that nobody was watching them as they played.
A flaw in the research design might have caused the results on Awareness. We will take a better look at this in the *reflection of research approach* part. In summary, in phase 1 (control group data was collected), students' attention in patient safety education was more guided towards patient safety and systematic analysis of errors; whereas in phase 2 (intervention groups data was collected) the focus was more towards general open and clear communication. We therefore think that the control group was more sensitive to recognizing adverse events during their following clerkship than the intervention groups and this might have caused insignificant results. For Stress and Awareness, the time spent in the intervention groups with the material is also a factor. This is why we did the same analysis for students who played AMS-1 for at least an hour and read the E-module at least 45minutes; but this had no effect. There was also no correlation between time played and score.

3. Evaluation

The evaluation, gives the ability to look further than the statistical analysis, giving some clarification on the effectiveness research which has been done above. The full conclusion of the evaluation can be found in section 4.3.

> **Sub-question 3.1**: *Do they evaluate the game better than the E-module and in what respect?*

This sub question is answered by the user questionnaire containing twenty-five questions divided over six items. At these six items, the game AMS-1 scored significantly better on engagement, user satisfaction and attitude towards using. In more detail this shows that the opinion of the AMS-1 group goes significantly towards:

- <u>Engagement</u>: making them have more attention, feel more actively involved and being more concentrated during the intervention.
- <u>Satisfaction recommending</u>: like AMS-1 as a learning tool, because it is more fun and pleasant to use.
- <u>Attitude towards using</u>: liking the idea of using a game (AMS-1) to learn about patient safety, overall enjoyed using it and would have like to continued using it.

The other items were all rated very positively for both AMS-1 and the E-module. This shows that the opinion of both groups significantly goes towards:

- <u>Perceived Usefulness</u>: helps being more effective at patient safety and are useful for educational purposes for patient safety.
- <u>Motivation</u>: finding it important to learn about patient safety, like what they are learning and can use what they learn in future work as a doctor.
- <u>Ease of use</u>: easy to use, not full of problems to operate, easily remember how to use it and come quickly skillful using it.

Sub-question 3.2: *What aspects of the game do students appreciate and dislike?*

This sub question is mainly answered by the interviews and partly by the user questionnaire.

Almost all participants started saying how much fun the game was to actually play and that they to continue playing the game because of the fun they had. Most of this fun came from the mission; it was educational, being able to make decisions, realistic, and generating stress. One part was explicitly mentioned in every interview: the red flag game. This was a single part of the game where a lecture, about the nine red flags, was put in a game format. Overall opinion was that this was a good example of how you should learn more about patient safety in a game format instead of lectures where you are only listening. It was mentioned that they would use what they had learned from the red flag game. The score system was the most motivating game aspect of AMS-1as mentioned in the interviews. It kept one participant motivated enough to watch all lectures. The storyline was the second best mentioned game aspect, but no specific reason was mentioned.

The breathing exercises were most frequently mentioned as a negative aspect of the game. Reasons being: susceptible to distraction, prone to errors (sweaty hands and not able to reduce heart rate), realism issues (are you really able to take a step back in clerkship and do a breathing exercise) and taking too long to complete.
The lectures were second most mentioned. Reason being that the lectures were boring, hard to keep attention, low engagement, English not always at its best, . Because of this, not a lot of lectures were watched by these participants.
Even though that 'speak up' (red flag game, say something when you think something is going wrong) was mentioned as a positive aspects, it became clear that participants were divided whether they would use it in their clerkship. Reasons being that students are not confident enough yet in their own knowledge and find it hard to criticize someone who will give them a grade. Opinions were divided if the game could really reduce stress. On the one hand the biofeedback exercises have potential to learn how to deal with stress. On the other hand participants did not find the missions realistic enough to generate stress. One person said there was not enough pressure, even when he was dealing with multiple patients, simply because you know it is a game, nobody is watching and you can fail without consequences.

MAIN research question: *How effective is serious gaming, used as a training tool, at improving medical student patient safety skills?*

To answer this question the game AMS-1 and an E-module that was created for this purpose were studied for their effectiveness in the following items: (1) knowledge, (2) self-reported skill and (3) self-reported behavior in stress and awareness on patient safety related issues. Analysis on these subjects show an significant increase in knowledge, but not in self-reported skill or in self-reported behavior. The problem is that increasing self-reported skill -and behavior is the major part of the game. Therefore it is very questionable that AMS-1 should be used as a training tool for the fifth year medical students that are about to go into their first clerkship. Also since a simpler and cheaper learning method (the E-module) had the same effect on the increase in patient safety knowledge. The probable cause for AMS-1 increasing medical knowledge is that students find the game engaging, are satisfied with the game, perceived usefulness of the subject patient safety is high, are motivated to learn about patient safety and learn with a serious game and find it easy it use. But most importantly it was fun and educational at the same time. The probable cause for the intervention groups not having

an effect on self-rated skill –and behavior are: wrong target group, bug problems with one specific part of the game, measurement bias, flaw in the research design and that the lectures in AMS-1 that had no gamification. This study does not deny that AMS-1 can work effectively for its intended purpose, but rather that this is not the target group for this game and the game needs improvements on some parts of the game, for instance a better training mechanism for the breathing exercises and more gamification in the lectures.

6 Limitations

When reflecting on this study, two things could have been improved. The first one is the randomization of the groups and the second one is the research population.

Randomization
In this study we had three groups; the control, E-module and AMS-1 group. It was considered important that these groups did not influence each other. For example, if the a participant from the AMS-1 let someone from the control group play the game then the data from this participant would not be valid for the control group. Therefore the control group was split from the intervention groups. This was done by first collecting the data of the control group, having a break of 10 weeks, and then randomizing those who would get one of the interventions. The problem was that this was not done in the same school period; school periods change in time. In this case the curriculum of patient safety that all participants got from their education, one week before their clerkship, changed:

1. Phase 1 (from April until September '13):
Human Factors approach of patient safety (recognize situations which enhance chance of errors). Instruction on systematic analyses of human errors, using the PRISMA approach (Prevention and Recovery Information System for Monitoring and Analysis). After a number of the clerkships, students have to describe an adverse event which they have encountered during their clerkship, using the PRISMA approach.

