E.V.A.C. – A VALIDATION STUDY
A content and construct validation of a serious game about fire safety

Oddur Vilhjálmsson

Faculty of Industrial Engineering, Mechanical Engineering and Computer Science
University of Iceland
E.V.A.C. – A VALIDATION STUDY
A content and construct validation of a serious game about fire safety

Oddur Vilhjálmsson

30 ECTS thesis submitted in partial fulfillment of a Magister Scientiarum degree in Computer Science

MS Committee
Per Backlund
Ebba Þóra Hvannberg

Master’s Examiner
Björn Berg Marklund

Faculty of Industrial Engineering, Mechanical Engineering and Computer Science
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, September 2017
E.V.A.C. – A Validation Study
A content and construct validation
of a serious game about fire safety
30 ECTS thesis submitted in partial fulfillment of a Magister Scientiarum degree in
Computer Science

Copyright © 2017 Oddur Vilhjálmsson
All rights reserved

Faculty of Industrial Engineering, Mechanical Engineering and Computer Science
School of Engineering and Natural Sciences
University of Iceland
Dunhagi 5
107, Reykjavik
Iceland

Telephone: 525 4700

Bibliographic information:
Oddur Vilhjálmsson, 2017, E.V.A.C. – A Validation Study, Master’s thesis, Faculty of
Industrial Engineering, Mechanical Engineering and Computer Science, University of
Iceland, pp. 58.

Printing: Háskólaprent
Reykjavik, Iceland, September 2017
Preface

This master project was carried out as a part of an exchange study programme at the University of Skövde in Sweden as a joint project between the University of Skövde and the University of Iceland. It is comprised of a 22.5 ECTS thesis that follows this preface and a 7.5 ECTS course called Serious Game Evaluation. Supervising the project are Ebba Dóra Hvannberg from the University of Iceland, and Per Backlund from the University of Skövde. The examiner is Björn Berg Marklund from the University of Skövde.

The aim of the course was to teach appropriate methods for experimental evaluation of games. Among its objectives were the planning, implementation and reporting of an experimental pilot study where a serious game was tested. The effects of the game were analysed. Qualitative and quantitative data was collected and analysed. Societal and ethical aspects on game studies were accounted for. The course covered various methods for analysing serious games, including statistical methods such as t-tests, chi-tests and linear regression.

There were three noteworthy assignments in the coursework. Firstly, an analysis of a research paper. Each of the students selected a peer reviewed research paper, examined the method presented in it and reported it to the class. Secondly, students were grouped up (3-4 in each group) and tasked with performing a small statistical study and conducting a statistical test. Lastly, each student had to plan and conduct an experimental pilot study, preferably related to the student’s thesis. The lessons learned during the pilot study were used as the basis of the method chapter in this thesis.
E.V.A.C. – A VALIDATION STUDY
A content and construct validation of a serious game about fire safety

Master Degree Project in Informatics
One year Level 22,5 ECTS
Spring term 2017

Oddur Vilhjálmsson

Supervisors: Per Backlund & Ebba Þóra Hvannberg
Examiner: Björn Berg Marklund
Abstract

Serious games are games with a purpose beyond entertaining the player. In order to prove that a serious game meets its purpose it must be validated, which is the process of collecting validity evidence to evaluate the various attributes of the game and the effect it has on the players. This thesis is a validation study of E.V.A.C. – a prototype serious game about fire safety. It is intended to demonstrate how one might go about gathering validity evidence for construct and content validity of a serious game. Construct validation evaluates whether the game does what it is supposed to do, make the players think and learn about fire safety. Content validation estimates the validity of the lessons based on a detailed examination by experts. A series of playtests was conducted with firemen and students aged 11-12. It was found that the game demonstrates good construct validity and acceptable content validity.

Keywords: Serious games, validation study, evacuation, construct validity, content validity
# Table of Contents

1 Introduction .................................................................................................................. 1

2 Background .................................................................................................................... 2
  2.1 Serious Games .......................................................................................................... 2
  2.2 Validity ...................................................................................................................... 4

3 Problem .......................................................................................................................... 6
  3.1 Method ...................................................................................................................... 8
    3.1.1 E.V.A.C. .............................................................................................................. 8
    3.1.2 Test Procedure .................................................................................................. 13
    3.1.3 Data Analysis ..................................................................................................... 14
    3.1.4 Collaboration ..................................................................................................... 14
  3.2 Ethical Considerations ............................................................................................... 15

4 Results ............................................................................................................................ 16

5 Analysis .......................................................................................................................... 18
  5.1 Construct Validity Analysis ....................................................................................... 18
  5.2 Content Validity Analysis ......................................................................................... 20

6 Conclusions ..................................................................................................................... 21
  6.1 Summary of Results ................................................................................................. 21
  6.2 Discussion .................................................................................................................. 21
  6.3 Challenges during the Validation Process ................................................................. 22
  6.4 Future work .............................................................................................................. 22

References .......................................................................................................................... 23
1 Introduction

Serious games are games with a purpose beyond entertaining the player (Mohan, Angus, Ricketts, Farris, Fischhoff, Rosengart, Yealy and Barnato, 2014, p. 2). This purpose can be anything ranging from education falling under digital game based learning (DGBL) to advertising.

Serious games’ content needs to be validated to prove that the games meet their purpose. Validation is a process where validity evidence is collected to evaluate the appropriateness of the decisions, interpretations and uses based on assessment results (Cook and Hatala, 2016, p. 2). Research shows that few games have been validated (Graafland, Schraagen and Schijven, 2012). Most thorough validation efforts found during the writing of this thesis seem to be focused on games related to the medical industry. This is to be expected since the accuracy of the lessons in those games is of critical importance.

This thesis is a validation study of E.V.A.C. – a prototype serious game that was developed at the University of Skövde at the request of the rescue service in Gothenburg. The purpose of the game is to get children to think about how they would behave in an emergency situation such as a fire. The game is intended to be used as supplementary material for firefighters during training sessions for children learning about fire safety. It is the intent of this thesis to demonstrate how one might go about gathering validity evidence for the construct and content validity of a serious game, using E.V.A.C. as an example. This work may then be of use when validating other serious games.
2 Background

2.1 Serious Games

Any game that has a purpose beyond entertaining the player, be it education, training, marketing, exercise etc. can be considered a serious game (Mohan, et al., 2014, p. 2). Serious games can be designed with this purpose in mind or they can be commercial games that have been modified to fit the purpose. Commercial off-the-shelf games can even be considered serious games if they are used for something beyond entertainment (Susi, Johannesson and Backlund, 2007, p. 5).

There are many different types of games that have been used in the field of rescue services. A few examples include:

ADMS-BART (Kobes, Helsloot, de Vries and Post, 2010) is a virtual environment simulator which was used to simulate a hotel and study which path the players took to the exit. This game was validated by comparing it to an evacuation of a real hotel on an individual basis. Based on their relative validation analysis, Kobes, et al. (2010, p. 49) claim that “the use of ADMS-BART can be considered valid as a research tool for research on wayfinding performance”.

The dangerous kitchen (Norfolk fire and rescue service, 2011) and Escape planning game (London fire brigade, 2016). These games fall under the point and click genre, as their objective is to identify the fire hazards in a room by clicking on them. No text mentioning validation of these games was found.

Eldracet (Trygg-Hansa, 2014) is a platformer game where the player is tasked to run out of a burning building while avoiding obstacles. No text mentioning validation of this game was found.

SPEED (Oliveira, Pereira, Almeida, Rossetti and Oliveira, 2015) and the prototype described by Chittaro and Ranon (2009) are first person view 3D games. These two games simulate a building, a university building and an auditorium respectively, and task the players with leaving as quickly as possible when they hear the fire alarm go off. Oliveira, et al. (2015, p. 129) describe how SPEED was used to gather data on “human behaviour when facing the urgent need of evacuating from an unknown building and having to deal with a set of unexpected situations and obstacles”. They compare the results from a test they performed on 19 children to a similar test on adults, but do not clarify whether that test was performed by them or someone else. No details were given on the characteristics of the sample of adults, and the comparison is not set up as a validation of the game itself. Chittaro and Ranon (2009, p. 79) tested their game on 7 adults but the focus was on their behaviour during gameplay and not on validating the game.
Other examples of serious games from different genres, that do not involve fire safety include:

**Vital Signs: ED** (Mohan, et al., 2014) is a cross between a simulator and a puzzle game in that it is simulating the environment of an emergency centre, but the gameplay consists of assigning correct treatments to patients within a time limit. The game was tested on 142 physicians who were randomly assigned to one of two groups, one control and one which played the game with increased cognitive load. The cognitive load group had poorer performance than the control group, indicating construct validity (Mohan, et al., 2014, p. 6).

