INTERNATIONAL DESIGN CONFERENCE - DESIGN 2014
Dubrovnik - Croatia, May 19 - 22, 2014.

# ISSUE OF USING STUDENTS AS TEST USERS – SOME FINDINGS IN VE TESTING

T. Tiainen and A. Ellman

*Keywords: user testing, virtual environment, virtual prototypes, students*

## 1. Introduction

Virtual prototype is a relatively recent practice used in various industrial domains, which aims at exploring a product that does not exist in reality yet. This practice can be used for evaluating the products. Generally this is done in virtual environment (VE) with user test [Kalawsky 1993] in which test users evaluate the visual image of products [Yang 2005]. As VE technology has developed, also other features of a product can be evaluated; such as its functional features, as well as its ergonomic and usability aspects [Bordegoni and Ferise 2013]. Next step might be virtual engineering which means that design and validation activities occur collaboraty in order to prove early product design, support decision making and enable continuous product optimisation [Ovtcharove 2010].

In all of the alternative ways to use virtual prototypes, the researchers assess user test as the main method to evaluate the virtual prototypes. Typically, the researchers take students as test users without pondering and discussing if the students differ significantly from the actual product users (e.g. older consumers or machine drivers). This though-model is seen, for example, in [Schwerdtfeger et al. 2011], in which the test was done by 34 users, half were students and half were citizens; however this study does not compare actions of these two groups.

When students are used as the test uses, it includes the assumption that students act similarly in the user tests than those people who are actual stakeholders of the experiment. In this paper we consider if the assumption is valid. We discuss the topic based on three VE test uses in which we compare student users to actual stakeholders of the test, which are consumers and machine drivers. The results show that in some cases students act differently in VE than the actual stakeholders of the test; especially when the user needs special skills to execute the task with virtual prototype, the students differed from actual users.

This paper is based on three empirical studies, in which test users are taken into 3D cave-like VEs. We first describe the test environment used. Then we outline the test results and compare student and actual users' actions in the VE. Finally we summarize the findings and discuss the possible field of application for students as test user.

## 2. Evaluation of product size – test 1

In test 1 we created a setup for studying how persons evaluated the size of products. We compare test users' evaluations of virtual products to their evaluations of physical products, the test setup also includes two parts: first, a VE for presenting virtual products, and second, a set up for evaluation of physical products.

### 2.1 Test 1: VE and test setup

A walk-in cave-like virtual environment was used in test 1 and 2. The used one is almost a cubic space consisting of five rear projection surfaces; three walls, floor and ceiling. The dimensions of the space are 3 x 3 meters with a height of 2.4 meters. The display resolution of each wall is 1280 x 1024 pixels, enabling the user to sense a view of up to 6.55 Million pixels. The users' view is rendered according his/her position and orientation using a magnetic tracking system. An active stereo image is observed through liquid crystal glasses with frequency 2 x 45 Hz. An ordinary Wand input device is used for controlling movements.

A shopping centre to which a user could walk in was used in the test 1 and 2. The layout of the shopping centre is presented in Figure 1. The dimensions of the virtual shops were from 3.0 to 4.3 meters to both directions. As the physical space is smaller, just 3 x 3 meters, a locomotion controlling device was needed.

**Figure 1. The room for evaluation of the virtual products**

The target in the first test use was to compare users' evaluation of the sizes of virtual and physical products. There were five product which sizes (height, length and amplitude) the users were asked to evaluate. The products were a coffee cup, a flower vase, a milk can, a cereal bowl and a large plant pot shown in Figure 2.

**Figure 2. The five products under size evaluation (a coffee cup, flower vase, milk can, cereal bowl and plant pot)**

Two groups of test users visited in the VE and evaluated the sizes of the products. At first they evaluated the virtual products which were located on a virtual table. The test users could walk around the table and watch the products from different perspectives, also from above by bowing down. Data of test users are presented in Table 1.

**Table 1. The test users of size evaluation**

| Group | average age | number of female | number of male | total |
|---|---|---|---|---|
| Students | 27 | 5 | 13 | 18 |
| Consumers | 45 | 9 | 11 | 20 |

## 2.2 Results of test 1

The test users evaluated two versions (a virtual and a physical one) of five products; all of them were evaluated by three attributes (height, length and amplitude). So each test user gave 30 values: the averages of them by the two groups are presented in Figure 3. The figure already gives an impression that the two groups got similar results in evaluating the product sizes.