2. Phase 2 (from September '13-now)
Focus on effective and open communication in different clinical situations, make assertive statements, reflection on communication patterns. After the first clerkship, students have to discuss their own experiences in the clinic and analyze these with each other. Themes discussed by students are communication and patient safety.

In summary, in phase 1 (control group data was collected), students' attention in patient safety education was more guided towards patient safety and systematic analysis of errors; whereas in phase 2 (intervention groups data was collected) the focus was more on general open and clear communication. We would expect the control group to have been more sensitive to recognizing adverse events during their following clerkship than the intervention groups. What should have been improved is that all groups should be randomized so that the data collection was in one time period. Students would then get asked in their evaluation if they swapped learning material with each other. If a student answered yes, the data would not have been used for the research.

Research population
The research population consists of fifth year medical students at the Erasmus University Medical Center, doing their preparatory education and 10-week Internal Medicine Clerkship (first clerkship). In general, a group of 13 students start their clerkships internal medicine at Erasmus MC or regional hospitals every 2 weeks. There were three problems with this population, first of which was that we could only ask 13 students every 2 weeks, so it took a year to collect the data which was not very efficient. This also had the effect that the preparatory education changed over time as mentioned earlier in this section. Second, a big part of AMS-1 was designed for stress management and in the effectiveness research we looked at how effectives this stress management was. The problem we found was that this group does not have a lot of stress, due to low responsibility they have in there clerkship. It is

hard to test how effective a stress management tool is, when the target group almost has no stress. An alternative target group would be AOIS (specialist in training). AOIS have a lot of responsibility and are well known for their high level of stress. Problem with AOIS is that it would be hard to convince them to play AMS-1 due to stress and the scarce free time they have at their disposal. Third and last, all participants agreed that they had too little time to play. This was due to the busy week they had which caused few possibilities to play in that week. All would have liked to continue playing the game. Another week should have been taken for the intervention groups so that the participants would have been able to undergo the intervention longer.

7 Future Research

To date, a lot of research has been done on the impact that patient safety has on the medical domain. Also many studies can be found on the different patient safety education topics, but our literature study shows that not a lot research has been done on the success of implemented patient safety education curriculums for undergraduates. In this study we showed a way to measure the success of a training tool to improve patient safety skills of medical students. This could be copied to measure the effectiveness of a curriculum.

It was hard to test how effective the stress management tool was, reason being that the target group almost had no stress. An alternative target group would be AOIS (specialist in training). AOIS have a lot of responsibility and are well known for their high level of stress. We recommend that this research is done again with this target group, focusing on the stress management.

References

[0] Van de Weerd, I., & Brinkkemper, S. Meta-modeling for situational analysis and design methods. In Handbook of Research on Modern Systems Analysis and Design Technologies and Applications (pp. 28–38).

[1] Institute of medicine. To Err Is Human: Building a Safer Health System. 1999.

[2] America, C. o., & Medicine, I. o. (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century.* America.

[3] Organization, W. H. (2013). *Patient-Saftey*. Retrieved on 03 18, 2013, van World Health Organization: http://www.euro.who.int/en/what-we-do/health-topics/Health-systems/patient-safety

[4] [15] Bruijne, M. C. D., Zegers, M., Hoonhout, L. H. F., and Wagner, C. Onbedoelde schade in nederlandse ziekenhuizen. dossieronderzoek van ziekenhuisopnames in 2004, 2004.

[5] Nie, Y., Li, L., Duan, Y., Chen, P., Barraclough, B. H., Zhang, M., et al. (2011). *Patient safety education for undergraduate medical students: a systematic review.* Sichuan Province, China: Pubmed.

[6]. Hall LW, Scott SD, Cox KR, Gosbee JW, Boshard BJ, Moylan K, et al. Effectiveness of patient safety training in equipping medical students to recognise safety hazards and propose robust interventions. Quality & safety in health care. 2010 Feb;19(1):3–8.

[7]. Leung GKK, Patil NG, Ip MSM. Introducing patient safety to undergraduate medical students--a pilot program delivered by health care administrators. Medical teacher. 2010 Jan;32(12):e547–51.

[8]. Paxton JH, Rubinfeld IS. Medical errors education: A prospective study of a new educational tool. American journal of medical qualityR: the official journal of the American College of Medical Quality [Internet]. 2010 [cited 2012 Nov 10];25(2):135– Available from: http://www.ncbi.nlm.nih.gov/pubmed/20110456

[9]. Halbach JL, Sullivan LL. Errors and Patient SafetyR: Evaluation of a Required Curriculum. 2005;80(6):600–6.

[10]. Moskowitz E, Veloski JJ, Fields SK, Nash DB. Development and evaluation of a 1-day interclerkship program for medical students on medical errors and patient safety. American journal of medical qualityR: the official journal of the American College of Medical Quality. 2007;22(1):13–7.

[11]. Patey R, Flin R, Cuthbertson BH, MacDonald L, Mearns K, Cleland J, et al. Patient safety: helping medical students understand error in healthcare. Quality & safety in health care. 2007 Aug;16(4):256–9.

[12]. Madigosky WS, Headrick L a, Nelson K, Cox KR, Anderson T. Changing and sustaining medical students' knowledge, skills, and attitudes about patient safety and medical fallibility. Academic Medicine. 2006 Jan;81(1):94–101.

[13] Zegers, M., Bruine, M. C., Wagner, C., Groenewegen, P. P., Waaijman, R., and Wal, G. V. D. Design of a retrospective patient record study on the occurrence of adverse events among patients in dutch hospitals. BMC Health Services Research 7, 27 (2007).