**SICKO (Surgical Improvement of Clinical Knowledge Ops)** (Lin, Park, Liebert and Lau, 2015) is a game that could be considered to fall under the puzzle game genre, in which the player must assign correct treatments to patients before a timer runs out. The game was made “with the intention of assessing the construct of surgical decision making” (Lin, et al., 2015, p.80). The game was validated by testing it on medical students and sub interns, 49 subjects in total. They were ranked into four groups according to their level of expertise and the scores were compared between groups. Each group significantly outperformed the less experienced groups, demonstrating construct validity (Lin, et al., 2015, p. 85).

**Bool the Miner** (Aufheimer, Bonenberger, Klein, Kurashvili and Rogers, 2016) is a combination of the puzzle and jump’n’run genres. It is an educational game intended to teach students about Boolean algebra. The game was tested on 84 participants by a pre-test, gameplay session and a post-test. The difference in the test scores indicated knowledge gain, but the article does not fit well with the processes of face, content, construct, concurrent, predictive or discriminate validation.
2.2 Validity

All of the games mentioned in chapter 2.1 were designed with some serious purpose in mind. In order to know to what extent a serious game fulfils its purpose it must be validated.

As stated by Cook and Hatala, (2016, p. 2) “Validation refers to the process of collecting validity evidence to evaluate the appropriateness of the interpretations, uses, and decisions based on assessment results.” In other words, the game must be tested and analysed. This process is important for several reasons. For instance, the validity evidence can reveal flaws in the design, can provide confidence in the quality of the education material, and reassurance that the score results can be trusted to correlate to actual proficiency in that which is being trained. As such there are several types of validation methods which focus on different aspects.

Content validity is “an estimate of the validity of a testing instrument based on a detailed examination of the contents of the test items” Gallagher, Ritter and Satava (2003, p. 1526). In essence this means that e.g. an educational game contains all the lessons that it is intended to cover. This test is largely subjective, and is reliant on expert reviews.

Face validity is “a type of validity that is assessed by having experts review the contents of a test to see if it seems appropriate” Gallagher, et al. (2003, p. 1526). This is about assessing the realism of the constructs, to see if the test measures what it is supposed to measure, or in the case of an educational game whether it teaches what it’s suppose to teach. This is subjective, and usually most useful in the earlier stages of development.

Construct validity is “a set of procedures for evaluating a testing instrument based on the degree to which the test items identify the quality, ability, or trait it was designed to measure” Gallagher et al. (2003, p. 1526). There is an inherent difference in the skill levels of experts and novices which should be visible in the test results. A common way to demonstrate construct validity is to show that experts perform better than novices at a given task. Other methods include factor analysis, correlation with another measure of the same construct, and change or stability over time (Cook and Hatala, 2016, p. 3).

Concurrent validity is “an evaluation in which the relationship between the test scores and the scores on another instrument purporting to measure the same construct are related” Gallagher et al. (2003, p. 1526). For example, comparing different media e.g. an educational game and a lecture or an educational movie.

Predictive validity is “the extent to which the scores on a test are predictive of actual performance” Gallagher et al. (2003, p. 1526). The game has predictive validity if it accurately predicts who will perform the actual tasks (evacuation) well and who will not. This requires a comparison to a performance in real life after the performers have been trained using the game. This can be prohibitively expensive and/or dangerous depending on the tasks being taught.

Discriminate validity is “an evaluation that reflects the extent to which the scores generated by the assessment tool actually correlate with factors with which they should correlate” Gallagher et al. (2003, p. 1526). In other words the results of a test should correlate strongly with factors that it’s supposed to correlate with, and it should not correlate with factors that it’s not supposed to correlate with.
Graafland, Dankbaar, Mert, Lagro, De Wit-Zuurendonk, Schuit, Schafstal and Schijven (2014, p. 5) suggest an order to the validation processes. First are content and face validity, they are relatively easy to test and they make sure that the project is on the right track. Then comes construct validity, to check that the representation of the lessons is well formulated and that the game works as intended. Next is concurrent validity, which will reveal whether the project is better or worse than different types of media or other similar projects and if there are any unexpected deviations. Last but not least is predictive validity, to examine the knowledge transfer and ascertain e.g. whether performance in the game correlates with performance in reality.

Graafland, et al. (2012) performed a literature review of 25 articles about serious games in the field of medicine with the intent of evaluating the validity testing of that type of games. The articles described 30 games, and they found that while eight of them had undergone some steps in validity testing, none of them had completed a full validation process. Of those eight games, three had undergone concurrent validation and one underwent face validation. The rest either did not meet the criteria for a specific validation process or the study design did not correspond to formal validity testing.

Barsom, Graafland and Schijven (2016) performed a literature review of articles concerning augmented reality (AR) applications for training medical professionals. They found 27 studies describing a total of seven AR applications. Two applications had undergone face-, construct- and concurrent validations, one underwent face- and construct validations and one was validated for face-, content- and construct validity. None of the applications had completed a full validation process as laid out by Graafland, et al. (2014, p. 5).

Wang, DeMaria, Goldberg and Katz (2016) performed a systematic review of serious games in training health care professionals. They identified 42 serious games for medical education and found that 33 of them (79%) had undergone some form of evaluation as a teaching intervention. The ones who hadn’t were mostly developmental or pilot studies. Only one of the studies evaluated more than just the gain in knowledge or skill (Wang, et al., 2016, p. 48).
3 Problem

As can be seen in the literature reviews by Graafland, et al. (2012), Barsom, et al. (2016) and Wang, et al. (2016) there are many serious games that are insufficiently validated. Games that have undergone some validation process are commonly tested with regards to construct validity (Lin, et al., 2015), (Mohan, et al., 2014), predictive validity (Kobes, et al., 2010), face validity (Graafland, et al., 2012) and concurrent validity (Graafland, et al., 2012) as can be seen in Table 1. However, studies of serious games are frequently formulated as attempts at measuring knowledge gain rather than as rigorous validation tests (Wang, et al., 2016, p. 48).

Cook and Hatala (2016, p. 8-9) point out that a successful construct validation provides only weak evidence for the quality of the test, as “the difference in scores may arise from a myriad of factors unrelated to the intended construct”. However, not finding a difference in the scores where a difference was expected is strong evidence that the test is flawed in some way.

The aim of this thesis is to demonstrate how a formal validation process of content and construct validity of a serious game could be set up. A prototype serious game called E.V.A.C. will be used as an example for the validation processes. It is a convenient subject as it was developed by the author of this thesis along with seven other students during the autumn of 2016, and has not undergone any validation process before.

Thus, the research question of this thesis is: How well does E.V.A.C. meet the criteria for achieving content and construct validity?

The content and construct validity processes were chosen to be the focus of this thesis because they are important and useful first steps to a full validation process. If any problems or defects are revealed during the content and construct validation processes then they should be addressed before the other, more resource demanding processes are performed. A concurrent validation, being a comparison to other similar games or other types of media, is more useful when it has been ensured that all the intended lessons are well represented in the game. Also, a predictive validation, where performance in the game is compared to performance in reality, would be prohibitively difficult to organize as it would require the test subjects to perform an actual evacuation after playing the game, thus it was deemed to be out of scope for this thesis.

According to Graafland, et al. (2012, p. 1324) the criteria for good content validity is: “A uniform and positive evaluation of game content and associated testing parameters” by a panel of experts. Content validity is largely subjective, and reliant on the experts’ judgement of how relevant the materials are (Gallagher, et al., 2003, p. 1526). The criteria for good construct validity is that the experts perform significantly better than the novices.
Table 1: An overview of which validation processes have been applied to the various serious games presented in section 2.1. Parentheses means that the process was only partially applied or that the test was not properly set up as a validation process although much or even all of the requisite validity evidence for that process had been gathered.