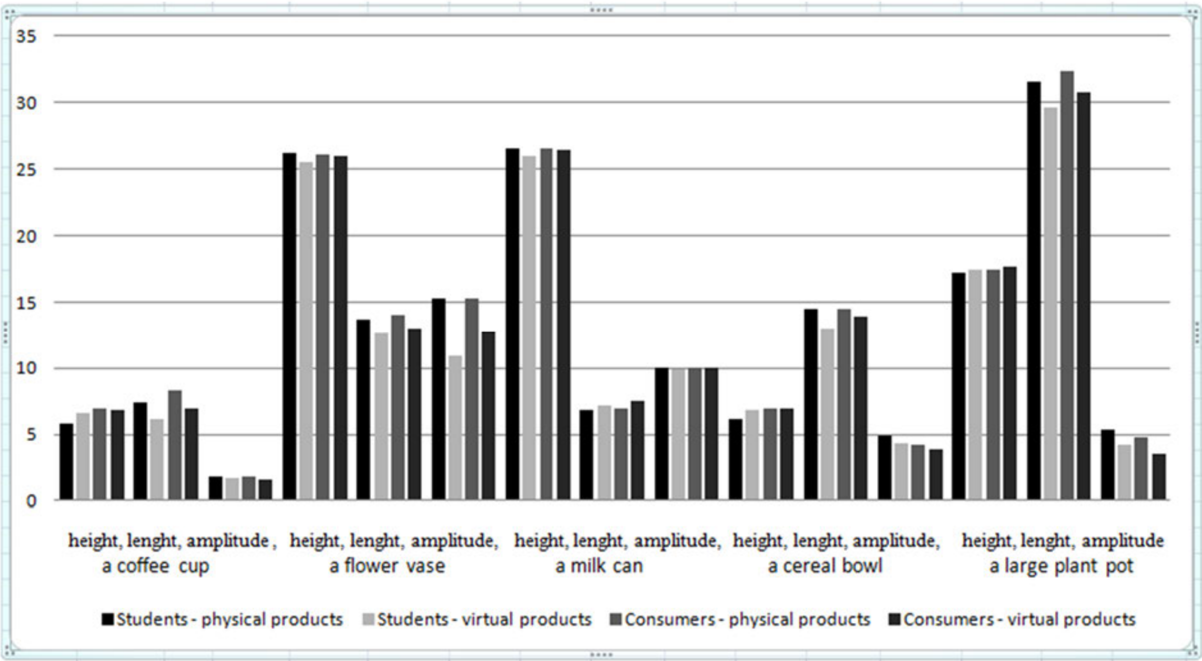


**Figure 3. Users' evaluation of the product size (average per group)**

We analysed if the individual test user's evaluated the virtual product similar than the physical product. This was done with calculating the correlation between user's estimated value for virtual and physical product. In the most cases there was a clear correlation between them: in students group, nine cases, and in consumers, 13. In some cases, there were low or none correlation: in students group, five case, and in consumers, one.

**Table 2. Correlation between size evaluation of virtual and physical products**

| | Student group | | | Consumer group | | |
|---|---|---|---|---|---|---|
| correlation | under 0.3 | 0.3 – 0.6 | over 0.6 | under 0.3 | 0.3 – 0.6 | over 0.6 |
| height | 2 | - | 3 | - | - | 5 |
| length | 1 | 1 | 3 | 1 | 1 | 3 |
| amplitude | 2 | - | 3 | - | - | 5 |

We continued by making T-test for analysing if the student test group differed from the consumer test group. The T-test was calculated by all evaluation tasks and none of them showed differences between the groups.

Based on above analysis and results, we state that both students and consumers evaluated similar product sizes when it is compared the results of evaluation virtual and physical products. So in this kind of cases it is acceptable to use students as test users.

## 3. Locomotion in VE – test 2

### 3.1 Test 2: VE and test setup

The second test focused on how users move in a VE. The possible alternatives of locomotion are physical changing of the user position, as taking some steps, and changing position in virtual imagine with a device. In this test the research question was how VE users choose to move in a VE when they can choose it by themselves. This test is also described in [Tiainen et al. 2007] and [Tiainen et al. 2013].