[14] Ogranisation, W. H. (2013). *curriculum*. Retrieved on 03 19, 2013, van World Health Ogranisation: http://www.who.int/patientsafety/education/curriculum/en/index.html

[15] Leape LL. Reporting of advers events. New England Journal of Medicine. 2002;347(20):1633–8.

[16] Special Eurobarometer 327. Patient safety and quality of healthcare. 2010.

[17] Lockwood JH, Sabharwal RK, Danoff D, Whitcomb ME. Quality improvement in medical students ' educationR: the AAMC medical school graduation questionnaire. 2004;234–6.

[18] Aubin D, King S, Boechler P, Burden M, Rockwell G, Henry M, et al. Serious games for patient safety education. Medical Teacher. 2012;(early Online):675.

[19] Kalkman CJ. Serious Play in the Virtual WorldR: Can We Use Games to Train Young DoctorsR? Journal of Graduate Medical Education. 2012;4(1):11–3.
[20] *international-crew-leaders*. (2011). Retrieved on 03 19, 2013, van airmedicsky1: http://www.airmedicsky1.org/international-crew-leaders.php
[21] Consortium, N. M. (2012). *Horizon report.* Horizon report, 2012: New Media Consortium.
[22] WHO, "Definition patient safety by WHO," 2014. [Online]. Available: http://www.euro.who.int/en/health-topics/Health-systems/patient-safety/patient-safety. [Accessed: 23-Jul-2014].
[23] The Institute of Medicine, "To err is human: building a safer healthsystem - Brief," 1999. .
[24] M. Zegers, M. C. de Bruijne, C. Wagner, P. P. Groenewegen, R. Waaijman, and G. van der Wal, "Design of a retrospective patient record study on the occurrence of adverse events among patients in Dutch hospitals.," *BMC Health Serv. Res.*, vol. 7, p. 27, Jan. 2007.
[25] IoM, "Institude of health website," 2014. [Online]. Available: http://www.iom.edu/About-IOM.aspx. [Accessed: 23-Jul-2014].
[26] Q. Chasm, C. Committee, H. Care, M. Isbn, T. Pdf, N. A. Press, and N. Academy, *Crossing the Quality Chasm A New Health System for the 21st Century*. 2001.
[27] C. Of, T. H. E. E. Union, L. Acts, and O. Instruments, "European commision public health- patient safety 2009," vol. 2009, no. June, 2009.
[28] S. Eurobarometer, "Patient safety and quality of healthcare Full report," no. October 2009, 2010.
[29] L. H. F. Hoonhout, *Onbedoelde schade in Nederlandse ziekenhuizen Dossieronderzoek van ziekenhuisopnames in 2004*. 2004.
[30] M. Langelaan, R. J. Baines, M. A. Broekens, K. M. Siemerink, H. Asscheman, M. C. De Bruijne, and C. Wagner, *Monitor Zorggerelateerde Schade 2008 Dossieronderzoek in Nederlandse ziekenhuizen*. .
[31] C. Blok, C. de, Koster, E., Schilp, J., Wagner, "VMS Veiligheidsprogramma 2008-2013."
[32] NIVEL, "Nederlandse-ziekenhuizen-zijn-veiliger-geworden," *18-11-2013*, 2013. [Online]. Available: http://www.nivel.nl/nieuws/nederlandse-ziekenhuizen-zijn-veiliger-geworden. [Accessed: 28-Jul-2014].
[33] M. Langelaan, M. C. De Bruijne, R. J. Baines, M. A. Broekens, K. Hammink, J. Schilp, L. Verweij, H. Asscheman, and C. Wagner, "Monitor Zorggerelateerde Schade 2011 / 2012."
[34] J. Kiesewetter, M. Kager, R. Lux, B. Zwissler, M. R. Fischer, and I. Dietz, "German undergraduate medical students' attitudes and needs regarding medical errors and patient safety - a national survey in Germany.," *Med. Teach.*, vol. 36, no. 6, pp. 505–10, Jun. 2014.
[35] J. Tingle, "The WHO Patient Safety Curriculum Guide."
[36] WHO, "Patient Safety Curriculum Guide Regional Launches," *22-11-2013*, 2013. [Online]. Available: http://www.who.int/patientsafety/education/curriculum/regional_launch/en/. [Accessed: 31-Jul-2014].
[37] E. Conroy-Smith, R. Herring, and G. Caldwell, "Learning safe prescribing during post-take ward rounds.," Jul. 2011.
[38] I. Shekhter, L. Rosen, J. Sanko, R. Everett-Thomas, M. Fitzpatrick, and D. Birnbach, "A patient safety course for preclinical medical students.," *Clin. Teach.*, vol. 9, no. 6, pp. 376–81, Dec. 2012.
[39] M. Gordon, "Non-technical skills training to enhance patient safety.," Jul. 2013.
[40] E. Alper, E. I. Rosenberg, K. E. O'Brien, M. Fischer, and S. J. Durning, "Patient safety education at U.S. and Canadian medical schools: results from the 2006 Clerkship Directors in Internal Medicine survey.," Dec. 2009.
[41] Y. Nie, L. Li, Y. Duan, P. Chen, B. H. Barraclough, M. Zhang, and J. Li, "Patient safety education for undergraduate medical students: a systematic review.," *BMC Med. Educ.*, vol. 11, no. 1, p. 33, Jan. 2011.
[42] B. Stokes, "Videogames have changed: time to consider 'Serious Games'?," 2005.
[43] F. W. Kron, C. L. Gjerde, A. Sen, and M. D. Fetters, "Medical student attitudes toward video games and related new media technologies in medical education.," *BMC Med. Educ.*, vol. 10, p. 50, Jan. 2010.
[44] K. Squire and H. Jenkins, "Harnessing the power of games in education," vol. 3, pp. 5–33, 2003.
[45] J. Kirriemuir, "Literature Review in Games and Learning Literature Review in Games and Learning," 2004.
[46] L. D. de Wit-Zuurendonk and S. G. Oei, "Serious gaming in women's health care.," *BJOG*, vol. 118 Suppl , pp. 17–21, Nov. 2011.