<table>
<thead>
<tr>
<th>Games presented in section 2.1</th>
<th>Face</th>
<th>Content</th>
<th>Construct</th>
<th>Concurrent</th>
<th>Predictive</th>
<th>Discriminate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADMS-BART</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>The dangerous kitchen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Escape planning game</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eldracet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEED</td>
<td></td>
<td></td>
<td></td>
<td>(X)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unnamed prototype by Chittaro, et al. (2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vital Sign: ED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>SICKO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Bool the miner</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below are games mentioned in the results section of the literature review by Graafland, et al. (2012, p. 1325-1327) as games that have undergone steps in validity testing.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virtual ED</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Virtual ED II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Triage Trainer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>CAVE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Nuclear Event Triage Challange</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(X)</td>
<td></td>
</tr>
<tr>
<td>Radiation Hazards Assessment Challenge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Knee Arthroplasty game</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>OLIVE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.1 Method

3.1.1 E.V.A.C.

E.V.A.C. is a serious game developed by the author of this thesis along with seven other students during the autumn of 2016 as a part of the Serious Games Project course at the University of Skövde. The iterative developmental cycle for E.V.A.C. evolved from paper prototypes, playtest sessions and assessment to the current version that is used for this comparative study. The game was requested by and developed in close contact with Räddningstjänsten Storgöteborg (RSG). The feedback received from the client had a considerable impact on the game’s overall design and structure during the iterative process.

RSG is the fire rescue service for the different municipalities of Gothenburg. They work to prevent and limit fires and other emergency situations in the entire region. A large part of the work with preventing fires involves education of society about different causes for emergencies and how they can be prevented.

One of the non-functional requirements set by RSG was that the game should be accessible to a broad audience. With that in mind, the game was designed to contain as little written language as possible. Instead, its instructions, objectives and feedback are presented in a symbolic manner. The symbols were designed to be self-referential, so that it would be easier to understand them no matter what background the user has. For example, to indicate that the button “c” is used to crouch, a picture of a crouching stick figure next to a picture of the button is used. An overview of the game controls was created to present to the players. It contains pictures of the relevant keyboard buttons along with images indicating what the buttons do. This overview can be seen in Appendix A.

Information in-game is also presented in a symbolic manner. For example, interactable objects have an icon floating on them, the elevator has pictures of keypad numbers next to it showing which floor it is on, and if a player triggers a fail state then a picture pops up showing which fail state was triggered.

The levels are structured in such a way as to minimize cognitive overload and build on each other like a tutorial, starting easy and getting progressively harder. Table 2 lists all the lessons in the game, and shows how they are distributed among the levels. Screenshot images of the levels can be seen in Appendix G.

The first level is intended to introduce the basic movement and navigation mechanics to the player. It does not have a fail state, unlike the other levels, and only focuses on two lessons: “Getting yourself to safety” and “Noticing the evacuation signs to know where to evacuate”.

The second level introduces fail states, fire, smoke and the alarms. This level starts with a fire burning next to the player. The fire alarm does not trigger automatically, causing a fail state if the player forgets to trigger it manually.

The third level introduces the elevator as well as forcing the player to learn how to crouch under smoke. When the player is about halfway through the level a fire starts. This time the alarm triggers automatically. The location of the fire and the speed of the smoke spread are designed in such a way that it should be impossible to get out of the building without crouching. At the end of the level the player is asked if there is anyone left inside the building. There is a possibility of one non-player character (NPC) being inside, as the player has already
just moved past him when the alarm is triggered. The purpose of the question is to get the player to try to memorize the NPCs he has encountered.

The fourth level is bigger and has a lot more NPCs, making it more difficult to memorize which ones, if any, are left inside.

Table 2: The lesson distribution in the levels of E.V.A.C.

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Getting yourself to safety</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Don’t walk into a fire</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Noticing the evacuation signs to know where to evacuate</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Alarming when upon noticing a fire</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Crouching when in smoke</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Close doors to prevent fire and smoke spread</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Noticing potential fire hazards</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Not using elevator in case of fire</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Don’t press the alarm when no apparent fire emergency</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Remembering who is left in the building</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Remembering on which floor the remaining NPCs are</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Guiding people out</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Presentation of the lessons:

Getting yourself to safety: The level ends successfully when the player gets outside after a fire has started, and ends unsuccessfully if the player fails in some way.

Don’t walk into a fire: The level ends when a player walks too close to a fire, and a picture of a stick figure standing in a fire is shown (Figure 2).

Noticing the evacuation signs to know where to evacuate: There are several evacuation signs in the level which point in the direction of the nearest exit.

Pressing the alarm after noticing a fire: There is an alarm trigger on every floor. If the player forgets to press it after a fire has broken out and exits the building, the level with end unsuccessfully with a picture of the alarm triggers (Figure 3).
Crouching when in smoke: The fire in the game generates smoke. If the player stands in it for too long, the level will end unsuccessfully with a picture of a stick figure crouching under a cloud of smoke (Figure 4). If the player crouches, he can stay under the smoke for as long as he likes.

Close doors to prevent fire and smoke spread: If the player closes a door before the smoke cloud has reached it, the smoke will stop by the door.

Noticing potential fire hazards: There are a couple of potential fire hazards in the game, such as a couch with a phone charger lying on it, which will catch fire in the beginning of level 2.

Not using the elevator in case of fire: When a fire starts the elevator closes its doors, shuts down and doesn’t accept commands from the player anymore. A red light starts flashing above the elevator doors, indicating the danger. If a player is inside the elevator when it shuts down it will carry him to the most recently selected floor to let him out. After that it is locked.

Don’t press the alarm when there is no apparent fire emergency: If the player presses the alarm before a fire has broken out, the level will end unsuccessfully with a picture of the alarm triggers (Figure 3).

Remembering who is left inside the building: When levels 3 and 4 are completed, the players are asked if they remember whether anyone else is still inside the building. For this prototype the question is asked by a test administrator. Answering this question is supposed to become very difficult in level 4, since it has 10 NPCs that use the same character model. The only difference between them is the colour pattern on their shirts.

Remembering on which floor the remaining NPCs are: After answering whether anyone is inside, the player is asked where the missing NPCs might be.

As a way to enhance the lessons and encourage the player to reflect on them E.V.A.C. uses the symbolic pictures, as seen in figures 2-4, after every fail state. If a player fails, the game stops and the player is presented with the picture. In order to continue the player has to navigate the mouse cursor to a green button with an arrow at the bottom-right corner of the picture (Figure 5). Upon noticing the player reaching a fail state, these reflective segments are further enforced by the test administrator with the question: “Why did the game stop?”

There is another type of reflective segment within the game that is not presented with a picture or after a fail state. This reflective segment occurs upon the completion of level 3 and 4. Instead of the “level complete” picture at the end of a completed level (Figure 6), the game stops at the point where the player can see all the NPCs that successfully got out of the building by the same exit as the player (Figure 7). This segment is further enforced by the test administrator asking the two following questions: “Is anyone left in the building” and “Do you know where they are?”
Figure 2: This picture indicates that the players should not walk too close to the fire.

Figure 3: This picture indicates that the players should trigger the alarm only when they can see fire or smoke.

Figure 4: This picture indicates that the player should crouch under smoke.
Figure 5: Level 2 - The player steps too close to the fire.

Figure 6: Level 2 completed.

Figure 7: Level 3 completed. The player can see which characters are outside.
3.1.2 Test Procedure

A manuscript of the test procedure was used to ensure that the test subjects would all receive the same information, and have as uniform experience as possible. The full manuscript can be seen in Appendix B.

First, the game and its controls are introduced to the players, and they are informed that they can stop the test at any time without giving a reason for doing so. Next, the players play the game. If they have any questions about the controls then those questions are answered, but if they have any questions about what to do they get limited answers such as “Explore the building”. While playing, the players are asked questions at specific times. If the level ends due to the players walking too close to the fire, standing in smoke or pressing the alarm prematurely then they are asked why the game stopped. At those times there is also a picture on the screen indicating what happened. At the end of levels 3 and 4 the players are asked if there is anyone left inside the building, and if so then where. Their answers to these questions are not recorded, as the purpose of them is primarily to make the players think about their actions. When the players have finished the fourth level, or when they have played for about 15 minutes, the play session is stopped and they are asked to answer the questionnaires. The players are not observed directly when answering the questionnaires, as the feeling of being watched may result in dishonest answers if a subject interprets it as a lack of anonymity. The test administrator is however close by in case the players have any questions about the questionnaires.