The testing situation was a small shopping centre in VE, shown in Figure 4. The centre included three small shops in which the test users visited in a planned order. In the first one, the test user was guided in and then s/he could watched the virtual objects by moving physically, as taking steps and turning around. In the second one, an operator moved the distance and perspective by using a device Wanda, shown also in Figure 4. The user could say what s/he wanted to observe and the operator moved the object closer. In the third shop, the user moved her/himself as s/he wanted.

35 test users participated to the test; half of them were students and half consumers. The groups are introduced in Table 3.

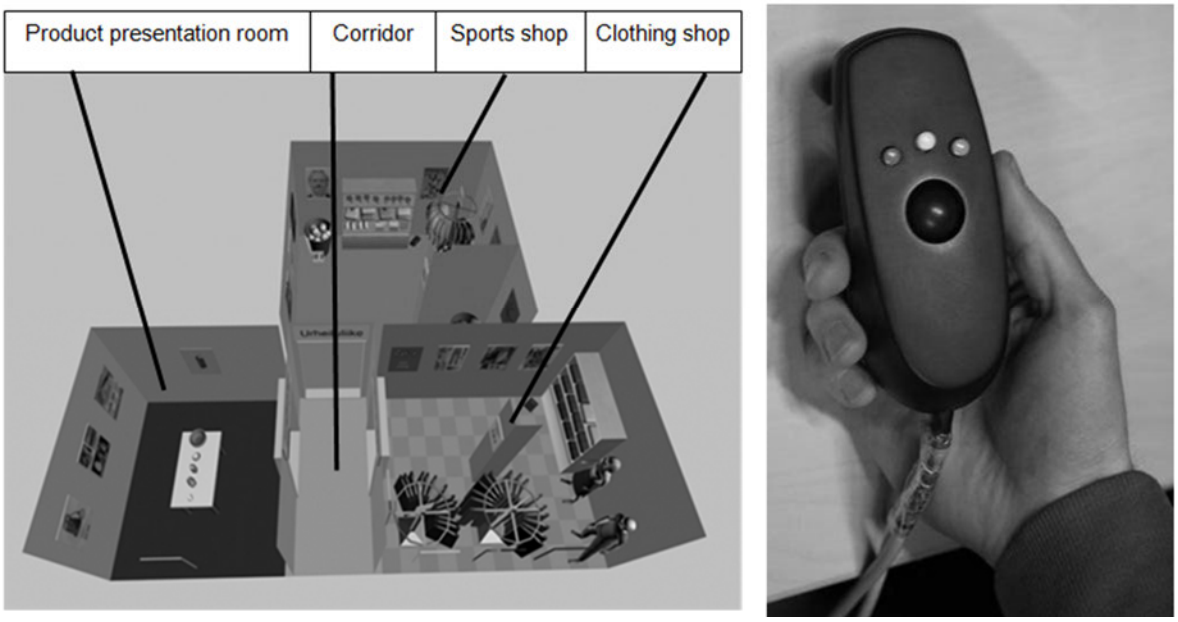


**Figure 4. The virtual shopping centre with three shops and locomotion control device Wanda**

**Table 3. The users of the test of moving in VE**

| Group | average age | number of female | number of male | Total |
|---|---|---|---|---|
| Students | 26 | 5 | 12 | 17 |
| Consumers | 45 | 9 | 11 | 18 |

### 3.2 Results of test 2

The analysis focused on understanding how users move in VE when they can choose it by themselves. Three features were analysed:

1. How easy did the users find the using of Wanda (the device for virtual locomotion in the VE)?
2. How many steps did the users took (called Walk)?
3. How many times did the users zoomed and rotated with Wanda (Wanda use)?

The test use material was analysed if the groups of students and consumers differ from each other. It was done by three statistical measurements. The average and standart deviation of all three analyses elements were different in the groups of students and consumers: Students found Wanda using easier and used it more, instead, consumers walked more, which is seen in Figure 5 and Table 4.
The analysis was continued by calculating the T-test, which shows if the differences between the groups of students and consumers are statistical remarkable. The T-test values of each element were compared to the limit of 95 per cent probability that the difference is not random (values in Table 4). Based on the results the groups of students and consumers are clearly different by finding Wanda use easy and by walking in the VE. The number of Wanda using between the two groups is also less different , its T-value value (1.41) is not meaningful with 95 per cent limit (1.66) but meaningful with lower 90 per cent limit (1.30) indicating some difference between the groups.