[47] S. J. Brown, D. a Lieberman, B. a Germeny, Y. C. Fan, D. M. Wilson, and D. J. Pasta, "Educational video game for juvenile diabetes: results of a controlled trial.," *Med. Inform. (Lond).*, vol. 22, no. 1, pp. 77–89, 1997.

[48] T. Susi, M. Johannesson, and P. Backlund, "Serious Games – An Overview," 2007.

[49] Kirkpatrick, D.I. (1998) Evaluating Training Programs: The Four Levels, 2nd edn (San Francisco, Berrett-Koehler).

[50] *self-efficacy*. (sd). Retrieved on 04 04, 2013, van medical-dictionary: http://medical-dictionary.thefreedictionary.com/self-efficacy

[51] Cohen, S., Kamarck, T., & Mermelstein, R. (1983). Perceived Stress Scale. Journal of Health and Social Behavior, 24, 385–396.

[52] Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface *Interface Evaluations* / Chin, John P. / Diehl, Virginia A. /Norman, Kent L. *Proceedings of ACM CHI'88 Conference on Human Factors in Computing Systems* 1988-05-15 p.213-218

[53] Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology / Davis, F. D. MIS Quarterly 1989 v.13 n.3 p.319-340

[54] Nielsen, J (1993) Usability Engineering, Boston: Academic Press

[55] IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use / Lewis, James R. International Journal of Human-Computer Interaction 1995 v.7 n.1 p.57-78

[56] Proceeding CHI EA '97 CHI '97 Extended Abstracts on Human Factors in Computing Systems Pages 168-169

[57] IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use / Lewis, James R. International Journal of Human-Computer Interaction 1995 v.7 n.1 p.57-78

[58] A Proposed Index of Usability: A Method for Comparing the Relative Usability of Different Software Systems *Usability Evaluation Methods* / Lin, Han X. / Choong, Yee-Yin / Salvendy, Gavriel Behaviour and Information Technology 1997 v.16 n.4/5 p.267-278

[59] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319-340. http://www.jstor.org/pss/249008

[60] SUMI: the Software Usability Measurement Inventory, Jurek Kirakowski and Mary Corbett, Article first published online: 27 OCT 2006

[61] Lund, A.M. (2001) Measuring Usability with the USE Questionnaire. STC Usability SIG Newsletter, 8:2.

[62] Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq)Educational and Psychological Measurement AUTUMN 1993 53: 801-813,

[64] Clifford N and Valteine G eds 2003 Key methods in geography Sage, London

Appendix

Appendix A: Topics in the WHO patient safety curriculum guide

1 What is patient safety?
2 What is human factors engineering, and why is it important to patient safety?
3 Understanding systems and the impact of complexity on patient care?
4 Being an effective team player
5 Understanding and learning from errors
6 Understanding and managing clinical risks
7 Introduction to quality improvement methods
8 Engaging with patients and carers
9 Minimising infection through improved infection control
10 Patient safety and invasive procedures
11 Improving medication safety

Appendix B: Knowledge, knowledge patient safety questionnaire

Communication
Pre-Briefing

Do the following components belong to a complete patient safety pre-briefing?

1.	A short description of the situation	Yes / **no**
2.	A summary	Yes / **no**
3.	Acknowledgement of all communication	**Yes** / no
4.	An invitation for participation	**Yes** / no
5.	The mechanism of injury	Yes / **no**
6.	Introducing yourself	**Yes** / no
7.	A motivational message	Yes / **no**
8.	Questions to ensure understanding	**Yes** / no
9.	A description of the task	Yes / **no**
10.	A checklist consisting the expected outcomes	**Yes** / no

11. Pre-briefing is particularly important before
    a. surgical procedure
    b. administration of medication
    c. any activity involving teamwork
    d. handing over a patient

12. A team huddle is:
    a. a group debriefing
    b. a pre-briefing
    c. a strategy meeting during action
    d. a post-hoc meeting to reconcile a dispute

Do the following skills belong to the 'showing good interpersonal skills' component of a pre-briefing?

13. Showing empathy Yes / **no**
14. Introducing yourself **Yes** / no
15. Making eye-contact **Yes** / no
16. Energizing the team Yes / **no**
17. Acknowledging all communications Yes / **no**
18. Getting to know all team members names **Yes** / no
19. Using adequate intonation while communicating Yes / **no**

20. Which item is not important during the communication of a message?
    a. Body language
    b. **The order of words**
    c. Word choice
    d. Pronunciation of words

Assertive Statement
'Dr Jones, don't you think we should treat this patient for his high bloodpressure?' is an example of an incomplete "effective assertive statement". Check which components are included, which are missing and which are not essential for an effective assertive statement.

21. Get the attention **included** / missing / not essential
22. Show your concern included / **missing** / not essential
23. Ask for help included / missing / **not essential**
24. State the problem included / **missing** / not essential
25. Wait for a solution included / missing / **not essential**
26. Propose a solution **included** / missing / not essential
27. Wait until the problem has been solved included / missing / **not essential**
28. Wait until a decision has been made included / **missing** / not essential

Call Back Read Back

29. A specialist asks you to give one 1ml ampulla morphine 10mg/ml. You reply: 'here is one 1ml ampulla morphine 10mg/ml' and you hand over the medication. This is an example of:
    a. 'Pre-briefing'
    b. 'Call back-read back' principle
    c. 'Debriefing'
    d. 'Hands off eyes on' principle
    e. The 'comprehension check'

30. The 'call back read back' principle teaches to always check the results of your previous ordered diagnostic tests.
    a. True
    b. False

31. The ambulance brings in an ill patient while you are on duty in the Emergency Room as a medical officer. You read the admission papers of the patient, which have been handed over by the ambulance nurse.
This is an example of the 'read-back' part of the 'call back-read back' principle.
    a. True

b. False

32. In a community hospital the statement 'Communication that goes unacknowledged did not happen' is:
    a. True
    b. False

Handovers

33. One appropriate way of conducting a handover is debriefing.
This statement is:
    a. True
    b. False

34. Part of a handover procedure is the 'call back-read back' action.
This statement is:
    a. True
    b. F**alse**

Do the following components belong to a complete handover?