The construct validity questionnaire starts with three background questions: Age, gender and game literacy. Then there are thirteen multiple-choice questions about the lessons in the game. Each of them has four choices, only one of which is correct. The choices were made to look very similar or appear to be equally likely to be true so that it would be hard to guess the right answer. The last four questions are relevant to the third and fourth levels of the game, so the player is asked if he played the third level. If he did not, then those questions are not presented to him.

The content validity questionnaire starts with an open question asking the subject to identify and list as many lessons from the game as possible. Then they are shown a list of the lessons that were designed to be in the game and asked if they agree that their representation in the game was good on a scale of 1 – 6, where 1 means disagree and 6 means agree. An even numbered scale was used in order to force the subjects to make a choice instead of allowing them the option to be completely neutral. After all twelve lessons had been rated, the subjects got two more open questions asking whether the game was missing any critical lessons, and if they had any comments on the representation of the lessons. Then they were asked how suitable they considered the game to be for children and for adults on a scale of 1 – 6. Lastly, they got an open question asking if they thought that the game is a useful tool for fire safety education.

Both questionnaires were implemented using Google Forms. The questionnaires can be seen in Appendix D.
3.1.3 Data Analysis
After the tests were carried out the results were sorted into three groups: Rescue personnel, students with feedback and students without feedback. Each question in the questionnaire has only one correct answer, so the statistic to be measured is the percentage of correct answers. The groups were compared using an unpaired t-test. First the rescue personnel group was compared to the student group where all the students were included. Then the groups were sorted into subgroups by whether they had played the game or if they saw someone else play it, whether they had the reflective feedback or not, and by the background questions, game literacy, age and gender. The subgroups were then compared using unpaired t-tests e.g. is there a difference between the firemen and students who just watched someone else play the game, or is there a difference between a student with low game literacy and a student with high game literacy? If the rescue personnel group generally performs better than the student groups, then that indicates that the game has good construct validity.

The game literacy score is derived from the background questions. There are three Likert scale questions asking how used the subject is to playing mobile, console and computer (PC, Mac) games. There are 6 options, ranging from “I never play …” to “I play nearly every day”. These answers are assigned a value of 0-5 respectively and the values from the three questions are added together to get a game literacy score ranging from 0 to 15.

3.1.4 Collaboration
This study was performed in collaboration with Anita Stenholm. The focus of Stenholm’s thesis is on the knowledge transfer of E.V.A.C. to the target audience, 10-16-year-old students, which made it convenient to use the same tests for both studies. As they are both studying the same serious game, and use the same tests, they have the same description of the game and test procedure. Those are chapters 3.1.1 and 3.1.2 in this thesis, and they were written in collaboration with Anita Stenholm.

Stenholm’s thesis does not use the results from the firemen. On the other hand, it adds a pre-test and an immersion questionnaire for the students, the results of which are not used in this thesis. All, Plovie, Nuñez Castellar and Van Looy (2015) researched the effect a pre-test has on the effectiveness assessment of serious games. They found that “in the traditional classroom context, the pre-test makes the participants more sensitive to the content treated in the intervention while administration of a pre-test does not influence outcomes of the DGBL treatment” (All, et al., 2015, p. 623). This thesis does not compare the effects of the game to that of a traditional classroom, so the pre-test should not have an effect on the results.

In order to test the reflective segments for Stenholm’s thesis, a value-added approach was used. A value added approach consists of testing the effects of adding features to an educational game on learning quality (Erhel and Jamet, 2013, p. 157). Two versions of E.V.A.C. were used, one with the reflective segments enabled and one without. The test administrator did not ask participants that played the game without reflective segments any questions.

\[\text{Stenholm’s thesis is expected to be published at the University of Skövde in August of 2017.}\]
3.2 Ethical Considerations

This study was performed in accordance with the ethical guidelines compiled by the Swedish Research Council (Vetenskapsrådet, 2009).

All participants are informed of the purpose of the playtest, and told that they can stop the tests at any time without giving any reason for doing so. This is ensured by following the test procedure manuscript in Appendix B - . By playing the game, the participant agrees to take part in the study. In the case of the students, a parent’s or guardian’s signature on the permission slip seen in Appendix E - was also required. All the data collected is anonymous and can’t be traced back to the testers. The data is used only by the researchers for this particular study.

The participants were not observed actively while answering the questionnaires, unless they needed help, as a perceived lack of anonymity, justified or not, may result in disingenuous answers to the questionnaires as the subject might feel that he will be judged based on his answers.

A decision was made during the design of the game to minimize potentially traumatic experiences. As such there are no deaths in the game, and the failure states consist of the game stopping and displaying one of the pictures seen in Appendix F - .
4 Results

Three test sessions were conducted, one at RSG and another at Räddningstjänsten Östra Skaraborg (RÖS), which is the fire station in Skövde, and lastly one at Käpplundaskolan, which is a local elementary and middle school in Skövde.

The test at RÖS went according to the planned procedure as described in the manuscript, and the game was tested by 11 firefighters. There were 4 computers available, and two test administrators, so four firefighters could play simultaneously with each administrator observing two subjects at once. No major problems were encountered in regards with asking the questions.

The test at RSG did not go exactly as planned, as there were only about 20 minutes available for testing an entire class of 19 students on 4 computers with two administrators. This problem was solved by the use of projectors. Two of the computers were connected to projectors in separate rooms, and the other computers had headphones. This way 10 students had an opportunity to play the game while the others watched. Not all of the students finished the game, and there was a lot of excitement causing a few of the questions to be missed. Due to the limited time the students were asked to answer the questionnaires once they got back to the school. Their teacher was asked to oversee it, and they ended up answering the questionnaires one week after playing the game. Two firemen were also tested at RSG, but that test was conducted before the students arrived, and thus in a much calmer environment and according to the plan.

The test at Käpplundaskolan was more organized and went more smoothly. There were 8 computers available and two test administrators. The game was set up on four of the computers, and the other four were used for the questionnaires. The students were brought in in groups of four and the game was introduced to them simultaneously. When the students had finished the game or played for 15-20 minutes they moved to answer the questionnaires and the next group of four was brought in. 17 students answered the questionnaire in this test.

A total of 13 firefighters and 36 students participated in the study. A summary of the t-test results can be seen in Table 3, and detailed results of the t-tests and their corresponding f-tests can be seen in Appendix C - .
Table 3: A summary of the t-test results.

<table>
<thead>
<tr>
<th>T-test</th>
<th>Number of students</th>
<th>Higher score</th>
<th>Significant difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>All 13 firemen (F) vs all students (S)</td>
<td>36</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F vs S who played the game</td>
<td>27</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F vs S who watched the game</td>
<td>6</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F vs S who didn’t see the game</td>
<td>3</td>
<td>F</td>
<td>No</td>
</tr>
<tr>
<td>F vs S with feedback (FB)</td>
<td>26</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F vs S who played the game with FB</td>
<td>21</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F vs S who saw the game with FB</td>
<td>5</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F vs S who played the game without FB</td>
<td>6</td>
<td>F</td>
<td>Yes</td>
</tr>
<tr>
<td>F with high game literacy (GL) vs F with low GL</td>
<td>-</td>
<td>Low GL</td>
<td>No</td>
</tr>
<tr>
<td>S with high GL vs S with low GL</td>
<td>-</td>
<td>High GL</td>
<td>Yes</td>
</tr>
</tbody>
</table>

As seen in Table 4 there was only one student who just watched the game and reported not seeing any of the feedback pictures, thus it does not make much sense to include that subgroup in the comparisons. That student is still included in the groups “students who watched the game” and “all students”. All the firemen played the game with reflective feedback.

Table 4: Student distribution across game modes

<table>
<thead>
<tr>
<th></th>
<th>Feedback</th>
<th>No feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Played the game</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>Watched the game</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Didn’t see the game</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Total number of students: 36
5 Analysis

5.1 Construct Validity Analysis

Figure 8 and Figure 9 are histograms of the questionnaire results, summarizing the data. Just by looking at the histograms it can be seen that the lowest score achieved by a fireman was 67%, but the scores of the students are more spread out.