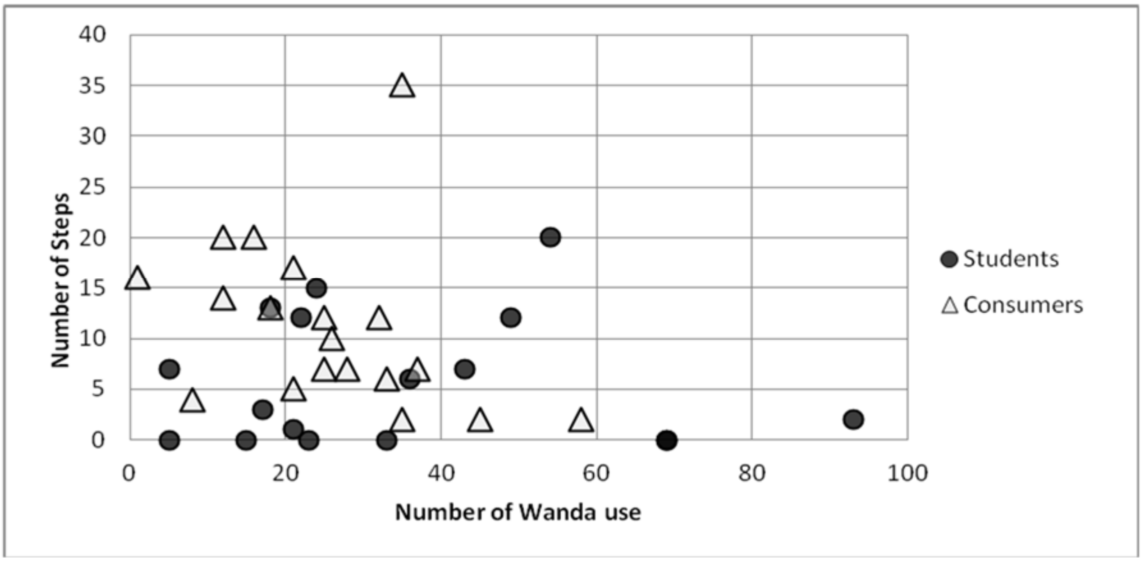


**Figure 5. The test users' moving in the VE**

**Table 4. The results of alternative ways of moving**

| Group | Easiness of using Wanda | | | Walk (number of steps) | | | Number of Wanda use | | |
|---|---|---|---|---|---|---|---|---|---|
| | ave | dev | min - max | ave | dev | min - max | ave | dev | min – max |
| Student | 6.0 | 3.0 | 0 – 9 | 5.8 | 6.5 | 0 – 20 | 35.1 | 24.7 | 5 – 93 |
| Consumers | 3.1 | 3.1 | -2 – 9 | 11.1 | 8.2 | 2 – 35 | 25.7 | 13.6 | 1 – 58 |
| T-test | t(10) = 3.41<br>95 % limit: 1.81<br>90 % limit: 1.37 | | | t(34) = 2.15<br>95 % limit: 1.70<br>90 % limit: 1.31 | | | t(91) = 1.41<br>95 % limit: 1.66<br>90 % limit: 1.30 | | |

Based on test 2 we state that students' behaviour differ from consumers' behaviour. Students prefer using a technical device, whereas consumers prefer physical movements, as taking steps. Taking students as test users is suspicious if the test users need to control the virtual objects or the environments.

## 4. Operating with a virtual mine loader machine – test 3

The third test focused on doing work tasks with a virtual heavy moving work machine. In the test the users drove a virtual mine loader and transported rock with it. The test task was based on real mine. The VE included its roads, collecting and releasing areas, as well as the real outfit, size, sounds and movements of the loader. This test situation is described also in [Tiainen et al. 2012].