35. A notation **Yes /** no
36. A summary Yes / **no**
37. A debriefing Yes / **no**
38. Information provided according an agreed structure **Yes /** no
39. Supplying a written summary Yes / **no**
40. Providing a differential diagnosis Yes / **no**
41. A dialogue **Yes /** no
42. A therapeutic plan Yes / **no**

43. The 'Hands off-eyes on principle' applies to
    a. Electric cardio version
    b. Handovers
    c. Pre-briefing
    d. Handling conflicts
    e. Call back Read back

44. The 'Comprehension check' should be carried out after:
    a. Call back Read back
    b. Hands off-eyes on principle
    c. Debriefing
    d. Handovers

45. Which of the following steps belong to a complete debriefing and put them in the right order:
    A. Checking for all unanswered questions/unaddressed problems
    B. Accurate description of what happened
    C. Things that need improvement
    D. Specific action for changing or improving next performance
    E. Thorough description of why we did what we chose to do

F. Things that went well

Correct order: B, E, D

46. A synonym for debriefing is a performance feedback session
    a. True
    b. False

47. Debriefing is important because:
    a. it makes patient handover safer
    b. it clarifies what all team members tasks are
    c. it is a check for unanswered questions
    d. its effect is continuous quality improvement

48. The goal of a debriefing is:
    a. to optimise the final preparation for a task
    b. to optimise the performance for the next task
    c. to make sure the task is completed totally

Focus under Stress

Focus on your Patient

Focus on the Now-through centred breathing

49. "Centred breathing" is a
    a. relax aimed breathing technique
    b. 5-step breathing technique
    c. focus directed breathing technique
    d. mid-abdominal breathing technique

Teamwork

Shared mental model

Which of the following statements about 'the shared mental model' are true and which are false? The shared mental model:

50. helps to detect possible mistakes of teammates
    a. True
    b. False
51. is based on clear communication between teammates
    a. True
    b. False
52. aims at intuitively knowing what teammates are doing
    a. True
    b. **False**
53. prevents depression in team mates
    a. True
    b. **False**
54. shows the distribution of stress applied to a team by external pressure
    a. True
    b. False
55. is a standard approach that every teammate knows about

a. True
b. False

56. spreads the effects of sleep deprivation among team mates
    a. True
    b. False

Teamwork: Ken Catchpole –teamwork to control stress level
Using Resources
Which of the following sentences are true and which are false?
The use of resources in dealing with patient safety issues is about:

57. reducing the stress of one team member to other members
    a. True
    b. False
58. knowing and using your equipment
    a. True
    b. False
59. using the knowledge and skills of other team members
    a. True
    b. False
60. asking a colleague for help
    a. True
    b. False

Sleep Deprivation
Signs of Sleep Deprivation

Which of the following statements are true and which are not true?

61. Sleep deprivation is associated with a significant increase in medical mistakes.
    a. True
    b. False

62. It is easier to recognize signs of sleep deprivation in yourself than in team mates
    a. True
    b. False

63. The 'call back-read back' principle is one of the best methods to prevent mistakes due to sleep deprivation.
    a. True
    b. False

64. Sleep inertia is the phenomenon that you feel really tired the first couple of minutes after a short nap
    a. True
    b. False

65. The best method to deal with sleep inertia is to adapt to the local day-night pattern
    a. True

b. False

66. To prevent sleep deprivation it is best to take an hour more sleep every day for 4-7 days.
    a. True
    b. False

67. The recommended time for a short nap is 10 to 15 minutes
    a. True
    b. False

68. Short naps do not make one more alert in the first half hour later
    a. True
    b. False

69. Self-awareness is impaired by sleep depriviation
    a. True
    b. False

70. When people ignore their feelings of tiredness, they can be able to get over their sleepy feelings and function normally.
    a. True
    b. False

71. Self-awareness is more related to self-confidence than to unsecurity.
    a. True
    b. **False**

Depression
Signs of Depression

72. Physicians are better skilled in recognising signs of depression in themselves than non-physicians.
    a. True
    b. False

Do the following statements belong to a sign of depression?

73	Insomnia or hypersomnia	**Yes /** no
74.	Feeling hopeless	**Yes /** no
75.	Having a flight of ideas	Yes / **no**
76.	Changes in weight	**Yes /** no
77.	Feelings of hurting yourself or others	**Yes /** no
78.	Feeling more sad in the evening	Yes / **no**

79. Working as a resident increases the risk of depression.
This statement is:
    a. True
    b. False

80. When you think a colleague is depressed, you should talk to a person of authority about it.
This statement is:

a. True
b. False

81. You should ask friends and family to pay attention to possible signs of depression.
This statement is:
a. True
b. False

Depression in Teammates

Depression & Drug Abuse
82. If you recognise signs of alcohol or drug abuse in fellow residents, you should talk with them about your suspicion.
a. true, because it is so easy for physicians to obtain narcotic drugs
b. true, because it is hard to recognise in oneself
c. false, because this only feeds feelings of distrust among residents
d. false, because this is a person's own responsibility

83. Alcohol or drug abuse can mask a depression; this is its major threat to patient safety.
This statement is:
a. True
b. False

84. Working as a resident increases the risk of alcohol or drug abuse
This statement is:
a. True
b. False

Making Mistakes when depressed
85. Residents showing evident signs of depression make about 10 times more errors then colleagues without such signs
a. this number is substantially higher
b. this number is substantially lower, but still higher than in non-depressed colleagues
c. it has not been documented that they make more mistakes