![Histogram of questionnaire results](image)

*Figure 8: The score distribution*

![Histogram of questionnaire results](image)

*Figure 9: A further breakdown of the score distribution*

Before every t-test comparison an f-test was performed in order to know whether the variances could be assumed to be equal or not. The details of the f-tests can be seen in Appendix C. The only test where the null hypothesis of equal variances was rejected was when the firefighters were compared to the subgroup of students who played the game without reflective feedback. In all other comparisons the null hypothesis could not be rejected.

All the t-tests except for the case mentioned above were performed assuming equal variances. The details of the t-tests can be seen in Appendix C. In almost all cases where the firemen were compared to the students in some way the null hypothesis of equal means was rejected. The one test where the null hypothesis could not be rejected was where the firefighters were compared to the students that did not see the game. However, that result is not very convincing as that was only a sample size of 3 students.
A t-test was also performed to see if the level of game literacy had an effect. No significant difference was found when comparing firefighters with low game literacy against firefighters with high game literacy. The same test with students did however reveal a significant difference, where the students with high game literacy performed slightly better on average than the students with low game literacy. Figure 10 shows the distribution of scores as a function of game literacy. Again, the details of the tests can be seen in Appendix C.

There were three frequently missed questions, where the rate of correct responses was less than 50%. The first one, with a rate of 9/32 was “What should you do if you notice an imminent potential fire hazard?”. The correct response was “all of the above”. The second one, with a rate of 13/32 was “When should you trigger the fire alarm (if it hasn’t already gone off automatically)?”. Again, the correct response was “all of the above”. Lastly, with a rate of 12/28 was “How can you best be prepared for fire emergencies?”. As before, the correct response was “all of the above”.

Figure 10: Game literacy does not appear to be a relevant factor to the firemen, but it seems to have a small effect on the students
5.2 Content Validity Analysis

The first question “What lessons are represented in the game? Identify and list them” does not seem to have very useful results. Only five test subjects (out of 13) listed multiple lessons, two of the others seem to have been somewhat confused as they wrote their names, and one wrote “asd”. Commonly mentioned lessons were “stay low under the smoke” and “follow the evacuation signs”.

There were five lessons (out of 12) that got an average score of less than 70% from the Likert scale questions about the representation of the lessons. Those lessons were:

- “Remembering who is left in the building” (67.9%)
- “Noticing potential fire hazards” (62.8%)
- “Remembering on which floor the remaining people are” (57.7%)
- “Guiding people out” (55.1%)
- “Not pressing the alarm when there’s no apparent fire emergency” (50%)

The three highest rated lessons were:

- "Crouching when in smoke" (94,9%)
- "Getting yourself to safety" (93,6%)
- "Noticing the evacuation signs to know where to evacuate" (91,0%)

When asked if any critical lessons were missing the most frequent answer was a desire to be able to put out the fire, and a comment that the priority order, i.e. what should you do first, could be better represented. Interestingly, no one mentioned that the game does not remind the user to call the emergency services.

When asked how suitable the game is the average score was 88.5% for students age 10-16, and 80.8% for adults.

When asked if the game is a useful tool for fire safety education all test subjects responded positively. Some added that it needed to be used in conjunction with traditional information and discussion.
6 Conclusions

6.1 Summary of Results

This thesis set out demonstrate one way to collect validity evidence for a serious game by attempting to answer the following question: How well does E.V.A.C. meet the criteria for achieving content and construct validity?

The construct validation was performed by play-testing the game on a group of firemen and a group of students and having them answer questionnaires after the play-test was completed. The game shows good construct validity, as there is a clear difference between the scores of experts and novices. Keep in mind however that, as Cook and Hatala (2016, p. 8-9) have pointed out, such a difference can stem from a myriad of sources, and not finding a difference where a difference was expected is much stronger evidence for invalidity than finding a difference is for validity.

The content validation was performed at the same time as the construct validation by presenting the firemen with an additional questionnaire which focused on the representation of the lessons in the game. The game shows acceptable content validity. There are a few lessons whose representation needs improving, the most important of which are: “Not pressing the alarm when there’s no apparent fire emergency”, “Guiding people out” and “Remembering on which floor the remaining people are” as well as adding a new lesson about calling the emergency services. There are also a couple of software bugs that need fixing. The worst bug that is known at the time of writing this thesis is that a fail state is sometimes triggered in level 2 when the level is supposed to be completed. On the whole the game is already a useable supplementary tool for fire safety education.

6.2 Discussion

While it would have been nice to have more test subjects, the results so far are promising. There seems to be a significant difference between the performance of the rescue personnel and the students, indicating that the game should pass a construct validity test.

The game did pretty well on the content validity test as well, and although there are still several things about the game that need to be improved, it is already good enough to be useful in engaging students to think about fire safety and reflect on the lessons that they learned at the fire station.

One of the firemen commented that the game would be most useful when combined with traditional educational material and discussion, and wouldn’t work as well on its own. This is an interesting comment, because that is exactly how the game was designed to be used. It is meant to be played during a break, or in some other way integrated into the schedule of a session where the firemen are instructing the students in fire safety.
6.3 Challenges during the Validation Process

There were several challenges that had to be overcome during the testing. The first and most serious was the time pressure during the playtest at RSG. There were only 20 minutes available to test the game on the class, which meant that not everybody would be able to play the game. This was solved by setting the game up on a couple of computers that were connected to projectors in two separate rooms. That way a few students could try the game while the others watched the progress. The construct validation questionnaire took this into account by asking whether the subject had played the game or just watched it, or if they hadn’t seen the game at all.

The second problem is also related to the time pressure. The students did not answer the questionnaires at the fire station as they should have, but instead they answered them a week later at the school. This can affect the results of the questionnaires as there is a good chance that the students will forget some of the lessons in the meantime. It does appear though that this did not affect the final results to a noticeable degree.

The third problem is due to the lack of reflective segments in some cases. This meant that some students were a lot more confused than they should have been when they encountered a failure state. Also, it was not possible to guarantee that the students answered correctly when they were asked whether they saw the pictures from the reflective feedback or not. It was necessary to implement this in the game as it is needed for Stenholm’s thesis².

The fourth problem is that none of the firemen who tried the game mentioned that nowhere in the game is there any reminder to call the rescue services (112). This is the kind of critical lesson that is supposed to be caught by a content validation process. The fact that this one didn’t catch it may indicate that the questionnaire needs improving, that more participants were needed, or that a questionnaire is insufficient on its own and should be accompanied by other methods, for example an interview.

These challenges give an idea of how difficult it can be to perform a validation process.

6.4 Future work

RSG has expressed interest in getting a finished version of the game to be hosted on their website. There are several features that need to be changed or added to the game before it is complete. Some lessons need to be improved, as revealed by the content validation, and some additions need to be made such as calling the emergency services and putting out a small fire. The task of asking the players questions periodically should be automated by a virtual firefighter in the game. It will however be a challenge to pose the questions symbolically and without any dialogue. The NPCs should be replaced with various different models so that the colour of their clothes is not the only thing that tells them apart. Once the game has been completed its content and construct validity might be evaluated again and then it may be validated further with regards to concurrent validity, or perhaps even predictive validity. The game was tested exclusively in Sweden, so it may be of interest to see if testing it in other countries will yield different results. It would be particularly interesting to see if the symbols are actually understood the same way by participants from different countries.

² Stenholm’s thesis is expected to be published at the University of Skövde in August of 2017.
References


Appendix A - E.V.A.C. Instructions

Appendix B - Test procedure manuscript

• Introduction
  o Hi. My name is [insert name here]
  o We want you to test a game that we have been working on for our thesis
  o After playing the game you will be asked to answer two questionnaires
  o You can stop the test whenever you want without having to give me a reason
  o The test is expected to take less than half an hour
  o Show an overview of the controls
    ▪ Mention what’s interactable
  o The purpose of the game is to get the user to think about fire safety
  o The context of the game is that it is intended to be used as supplementary
    material for firemen when they are teaching children about fire safety

• Play the game
  o Observe the player as he is playing
  o Answer any questions the player has about how to control the game, as that’s
    not the focus of this study
  o Do not answer questions about what the player is supposed to do, except
    “explore the building” if applicable.
  o Stop the player and ask questions at the end of levels 3 and 4
    ▪ Is anyone left inside the building?
    ▪ If so, do you remember where?
    ▪ If the player failed a level: What happened? Why did the level end?
    ▪ There is no need to record the answer the player gives, as the point of
      the questions is to get the player to think about them.

