Measuring involvement with audio/video content

Nele Van den Ende
Eindhoven University of Technology
5600 MB Eindhoven, the Netherlands
n.v.d.ende@tue.nl

Jettie Hoonhout
Philips Research Laboratories
5656 AE Eindhoven, the Netherlands
jettie.hoonhout@philips.com

Lydia Meesters
Eindhoven University of Technology
5600 MB Eindhoven, the Netherlands
l.m.j.meesters@tue.nl

The authors believe that involvement is one of the key cognitive mediating factors between content and viewing experience. A questionnaire to measure involvement with audio/video content was developed. Results showed that 25 questions are probably sufficient to cover the aspects of the construct. . Creating a valid and reliable questionnaire is not an easy process, but the process established in this paper can hopefully contribute to the creation of well-defined constructs, and valid and reliable measurements for these constructs.

Involvement, Factor Analysis, Audio/video Content, Questionnaire

1. INTRODUCTION

Elizabeth is watching a video clip on YouTube. Suddenly the video becomes very blurry and is out of sync with the audio. Elizabeth is annoyed but since she really wonders what is going to happen next, she continues to watch the video and tries to follow what happens as well as possible.

The authors believe that involvement is one of the key cognitive mediating factors between content and viewing experience. Furthermore, it is a construct that allows personalization of applications, such as social media, social TV or recommender systems. However, involvement is not an easy construct to define or measure. The influence of involvement as a mediating factor between content and viewing experience can only be decided upon when there is a definition of involvement and a reliable and valid measurement tool.

Based on previous research [1], six attributes underlying the construct for involvement with audio/video content are identified (see figure 1). The attributes were identified through the process of concept mapping [2]. Concept mapping as applied started with generating statements about involvement through semi-structured interviews. Next, these statements were used in both structured and unstructured card sorting tasks to examine how these statements relate to each other. A hierarchical clustering analysis was then conducted to create a representation of the relation between the statements. To interpret this representation, participants were asked to come in groups, such that the interpretation was done in accordance with several others. One last hierarchical clustering analysis and visualization of the results led to the six attributes as shown in Figure 1.

Figure 1. Clusters (together with their statements) on which the construct involvement with audio/video is based.

How involvement is defined and measured depends on the area one wants to use involvement for, e.g. predicting buying behavior [3], presence detection in 3D television [4] and immersion in video games [5]. In consumer behavior research, an accepted definition of involvement is "a person's perceived relevance of the object based on inherent needs, values and interests" [6, 7]. Zaichowsky [6] proposed the Personal Involvement Inventory (PII), a semantic differential scale. Examples of semantic differentials include: 'important – unimportant', 'irrelevant – relevant', 'uninterested – interested',

and 'undesirable – desirable'. The assumption of the PII is that the level of involvement varies on a bipolar scale, from low to high involvement. However, this research concerned only products (e.g. red wine, colour TV, jeans).

Witmer & Singer [4] developed the Immersive Tendencies Questionnaire (ITQ) intended to measure tendencies of people to become involved in everyday activities and their ability to focus on a specific activity. Additionally, it is unclear whether there are not better ways to measure focus than self-reporting via a questionnaire. Witmer & Singer [4] provide a definition for involvement, from which they developed a subscale with the same description in the ITQ: "Involvement is a psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events. Involvement depends on the degree of significance or meaning that the individual attaches to the stimuli, activities or events. ... Involvement can occur in practically any setting or environment and with regard to a variety of activities or events; however, the amount of involvement will vary according to how well the activities and events attract and hold the observer's attention." Following from this definition, factors necessary for involvement are attention (controlled-deliberate and automatic) and meaningfulness or significance of the stimuli to the observer. While the ITQ predicts presence as measured by Witmer & Singer [4] Presence Questionnaire (PQ), it is clear that the ITQ would not be useful to predict involvement with audio/video content. In particular not when involvement needs to be determined after each audio/video fragment shown, since the ITQ questions are rather general. Another direction to involvement research is presented by Klimmt & Vorderer [8], who argue that experiencing involvement is realized via a perceptual focus on mediated information, while avoiding/suppressing stimuli that are not important for the mediated information. In other words, when somebody is involved with media information, perception, thought and emotion are directed towards the media information as much as possible, while distractors are ignored as much as possible. If involvement can be seen as having low and high ends, two levels could be identified: "... a distant, analytical way of witnessing the events presented by the medium (low involvement) and, in contrast, a fascinated, emotionally and cognitively engaged way of enjoying the media information (high involvement)" [8]. Emotional engagement could also be seen as creating more experience, and analytical engagement as requiring more processing, in which case high involvement could be seen as the processing of media in a fully experiential way. Flow and immersion are other constructs that would be considered as related to immersion, that one often comes across in games research [9, 10].