86. The 4 most probable reasons for making more mistakes while being depressed are:
a. a slower reaction time
b. feelings of depression
c. a lack of motivation
d. making less eye-contact in communication
e. being very tired the first 10-15 minutes after a nap
f. increased forgetfulness
g. feelings of disassociation
h. thoughts of hurting oneself or other people

Patient Safety
Introduction to Patient Safety

87. Patients admitted to hospital encounter unintended harm due to adverse events during their stay. The correct percentage is:

a. 1-3%
b. 3-6%
c. 6-12%
d. 12-20%
e. 20-30%

88. Unintended adverse events in patients during hospitalization are a more common cause of death in the USA than car accidents.
This statement is:

a. **True**
b. False

89. The majority of unintended adverse events in patients during hospitalization is due to:

a. a lack of motivation in the medical staff to observe patient safety rules
b. a lack of knowledge and/or skills medical staff about patient safety
c. system problems of the hospital in dealing with patient safety

90. The Systems Engineering Initiative for Patient Safety (SEIPS) work system model is an excellent tool to evaluate a working environment.
This statement is:

a. True
b. False

Which item(s) belong(s) to an element of the SEIPS work system model?

91. The tools and technology **Yes** / no
92. The working team Yes / **no**
93. The individual **Yes /** no
94. The training and education Yes / **no**
95. The task **Yes /** no
96. The organisation **Yes /** no

The 4 challenges of Patient Safety
Do the following statements belong to the challenges of patient safety?

97. High level of ambiguity **Yes /** no
98. High level of carelessness Yes **/ no**
99. Problems are hardly visible **Yes /** no
100. High level of complexity **Yes /** no
101. High level of indifference Yes **/ no**
102. Conflicting interests Yes **/ no**
103. The problem is linked to several side-effects of professionalism **Yes /** no

104. The high level of complexity is one of major challenges in:

a. Patient safety matters
b. Sleep deprivation
c. Identifying red flags in teamwork

d. Providing a good handover

105. A high level of ambiguity can mean:
a. The discussion of system failure versus individual failure
b. Hierarchical pressure
c. Peer pressure
d. Difficulties asking colleagues

9 Red Flags
Which are the nine red flags and which are not?

106.	Conflicting inputs	**Yes /** no
107.	Ambiguity	Yes / **no**
108.	Preoccupation	**Yes /** no
109.	Lazyness	Yes / **no**
110.	Imprudence	Yes / **no**
111.	Complex situations	Yes / **no**
112.	Not communicating	**Yes /** no
113.	Confusion	**Yes /** no
114.	Rudeness	Yes / **no**
115.	Violating policy or procedure	**Yes /** no
116.	Failure to meet a target	**Yes /** no
117.	Prematurity	Yes / **no**
118.	Inaccuracy	Yes / **no**
119.	Indifference about a situation	Yes / **no**
120.	Not addressing a discrepancy	**Yes /** no
121.	Fatigue	**Yes /** no
122.	Carelessness	Yes / **no**
123.	Indiscretion	Yes / **no**
124.	Stress	**Yes /** no

125. The red flags are warning signs of:
a. Depression
b. Sleep Deprivation
c. Teamwork
d. Patient status

126. A red flag should prompt all attending personal to initiate advanced live support.
This statement is:
a. True
b. **False**

Appendix C: Skill, Self-efficacy questionnaire

100 mm visual analogue scale self-efficacy questions

"Please indicate how you assess your ability to …

1 Physically reduce high levels of stress in yourself
2 Focus on one important task when dealing with multiple things at a time
3 Perform a debriefing within a team, e.g. before an operation
4 Hand over patient information
5 Recognize signals of threats to patient safety during teamwork
6 Conduct a debriefing after a team task
7 Raise issues of threat to safety of a patient among the medical staff
8 Recognize signs of depression in yourself
9 Recognize signs of depression in colleagues
10 Recognize signs of sleep deprivation in yourself
11 Recognize signs of sleep deprivation in colleagues
12 Approach a senior staff member personally for his or her negligence

Appendix D: Attitude, Perceived Stress Scale questionnaire

The questions in this scale ask you about your feelings and thoughts **during the last month**. In each case, you will be asked to indicate by circling *how often* you felt or thought a certain way.

Name _____ Date _____

Age _____ Gender (*Circle*): **M F** Other _____

0 = Never 1 = Almost Never 2 = Sometimes 3 = Fairly Often 4 = Very Often

1. In the last month, how often have you been upset because of something that happened unexpectedly?..... **0 1 2 3 4**
2. In the last month, how often have you felt that you were unable to control the important things in your life?..... **0 1 2 3 4**
3. **In the last month, how often have you felt nervous and "stressed" during Clerkship?** **0 1 2 3 4**
4. In the last month, how often have you felt confident about your ability to handle your personal problems? **0 1 2 3 4**
5. In the last month, how often have you felt that things were going your way? **0 1 2 3 4**
6. **In the last month, how often have you found that you could not cope with all the things that you had to do during clerpship?** **0 1 2 3 4**
7. In the last month, how often have you been able to control irritations in your life **0 1 2 3 4**
8. **In the last month, how often have you felt that you were on top of things during clerkship?** **0 1 2 3 4**
9. In the last month, how often have you been angered because of things that were outside of your control? **0 1 2 3 4**
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them? **0 1 2 3 4**

Appendix E: Manual AMS-1

Handleiding Air Medic Sky-1

De drie onderdelen van AMS-1

AMS-1 bestaat uit drie onderdelen:

1. In het begin van het spel zal je in de 'Biodome' terecht komen. Dit is een gebouw met drie verdiepingen waar je doorheen kunt klikken. Het eerste onderdeel die je in deze biodome tegen zult komen zijn ademhalingsoefeningen. Deze worden gespeeld met de biofeedback vingers. Zorg ervoor dat je deze vingers de gehele game om hebt. Hier krijg je 'self regulation' punten voor.
2. Het tweede onderdeel zul je ook in de 'Biodome' tegenkomen. Hier worden verschillende lectures gegeven over patiënt veiligheidsonderwerpen zoals communicatie, samenwerken, werken onder stress en slaapaanvallen. Hier krijg je 'knowledge points' voor.
3. Wanneer je genoeg 'self regulation' punten en 'knowledge punten' hebt wordt je automatisch uitgenodigd om mee op missie te gaan. Wegens het onderzoek willen wij graag dat je vrijwel alles in de Biodome voltooid, voordat je door klikt naar deze missies, omdat een deel van de onderzoeksvragen die je aanstaande vrijdag krijgt hierover gaan.