• Questionnaires after the game
  o Remember not to watch as the player answers the questionnaires. The feeling
    of being watched may result in dishonest answers if the subject interprets it as
    a lack of anonymity.
Appendix C - Test results

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>All students</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>66.55</td>
<td>86.54</td>
</tr>
<tr>
<td>Variance</td>
<td>283.22</td>
<td>145.12</td>
</tr>
<tr>
<td>Observations</td>
<td>36</td>
<td>13</td>
</tr>
<tr>
<td>df</td>
<td>35</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>1.95</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>2.44</td>
<td></td>
</tr>
</tbody>
</table>

Result: F<F-Crit => Can't reject the null hypothesis that the variances are equal

T-test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>All students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86.54</td>
<td>66.55</td>
</tr>
<tr>
<td>Variance</td>
<td>145.12</td>
<td>283.22</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td>Paired variance</td>
<td>247.96</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>3.92</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.68</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.01</td>
<td></td>
</tr>
</tbody>
</table>

Result: t Stat > t Crit => Reject the null hypothesis that the means are equal

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students who played the game</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>66.98</td>
<td>86.54</td>
</tr>
<tr>
<td>Variance</td>
<td>323.08</td>
<td>145.12</td>
</tr>
<tr>
<td>Observations</td>
<td>27</td>
<td>13</td>
</tr>
<tr>
<td>df</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>2.23</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>2.49</td>
<td></td>
</tr>
</tbody>
</table>

Result: F<F-Crit => Can't reject the null hypothesis that the variances are equal
t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students who played the game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86.54</td>
<td>66.98</td>
</tr>
<tr>
<td>Variance</td>
<td>145.12</td>
<td>323.08</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>27</td>
</tr>
<tr>
<td>Paired variance</td>
<td>266.89</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>3.55</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.0005</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.69</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.02</td>
<td></td>
</tr>
</tbody>
</table>

Result: t Stat > t Crit => Reject the null hypothesis that the means are equal

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students who didn’t see the game</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>75.00</td>
<td>86.54</td>
</tr>
<tr>
<td>Variance</td>
<td>156.25</td>
<td>145.12</td>
</tr>
<tr>
<td>Observations</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>df</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>3.89</td>
<td></td>
</tr>
</tbody>
</table>

Result: F<F-Crit => Can’t reject the null hypothesis that the variances are equal

t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students who didn’t see the game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86.54</td>
<td>75.00</td>
</tr>
<tr>
<td>Variance</td>
<td>145.12</td>
<td>156.25</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Paired variance</td>
<td>146.71</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.76</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.14</td>
<td></td>
</tr>
</tbody>
</table>

Result: t Stat < t Crit => Can’t reject the null hypothesis that the means are equal
F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students who watched the game</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>60,42</td>
<td>86,54</td>
</tr>
<tr>
<td>Variance</td>
<td>151,04</td>
<td>145,12</td>
</tr>
<tr>
<td>Observations</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Df</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>1,04</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0,44</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>3,11</td>
<td></td>
</tr>
</tbody>
</table>

Result F<F-Crit => Can't reject the null hypothesis that the variances are equal

t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students who watched the game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86,54</td>
<td>60,42</td>
</tr>
<tr>
<td>Variance</td>
<td>145,12</td>
<td>151,04</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Paired variance</td>
<td>146,86</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Df</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>4,37</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0,0002</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1,74</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0,0004</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2,11</td>
<td></td>
</tr>
</tbody>
</table>

Result t Stat > t Crit => Reject the null hypothesis that the means are equal

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students with feedback</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>67,31</td>
<td>86,54</td>
</tr>
<tr>
<td>Variance</td>
<td>266,24</td>
<td>145,12</td>
</tr>
<tr>
<td>Observations</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>Df</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>1,83</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0,14</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>2,50</td>
<td></td>
</tr>
</tbody>
</table>

Result F<F-Crit => Can't reject the null hypothesis that the variances are equal
t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students with feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86.54</td>
<td>67.31</td>
</tr>
<tr>
<td>Variance</td>
<td>145.12</td>
<td>266.24</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>Paired variance</td>
<td>226.96</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Df</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>3.76</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.0003</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.69</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.0006</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.03</td>
<td></td>
</tr>
</tbody>
</table>

Result: t Stat > t Crit => Reject the null hypothesis that the means are equal

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students who played with feedback</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>69.05</td>
<td>86.54</td>
</tr>
<tr>
<td>Variance</td>
<td>278.77</td>
<td>145.12</td>
</tr>
<tr>
<td>Observations</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>Df</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>1.92</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>2.54</td>
<td></td>
</tr>
</tbody>
</table>

Result: F<F-Crit => Can't reject the null hypothesis that the variances are equal

t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students who played with feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86.54</td>
<td>69.05</td>
</tr>
<tr>
<td>Variance</td>
<td>145.12</td>
<td>278.77</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>Paired variance</td>
<td>228.65</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Df</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>3.28</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.69</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.0025</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.04</td>
<td></td>
</tr>
</tbody>
</table>

Result: t Stat > t Crit => Reject the null hypothesis that the means are equal
F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students who saw with feedback</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>60,00</td>
<td>86,54</td>
</tr>
<tr>
<td>Variance</td>
<td>187,50</td>
<td>145,12</td>
</tr>
<tr>
<td>Observations</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>df</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>1,29</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0,33</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>3,26</td>
<td></td>
</tr>
</tbody>
</table>

Result: F<F-Crit => Can't reject the null hypothesis that the variances are equal

\[t\]-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students who saw with feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86,54</td>
<td>60,00</td>
</tr>
<tr>
<td>Variance</td>
<td>145,12</td>
<td>187,50</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>Paired variance</td>
<td>155,72</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>4,04</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0,0005</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1,75</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0,0009</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2,12</td>
<td></td>
</tr>
</tbody>
</table>

Result: t Stat > t Crit => Reject the null hypothesis that the means are equal

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students who played without feedback</th>
<th>Firefighters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>59,72</td>
<td>86,54</td>
</tr>
<tr>
<td>Variance</td>
<td>483,80</td>
<td>145,12</td>
</tr>
<tr>
<td>Observations</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>df</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>3,33</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0,04</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>3,11</td>
<td></td>
</tr>
</tbody>
</table>

Result: F>F-Crit => Must reject the null hypothesis that the variances are equal
t-Test: Two Sample Assuming Unequal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters</th>
<th>Students who played without feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>86,54</td>
<td>59,72</td>
</tr>
<tr>
<td>Variance</td>
<td>145,12</td>
<td>483,80</td>
</tr>
<tr>
<td>Observations</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>2,80</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1,94</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.45</td>
<td></td>
</tr>
</tbody>
</table>

Result  
\( t \text{ Stat} > t \text{ Crit} \Rightarrow \text{Reject the null hypothesis that the means are equal} \)

F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters (high G.L.)</th>
<th>Firefighters (low G.L.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>81,67</td>
<td>89,58</td>
</tr>
<tr>
<td>Variance</td>
<td>152,78</td>
<td>133,93</td>
</tr>
<tr>
<td>Observations</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>df</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>F</td>
<td>1,14</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>4.12</td>
<td></td>
</tr>
</tbody>
</table>

Result  
\( F<F-\text{Crit} \Rightarrow \text{Can't reject the null hypothesis that the variances are equal} \)

t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Firefighters (low G.L.)</th>
<th>Firefighters (high G.L.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>89,58</td>
<td>81,67</td>
</tr>
<tr>
<td>Variance</td>
<td>133,93</td>
<td>152,78</td>
</tr>
<tr>
<td>Observations</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Paired variance</td>
<td>140,78</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>1,17</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.80</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.20</td>
<td></td>
</tr>
</tbody>
</table>

Result  
\( t \text{ Crit} < t \text{ Stat} < t \text{ Crit} \Rightarrow \text{Can't reject the null hypothesis that the means are equal} \)
### F-test: Two-Sample for Variances