### 4.1 Test 3: VE and test setup

Virtual mine loader machine was a multisensory game-like setup consisting of physical parts such as driver's chair and vehicle control devices which signals were connected to a real-time simulator. This enabled computation of machine movements, behaviour of rocks etc. Furthermore, there was a motion platform, so that the driver could feel the bumps, speedups and stops. Also aurial feedback from state of machine and events such as collisions to walls were given audio system of the environment. The image from the cabin and from virtual environment is presented in Figure 6.
The test included first a practising part, so that the users became familiar with the actions of a loader. The second part of the test was a rock transporting task. In it the user drove the loader down to the rock pile, fill the bucket with rock, lift the bucket up, drive the same road back up and empty the bucket to a releasing spot. The amount of collected rock and transported rock was measured.
The task was done by twelve test users; five of them were students of technical university and seven were from an vocational institute which educates drivers for heavy machines. All the users were male. The groups are introduced in Table 5.

**Figure 6. The view from the virtual cocpit of the mine loader machine and VE of the test 3**

**Table 5. The test users of the mine loader machine test**

| Group | average age | number of female | number of male | Total |
|---|---|---|---|---|
| Students | 22 | 0 | 5 | 5 |
| Drivers | 24 | 0 | 7 | 7 |

### 4.2 Results of test 3

The groups' results of transporting the rock were different (see Figure 7 and Table 6). In the students group, only three of five students got some rock into the bucket and two of them didn't get any. One of them lost it all his load during the driving. In the drivers' group, they all got some rock into the bucket, but also there one driver lost all his load during the driving. Mass of an individual virtual rock was 300 kg.
The visual image of the result of transporting task (Figure 7) and average of treated rock (Table 6) outline a clear picture of the differences between students and drivers. Furthermore, we verify the existence of the differences between the groups by T-test. Its results in both loading and transporting task indicates that by over 90 per cent probability the groups differ from each other. As the number of test users is so low, this uncertainty is acceptable.

**Table 6. Average of the loaded and transported rock and results of T-test**

| | | The amount of rock (kg) | |
|---|---|---|---|
| Groups | Statistical test | into bucket | transported |
| Student | Average | 2253 | 1253 |
| Drivers | Average | 4524 | 3669 |

| | | | |
|---|---|---|---|
| | T-test | t(22)=1.48 | t(22)=1.52 |
| | Limit (95%) | 1.72 | 1.72 |
| | Limit (90%) | 1.32 | 1.32 |

Based on this test we state that when the VE test is realistic and the task demands skills, student users' actions differ from real workers' actions too much. In such cases only skillful workers should be used as VE test users for reaching reliable results.

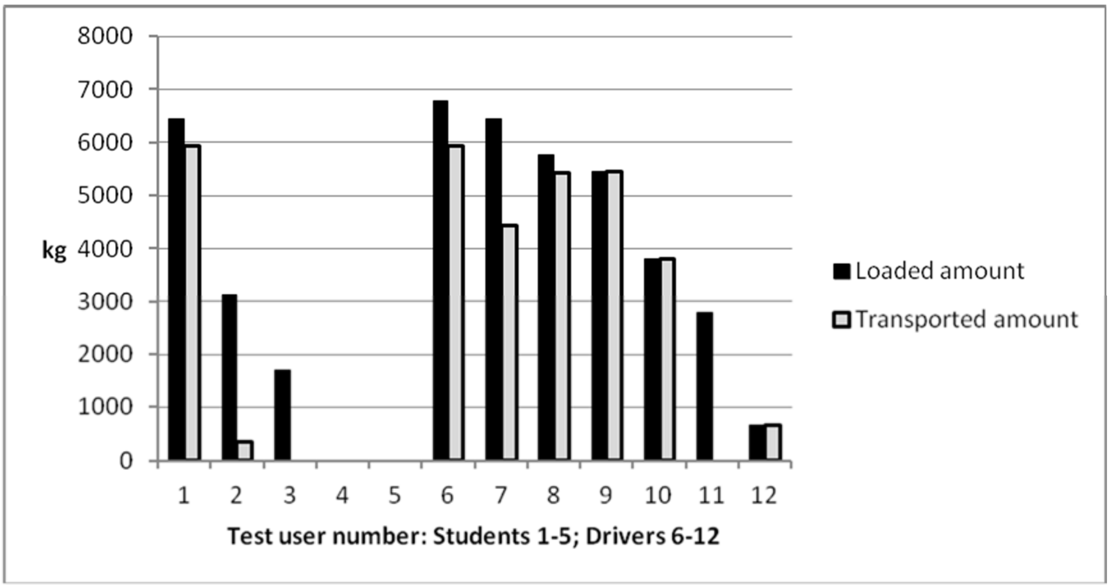


**Figure 7. Users' loaded and transported amount of rocks**

## 5. Discussion

It is common to use students as test users in VE studies which focus on the evaluation of virtual products. In this paper we examined on the basis of that: if students behave in VE similarly than actual product users do. We did this with three test uses in VE.