Saveknop en logdata

Sla het spel altijd op wanneer je wilt stoppen op de usb-stick! Indien je AMS-1 lang achter elkaar gaat spelen raden wij aan het ieder half uur even op te slaan. Het is van uiterst belang dat je dit doet, omdat wij uit deze savebestanden data halen voor het onderzoek. De 'saveknop' zit een beetje verstopt. Het is te vinden wanneer je linksonder op 'Menu' klikt, vervolgens op 'Exit game' drukt. Hierna zal linksboven 'save game' verschijnen.

Lukken de ademhalingsoefeningen niet?

Het kan voorkomen dat je van naturen een hele rustige hartslag hebt, dit de ademhalingsoefeningen verstoort en zo niet alles kunt halen. Probeer ze in ieder geval wel te doen zodat je weet wat je moest doen. Als de ademhalingsoefeningen niet lukken kan je ook genoeg punten halen met lectures kijken om 'intern skill level' te krijgen en door te gaan naar de missies.

Debriefing bij de missies

Er vindt bij elke patient die je gaat behandeld een debriefing plaats door een zuster. Houdt er rekening mee dat dit net als in het echt werkt. Een zuster zal je niet alles zelf in één keer vertellen. Door meerdere keren op de knoppen te drukken vraag je haar extra informatie te vertellen. Maak hier gebruik van!

Hoe ver ben ik in het spel en tot waar moet ik spelen?

Wanneer je linksonder op 'menu' klikt en vervolgens op 'Your level' kun je zien hoeveel punten je voor alle onderdelen tot nu toe hebt gehaald. In totaal kun je vijf levels bereiken met deze punten. We hopen dat je level vier 'Attending' in ieder geval zult halen halen voor vrijdag.

Problemen?

Indien er een probleem is waarbij je vastloopt, je niet verder kunt spelen of je een vraag hebt over het onderzoek zelf stuur dan een mail naar olivierrichters@hotmail.com of bel 0658814550.

Het spel terug geven

Vrijdag tijdens wanneer de AMS-1 vragenlijst en ICK-toets maakt verwachten wij dat je het spel terug geeft aan ons.

Appendix F: Evaluation questionnaire

Dear student,

We would like to thank you for participating in our research effectiveness study of education in Patient Safety. We would like to ask you some question to evaluate the game Air Medic Sky-1.

External variables:

1. How experienced are you in electronic games such as gaming apps on your mobile, games for the pc, consoles etc. ?

   0 - NONE: I have no experience in gaming. I do not play games at all.
   1 - BASIC: I have some basic gaming experiences, I play 0– 2 hours per week.
   2 - MODERATE: I have basic gaming experiences. I play 2-4 hours a week .
   3 - INTERMEDIATE: I consider myself experienced. I play 4-6 hours a week.
   4 - EXPERT: I consider myself being an expert in gaming. I play 6+ hours a week.

2. I played AMS-1 until the end: YES/NO
   If ''no'', where did you stop and why? …….
3. I have played AMS-1 : …..hours…..minutes
4. Extra vraag: During this week, did you allow access to the e-module or AMS1 to someone participating in the research who was not officially allowed access?

		Strongly disagree	disagree	Neutral	Agree	Strongly agree
	Usefulness:					
1	AMS-1 helps me to be more effective at Patient Safety					
2	AMS-1 is useful for Patient Safety education					
3	AMS-1 gives me control on Patient Safety activities during my clerkship					
4	AMS-1 meets my needs					
	Ease of use:					
5	AMS-1 is easy to use					
6	Learning to operate AMS-1 initially is full of problems					
7	I easily remember how to use AMS-1					
8	I quickly became skillful with AMS-1					
	Satisfaction:					
9	I am satisfied with AMS-1					
10	I would recommend AMS-1 to a fellow students					
11	AMS-1 is fun to use					
12	AMS-1 is pleasant to use					

	Engagement:					
13	When playing AMS-1, my attention was entirely on the game					
14	When playing AMS-1, I felt actively involved in the game					
15	I could concentrate fully throughout the game					
16	The content of the game is educational					
17	I play AMS-1 without thinking about how to play					
	Attitude towards using:					
18	It is a good idea to use the AMS-1 for my study					
19	I like the idea of using a game (AMS-1) to learn about patient safety topics					
20	Using AMS-1 provided me with a lot of Enjoyment					
	Behavioural Intention to Use:					
21	I would have liked to continue playing AMS-1					
	Motivation:					
22	I think it's important for medical students to learn about patient safety topics					
23	I like what we are learning in this game					
24	I can use the things we learn in this game in my future work as a doctor					
25	I try to learn more about patient safety to be a better doctor in the future					

List the most **negative** aspect(s):

1.
2.
3.

List the most **positive** aspect(s):

1.
2.
3.

Appendix G: **Summary interviews**

Interview 1 (2 participanten)

Difficult, breathing exercises took too long. A small distraction in your surrounding and you needed to start over, again 10minutes. Because of this I skipped a lot. It would be fun to combine it with an action. Because I needed to focus a lot, my heart rate went up instead of down. We believe it works (gave example), we would use this during our clerkship, but behind a pc it is hard to practice.