<table>
<thead>
<tr>
<th></th>
<th>Students (high G.L.)</th>
<th>Students (low G.L.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>70.33</td>
<td>57.95</td>
</tr>
<tr>
<td>Variance</td>
<td>279.69</td>
<td>202.97</td>
</tr>
<tr>
<td>Observations</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>df</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>F</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td>P(F&lt;=f) one-tail</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>F Critical one-tail</td>
<td>2.74</td>
<td></td>
</tr>
</tbody>
</table>

**Result**: $F < F_{\text{Crit}} \Rightarrow$ Can't reject the null hypothesis that the variances are equal

### t-Test: Two Sample Assuming Equal Variances

<table>
<thead>
<tr>
<th></th>
<th>Students (low G.L.)</th>
<th>Students (high G.L.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>57.95</td>
<td>70.33</td>
</tr>
<tr>
<td>Variance</td>
<td>202.97</td>
<td>279.69</td>
</tr>
<tr>
<td>Observations</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td>Paired variance</td>
<td>257.12</td>
<td></td>
</tr>
<tr>
<td>Hypothesized mean difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>df</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>t Stat</td>
<td>-2.13</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tail</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>t Critical one-tail</td>
<td>1.69</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tail</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>t Critical two-tail</td>
<td>2.03</td>
<td></td>
</tr>
</tbody>
</table>

**Result**: $t_{\text{Stat}} < -t_{\text{Crit}} < t_{\text{Crit}} \Rightarrow$ Must reject the null hypothesis that the means are equal
Appendix D - Questionnaires

Construct validation questionnaire – English version

Did you play the game or watch someone else play it?
   I played the game
   I watched someone else play
   I did not see the game

Please write your tester ID. This is only to match the results with the pre-test, so if you did not participate in the pre-test then you should not write anything here.

Did you see any of the following pictures while playing the game?
   I saw the pictures
   I did not see any of these pictures

How old are you? *
   9 or younger
   10-11
   12-13
   14-16
   17-20
   21-30
   31-40
   41-50
   51-60
   61 or older

What is your gender? *
   Male
   Female
   Prefer not to say
   Other:

How used to playing mobile games are you? *
   I have never played a mobile game before
   I play mobile games about once a year or less
   I play mobile games a few times a year
   I play mobile games a few times a month
   I play mobile games a few times a week
   I play mobile games almost every day
How used to playing console games (XBOX, PS4 etc.) are you? *
   I have never played a console game before
   I play console games about once a year or less
   I play console games a few times a year
   I play console games a few times a month
   I play console games a few times a week
   I play console games almost every day

How used to playing computer games (PC, Mac etc.) are you? *
   I have never played a computer game before
   I play computer games about once a year or less
   I play computer games a few times a year
   I play computer games a few times a month
   I play computer games a few times a week
   I play computer games almost every day

If the fire alarm goes off and you don’t see a fire you should *
   Gather your belongings first and then evacuate
   Evacuate immediately to the nearest emergency exit
   Wait it out, it might be a false alarm
   Find someone of authority and ask them what to do

There is a fire blocking your path. What’s the first thing you should do? *
   Just run through. It’s safe if you’re fast enough.
   Try to find another way out
   Hide and wait for rescue
   Try to put out the fire

How do you know where to find the nearest emergency exit? *
   Look for the evacuation signs
   Follow the sound of the alarm bell
   Follow the other people
   All of the above

What should you do if you notice an imminent potential fire hazard? *
   Alert the person responsible for it
   If you have an opportunity to do something about it, you should remove the risk
   Tell someone else about it, for example a teacher or a janitor
   All of the above

If you are unable to evacuate a burning building, what should you do? *
   You should open the doors and windows to let the smoke out
   Stay low under the smoke
   All of the above
   None of the above
When should you trigger the fire alarm (if it hasn't already gone off automatically)? *
   As soon as you notice fire
   As soon as you notice smoke
   As soon as you hear someone shout “FIRE!”
   All of the above

How can you stop or slow the spreading of smoke in a room? *
   Close doors and put wet rags around the edges
   Open a window to let the smoke out
   Using a fire extinguisher
   All of the above

If there’s a fire emergency you should do the following in this order *
   Rescue, extinguish, alert and alarm
   Rescue, alert, alarm and extinguish
   Extinguish, alert, alarm and rescue
   Alert, alarm, extinguish and rescue

Did you play (or see someone else play) a level with an elevator? *
   Yes
   No

In case of a fire emergency, should you use the elevator or the staircase to get out? *
   I should never use the elevator, unless it is clearly labelled as being designed for evacuation in case of fire
   I can use the elevator if the staircase is on fire and there is no other way down
   I should use the elevator if it is closer to me than the staircase
   I should only use the elevator if I know the fire is far away from me in the building

How can you best be prepared for fire emergencies? *
   Practice evacuation
   Make sure to have a plan (sketched)
   Memorize the location of the emergency exits
   All of the above

You have safely evacuated the premises and the rescue service arrives, is there anything else you should do? *
   If you remember seeing a person left in the building, you should tell them where you think they are
   If you think you know how many are left inside, you should let them know
   If you saw the fire, you should tell them where it is and what kind of fire it is
   All of the above

You come across a person who seems lost and confused during a fire emergency. What should you do? *
   Ignore them and save yourself
   Guide them to the nearest exit and follow them out
   Tell them where the exit is, then run out
   None of the above
Construct validation questionnaire – Swedish version

Spelade du spelet eller tittade du på när någon spelade? *
   Jag spelade spelet
   Jag tittade på när någon annan spelade
   Jag såg inte spelet

Vänligen skriv ditt namn. Detta är bara för att matcha resultaten med förprovet. *

Såg du några av dessa bilder medan du spelade spelet? *
   Jag såg bilderna
   Jag såg inte några av dessa bilder

Hur gammal är du? *
   9 eller yngre
   10-11
   12-13
   14-16
   17-20
   21-30
   31-40
   41-50
   51-60
   61 eller äldre

Vilket är ditt kön? *
   man
   kvinna
   föredrar att inte säga
   Other:

Hur van vid att spela mobilspel är du? *
   Jag spelar aldrig mobilspel
   Jag spelar mobilspel ungefär en gång per år eller mindre
   Jag spelar mobilspel ett par gånger om året
   Jag spelar mobilspel några gånger i månaden
   Jag spelar mobilspel några gånger i veckan
   Jag spelar mobilspel nästan varje dag

Hur van vid att spela konsolspel (XBOX, Playstation, Nintendo etc.) är du? *
   Jag spelar aldrig konsolspel
   Jag spelar konsolspel ungefär en gång per år eller mindre
   Jag spelar konsolspel ett par gånger om året
   Jag spelar konsolspel några gånger i månaden
   Jag spelar konsolspel några gånger i veckan
   Jag spelar konsolspel nästan varje dag
Hur van vid att spela datorspel (PC, MAC etc.) är du? *
Jag spelar aldrig datorspel
Jag spelar datorspel ungefär en gång per år eller mindre
Jag spelar datorspel ett par gånger om året
Jag spelar datorspel några gånger i månaden
Jag spelar datorspel några gånger i veckan
Jag spelar datorspel nästan varje dag

Om brandlarmet börjar tjuta och du inte ser eld då ska du *
Uttrym direkt och gå till närmaste nödutgång
Vänta till det slutar, det kanske är ett falskt alarm
Hämta dina saker först och evakuera
Leta efter en auktoritär person och fråga hen vad man ska göra

Om en eld blockerar din väg, vad ska du då göra? *
Spring igenom elden, det är säkert om du är snabb nog
Försök att hitta en annan väg ut
Göm och vänta
Försök att släcka elden

Hur kan du bäst hitta närmaste nödutgång? *
Leta efter nödutgångsskyltar
Följ ljudet från brandlarmet
Gå dit där alla andra går
Alla av ovanstående

Vad ska du göra ifall du lägger märke till en omedelbar möjlig brandrisk? *
Säg till personen som är ansvarig
Om du har möjlighet att göra någonting åt situationen, så ska du ta bort risken
Säg till någon annan t.ex. lärare eller vaktmästare på skolan
Alla av ovanstående

Om du inte kan utrymma en brinnande byggnad, vad ska du göra? *
Öppna dörrar och fönster för att släppa ut röken
Håll dig lågt under röken
Alla av ovanstående
Inget av ovanstående