The results presents that in some cases students actions do not differ from actual users' ones. In those cases it is viable to use students as test users. However, in some cases students' act different than actual users.

In the design context the most common use of VE is to evaluate product prototypes. Students' use is viable when test users' task is to evaluate virtual products (as above in the test 1). In such case student act similar than consumers: both of them evaluated virtual and physical products similarly. However, we only focused the evaluation of product sizes, not if their taste is similar.

When the test users need to locomote in the VE, it is in the border line, if the use of students indead of consumers is viable. As out test 2 presents students prefer the use of technical device, whereas the consumers prefer bodily movements, as taking steps. Our VE test use did not include users locomotion and navigation in a large VE. Based on our user test, we assume that students' actions differ clearly from consumers' actions. However, this assumption needs to be studied in future studies.

Our test 3 uncovers the situation in which students' use is inappropriate. The task in the test 3 was to tranport rock with a mine loader machine. In that case the differences between students' and drivers' coping was clear. In this kind of task users' skills are fundamental. If VE is used for developing or evaluating how a tool fits to work practice, students use as test users is inappropriate. Instead, the test users need to have occupational skill. This is understandable when thinking a familar case: driving a car. If the design of car is unnormal, for exampel, the order of gear in a gear stick is unnormal. When such test users who have no experience of driving normal car, that is not a problem to them. Instead, when test users have strong experiences on driving car, they have learned where different gears are and they assume to find the gears in their normal place.

When new products are designed and evaluated in VE, it is necessary to consider which kind of test users represent probable actual user's of the product. Furthermore, in designing the VE test setup it is needed to consider what kinds of actions (as the ways of locomotion) is familiar to the test users.

**References**

*Bordegoni, M., Ferise, F., "Designing interaction with consumer products in a multisensory virtual reality environment", Virtual and Physical Prototyping, 8(1), 2013, pp. 51-64.*

*Kalawsky, R., "The Science of Virtual Reality and Virtual Environments: A Technical, Scientific and Engineering Reference on Virtual Environments", Addison Wesley Publishing Company, 1993.*

*Ovtcharova, J. G., "Virtual engineering – proinciples, methods and application", Internatioal Design Conference, Dubrovnik, Croatia, 17-20 May, 2010, pp. 1267-1274.*

*Schwerdtfeger, B., Reif, R., Gunthner, W. A., Klinker, G., "Pick-by-vision: there is something to pick at the end of the augmented tunnel", Virtual Reality 15, 2011, pp. 2013-223.*

*Tiainen, T. Kaapu, T., Ellman, A., "Evidence against correlation between ease of use and actual use of a device in a walk-in virtual environment", Human Technology: An Interdisciplinary Journal on Humans in ICT Environments, 9(1), 2013, pp. 56-71.*

*Tiainen, T., Ellman, A., Kaapu, T., Roberts, D., "Effect of Navigation Task on Recalling Content: The Case of Occasional in Restricted, Cavelike Virtual Environment", In DJ Roberts, GK Theodoropoulus, and A El Saddik (Eds.), Proceedings of 11th IEEE / ACM International Symposium on Distributed Simulation and Real-Time Applications (October 22 - 24, 2007, Chania, Crete Island, Greece), IEEE Computer Society, Los Alamitos, California, USA (ISBN 0-7695-3011-6.), 2007, pp. 209-216.*

*Tiainen, T., Kaapu, T., Ellman, A., Kuusisto, J., "A Co-creation Tool in Walk-in Virtual Environment: Making Prospective Work Visible", ECIS2012, Barcelona, Spain, June 11-14, 2012, 13 pages.*

*Yang, M. C., "A study of prototypes, design activity, and design outcome", Design Studies, 26, 2005, pp. 649-669.*

Asko Ellman, Dr.Tech., Professor
Department of Engineering Design, Tampere University of Technology
P.O.Box 589 Tampere, Finland
Telephone: +358 40 7574247
Email: asko.ellman@tut.fi