It was hard to keep our attention at the lectures. It was not interactive like the 9 flag game, it should be like the 9flag game. At the 9flag game I remembered it, not with the use of the lectures. Because of this I did not watch a lot of lectures, boring, not a lot of time during the week. The score system was nice, you get further the more you played, but as soon as I could continue to the missions I did. The aircraft lecture was good, but I would not try as a co-assistant to say something to a high ranked surgeon: ''I think you are doing something wrong here''.

The missions were very realistic . It took a while before I got used to the UI, but I had a lot of fun after that. I had the idea that playing wise I learned a lot.

The storyline was a good/fun thing, rewarding system (scores) to get to the missions. The rewarding system could be more elaborated, because now you only get a different title.

We would not make this game obligated, it is not ready for that, things need to be changed.

We had to little time to play. We really liked it and would have liked to play it for another week.

The self-rated skill questions were hard to estimate, but we think you can improve these skills with the game.

At the evaluation questionnaire I put a lot of ''neutral'', because I did not play the game fully till the end to get a good opinion about this.

Interview 2
I enjoyed playing the game, but the breathing exercises did not work well.
Ik vond het heel leuk speel. Maar de ademhalingsoefeningen werkte niet goed. Prone to errors, I had sweaty hands and this interfered.

The game is very educational, but the lectures were very boring. The red flag game was fun to learn theorie. I did learn something from the lectures, but at some I lost my attention because I did not get the relevance (aircraft).

The missions were fun, but I got an error whereby I had to restart the game. I did not have a save game, because I was just enjoying playing and continued playing without thinking about that. I enjoyed the first two missions so much that it was worth to do it all over again.

It is fun learning thoughout a game, making your own decisions. I did not play long enough, would have liked to continue.

My engagement was low because of the lectures.

I don't think this game is important to be played before the clerkships, because we do not have a lot of responsibility there. It might be better for AOIS (specialist in training).

I would not dare to just point my finger up at the clerkship if I think that something is going wrong, because that person will judge/rate me for my clerkship and because my opinion won't be taken into account. The game does not give me more self-confidence. It give me more knowledge what to do when a patient comes in. The game is not realistic enough to give me stress.

Why does patient safety topics not return in the missions, that is something I wondered

Interview 3
Played for 3.5 hours. Nice game to play, being interactive. Biofeedback worked ok, not super. It was kind of boring. I did use one of the exercises during my clerkship and it helped me relax. I did not know this exercise until I played AMS-1.

I did almost all lectures. Lectures were fine, except for the English sometimes. One lecture was not clear enough; about addiction. I missed a pause button during the lectures, but it worked good, no extra teacher in real life is needed for this.

Missions were fun, I did four. One mission I could not do anything, it kind of froze, no options were available anymore. I missed realism, so I did not feel any pressure or stress.

The score system did motivate me to continue with more lectures etc.

For self-management it may have it effects, also for knowledge, but for our real clerkship not much effect. I will not work in teamwork there.

I had engagement problems because it was a busy weak, but also because the lectures were boring. I would have liked to play longer.

The red flags was a perfect example how you should get more knowledge on patient safety in the game instead of the lectures.

Interview 4
It was fun and educational. The missions were the best part, making your own decisions.

I liked that there was a story so that you do not only click in the game. Realism missed sometimes, for example the earthquake. Why did people have a heart attack at earthquake? Only had one person with a fracture.

I have a low heart rate (55), so the biofeedback was very hard. I was running in my room before the exercise so I could finish this part. I think people can and will use this in the clerkship, because they did see that their heart rate dropped because of controlling your breath.

I did see all the lectures, because of the points. I wanted to get the highest rank. The knowledge test was represented for the lectures. Only a few questions where I thought: where does this come from?

I did not see patient safety back in the missions, only in the lectures. The realism missed that I was working in a team.

The red flags made me more aware, I will use this. I can see my supervisor as an equal, so no problem there.

I liked that the lectures were short, so I did not lose my attention to the video. I missed the game element here.

I would advise my fellow students this game. All parts are useful, reducing stress, knowledge etc.

The game did not professionalism. Different speakers from all over. You can see there has been put a lot of time in this game. I would like to see more of leveling in the game, not just points. For example that you have to unlock certain capabilities and it gets harder the longer you play.

Interview 5
The first part was very long, the breathing exercises. Did not always work. I was happy when I got to the mission part. I do think it works and people would use it to reduce stress, but it does not always work so well.

I have seen all lectures. I was concentrated because they were short. Had no problem watching them.

It was not hard to learn how to play the missions. A little try and error in the begin.

I did not get any stress in the game, because nobody was watching and it is fake.

I think I would gain equal knowledge with AMS-1 or a tutor. The difference might be that I like the lectures more so I do not have to read at home.

The score system did motive me. I wanted to get to the second highest level.

I only had patient safety educations once, but it was a very short one. It was about mistakes in hospitals. So I think it is a good idea if we get more lessons in patient safety.

I think the red flag game is important, but I'm not confident enough yet to always say what I think. Do I have enough knowledge to say something about this?

I don't know how much effect this game will have on me. In real life things always will go differently, realism.

Interview 6
It was fun, but the breathing exercises did not add any value. You can't use this in real life ''hey all wait please and let me rest 1 minute with my breathing''.

I liked that there was a board where text was written on during the lectures.

I absolutely loved the missions, but sometimes I wanted to do something and it wasn't there. It has its limits, you can't do everything like in real life.

I will use the red flags. But not because I did this game, it is kind of obvious that if the nurse is talking about the left leg and the surgeon the right leg that I say something.

I learn best when I read and do things at the same times. Since the game is interactive it worked for me.

I would advise this game, but not being mandatory. The red flags and missions are fine, but the game as a whole not perfect.