När ska du sätta igång brandlarmet, om det inte redan är igång? *
Så fort du märker av eld
Så fort du märker av rök
Så fort du hör någon ropa “det brinner!”
Alla av ovanstående

Hur kan du hindra spridningen av rök in till ett rum? *
Stäng dörrar och lägg blöta trasor runt springorna
Öppna ett fönster och låt röken gå ut
Använd en brandsläckare
Alla av ovanstående
Om det uppstår en brandfara ska du agera i följande ordning *
rädda, varna, larma och släck
släck, varna, larma och rädda
varna, larma, släck och rädda
rädda, släck, varna och larma

Har du spelat (eller tittat på när någon spelade) en nivå som har en hiss? *
Ja
Nej

I en nödsituation, ska du använda hissen eller trappan för att utrymma platsen? *
Jag ska använda hissen om den är närmast mig
Jag ska aldrig använda hissen (om det inte är en uttalad och uppskyltad hiss som är avsedd för evakuering vid brand)
Jag ska endast använda hissen om jag vet om att branden är långt borta från mig
Jag kan använda hissen om branden är i trappen

Hur kan man förbereda sig på bästa sätt om en brand skulle hända? *
Öva utrymning
Se till att ha en plan (ritad som skiss)
Lär dig platserna för nödutgångarna
Alla av ovanstående

Du har utrymt byggnaden och räddningstjänsten ankommer, finns det någonting mer du ska göra? *
Om du kommer ihåg att det finns en person kvar i byggnaden så ska du säga till vart någonstans i byggnaden du tror att de är
Om du vet hur många som finns kvar i byggnaden så ska du säga till
Om du såg elden så ska du säga till vart den finns och vilken typ av brand det är
Alla av ovanstående

Under en nödsituation passerar du en person som verkar förvirrad. Vad ska du göra? *
Ignorera dem och rädda dig själv
Hjälp dem att hitta till närmsta nödutgång och följ dem ut
Säg till dem vart utgången är och spring sedan ut
Inget av ovanstående
Content validation questionnaire – English version

What lessons are represented in the game? Identify and list them (you can press enter) *

Here is a list of the lessons we tried to represent in our game.

- Getting yourself to safety
- Not walking into a fire
- Noticing the evacuation signs to know where to evacuate
- Noticing potential fire hazards
- Not using elevator in case of fire
- Crouching when in smoke
- Alarming when upon noticing a fire
- Not pressing the alarm when there's no apparent fire emergency
- Remembering who is left in the building
- Remembering on which floor the remaining people are
- Guiding people out
- Closing doors to prevent fire and smoke spread

The lesson "Getting yourself to safety" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Not walking into a fire" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Noticing the evacuation signs to know where to evacuate" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Noticing potential fire hazards" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Not using elevator in case of fire" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Crouching when in smoke" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Alarming when upon noticing a fire" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Not pressing the alarm when there's no apparent fire emergency" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree
The lesson "Remembering who is left in the building" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Remembering on which floor the remaining people are" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Guiding people out" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

The lesson "Closing doors to prevent fire and smoke spread" is well represented in the game *
Disagree 1 2 3 4 5 6 Agree

Is the game missing any critical lessons?

Do you have any comments on the representation of the lessons?

How suitable do you think this game is for children (age 10-16)? *
Not suitable 1 2 3 4 5 6 Very suitable

How suitable do you think this game is for adults? *
Not suitable 1 2 3 4 5 6 Very suitable

Do you think this game is a useful tool for fire safety education? *
Content validation questionnaire – Swedish version

Vilka lärdomar finns representerade i spelet? Identifiera och lista dem (du kan trycka på enter)

Här är en lista över de lärdomar vi försökt att representera i vårt spel.

- Ta dig till säkerhet
- Gå inte in i en eld
- Var uppmärksam på utrymningsskyltarna för att veta vart du ska utrymma
- Var uppmärksamt på potentiella brand faror
- Inte använda hissen vid en brand
- Hålla sig lågt under rök
- Sätta igång brandlarmet när man märker en utbruten eld
- Tryck inte igång brandlarmet när det inte är ett nödläge
- Komma ihåg vem som är kvar i byggnaden
- Komma ihåg i vilken våning de som är kvar i byggnaden är
- Vägled andra ut (vid nödläge)
- Stänga dörrar för att förhindra spridning av eld

Lektionen "Ta dig till säkerhet" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Gå inte in i en eld" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Var uppmärksam på utrymningsskyltarna för att veta vart du ska utrymma" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Var uppmärksamt på potentiella brand faror" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Inte använda hissen vid en brand" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Hålla sig lågt under rök" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Sätta igång brandlarmet när man märker en utbruten eld" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Tryck inte igång brandlarmet när det inte är ett nödläge" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer
Lektionen "Komma ihåg vem som är kvar i byggnaden" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Komma ihåg i vilken våning de som är kvar i byggnaden är" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Vägled andra ut (vid nödläge)" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Lektionen "Stänga dörrar för att förhindra spridning av eld" är väl representerat i spelet *
Instämmer inte 1 2 3 4 5 6 Instämmer

Saknade spelet några kritiska lektioner?

Har du några kommentarer på representationen av lektionerna?

Hur lämpligt tycker du att spelet är för barn (ålder 10-16)? *
Inte lämpligt 1 2 3 4 5 6 Mycket lämpligt

Hur lämpligt tycker du att spelet är för vuxna? *
Inte lämpligt 1 2 3 4 5 6 Mycket lämpligt

Tycker du att detta spel är användbart för att lära ut brandsäkerhet? *
Appendix E - Permission slip

Skövde 2017-04-29

INFORMATION AVSEENDE DELTAGANDE I FÖRKNINGSSTUDIE
Projekt för inkluderande speldesign

Mitt namn är Anita Stenholm och jag är en magisterstudent på Högskolan i Skövde. Vi är två studenter som gör vår magisterexamen tillsammans och vi skulle behöva några elever (runt 11 år) att testa vårt spel på.

Spelet är utvecklat tillsammans med en klient på Räddningstjänsten StorGöteborg och tanken är att spelet ska lära ut barn om brandsäkerhet samt hur man evakuerar en brinnande byggnad. Målet med spelet är att det ska kunna fungera som ett komplement till den undervisning brandmän ger 5:e klassare om brandsäkerhet på brandstationen eller i klassrummet.

Denna undersökning syftar till att utvärdera ett spel. Vi kommer att observera deltagaren när denne spelar spelet för att analysera om spelet fungerar på det sätt vi har tänkt. Ingen beklöver kommer att göras på spelarens förmåga att spela spelet.

Spelet innehåller inget stötande material.

Deltagandet är helt frivilligt och kan när som helst avbrytas utan att deltagaren behöver förklara varför. Vi försäkrar att ingen information om deltagaren kan komma att kopplas till dennes identitet. Medverkan är anonym.

Vi förväntar oss att varje testtillfälle tar ca 30 min.

För din medverkan behöver vi ett formellt medgivande från dig att du har tagit del av och accepterat villkoren för undersökningen.

Om du har några frågor, vänta pleased kontakta:
Anita Stenholm, Student. a16anist@student.his.se, 076-0150347

Jag godkänner ovanstående villkor

Ort & Datum

Namn (målsmans underskrift krävs för testpersoner under 15 år)
Appendix F - Reflective feedback

The following pictures are displayed when the player encounters a failure state in the game.

Figure 11: This picture indicates that the players should trigger the alarm only when they can see fire or smoke.

Figure 12: This picture indicates that the player should not walk too close to the fire.

Figure 13: This picture indicates that the player should crouch under smoke.
Appendix G - Screenshots of E.V.A.C.

Figure 14: Level 1.

Figure 15: Level 2.
Figure 16: Level 2 – The player encounters smoke for the first time.

Figure 17: Level 2 - The player steps too close to the fire.
Figure 18: Level 2 – The player triggers a fire alarm.

Figure 19: Level 2 completed.
Figure 20: Level 3.

Figure 21: Level 3 completed. The player can see which characters are outside.
Figure 22: Level 4.

Figure 23: Level 4 complete. There are a lot more characters to remember this time.
Figure 24: Level 4 complete - alternate exit.