Sweetser & Wyeth [9] studied immersion, which they defined as 'deep but effortless involvement in a game'. According to them, immersion is expressed by people through a loss of concern for self and everyday life, an altered sense of time, forgetting that players are participating through a medium, making players linger, and drawing players into the narrative with characters, storyline and background. Witmer & Singer [4] define immersion as "... a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences". Although immersion and involvement seem linked, in general watching video material is a passive activity and does not necessarily demand interaction with the content. This characterizes an important difference between involvement and immersion.

So, although several questionnaires are proposed to measure involvement [4, 6], these questionnaires address different situations (e.g. involvement with products [6]), and are not specific enough to measure involvement with audio/video content. A drawback of questionnaires is that the viewers are asked to report their experience afterwards. Behavioural measures, on the other hand, allow direct measurement of the viewer experience. Unfortunately, no behavioural measurements are available to evaluate the involvement construct. Behavioural measurements can be informative about participant behaviour and experience. However, the data is often very noisy and difficult to interpret. Since there is no appropriate self-report measure to assess the construct of involvement, an accurate interpretation of behavioural measures would be quite a challenge. Therefore, the decision was made to develop a questionnaire to measure the construct of involvement with audio/video content.

This paper describes the development of a questionnaire to measure user involvement with audio/video content. The first step in developing a questionnaire for the involvement construct is to make an item pool of questions based on the operational definition [11, 12]. The six attributes (consisting of clusters of coherent words) underlying the involvement construct (see Figure 1) served as operational definition. Additionally, the attributes were used to create fifty items for the involvement questionnaire (from now on called invQ). Section 3 contains a description of the initial invQ items and how the items were tested with cognitive interviewing. The updated item pool was tested in an online survey. The paper ends with a discussion of how the structure of the questionnaire gives further information about the construct of involvement with audio/video content.

2. COGNITIVE INTERVIEWS

Cognitive interviews were employed to test whether the first 50 items are clear, concise, concrete and free of ambiguity [11].

2.1 Method

Four people (two female, two male, mean age 35 years, standard deviation 4.5 years) were invited to participate in the cognitive interview. Participants' backgrounds were in psychology, human technology interaction and computer science.

Participants were asked to read the instructions, and if they had no further questions, were shown five video fragments. The five fragments were chosen to represent five categories: action, sports, animation, documentary and soap-operas. After each video fragment the participants filled out the involvement questionnaire.

The involvement questionnaire consisted of 50 English-language items with a 7-point Likert-scale, which ranged from strongly disagree to strongly agree (see table 2 for the exact wording of the items).

After the participants answered the InvQ for the five video clips they were asked which questions were unclear to them, and how they would interpret these questions, or rewrite questions to make them clearer. After those questions were covered participants were asked to go over the remaining questions and state how they interpreted them. Participants were also asked whether they had seen some of the video fragments before, and whether they found any spelling or grammar mistakes.

2.2 Results and Discussion

All interviews were transcribed. Two raters judged and agreed whether the participants interpreted items similarly or not. Items were judged similar if the explanation of the participants contained the same words or synonyms.

Ten out of 50 questions were interpreted differently by the participants, and were removed since they were not free of ambiguity. For example, "I laughed out loud" and "I feel sleepy" were removed, since two participants pointed out that these questions were hard to answer using a Likert scale. Additionally, two questions were only appropriate with a yes/no scale, and so they were removed. Table 2 shows the items used for the online survey, before and after the cognitive interviews. The new version of the involvement questionnaire ended up with thirty-eight questions.

3. TESTING THE ITEM POOL VIA AN ONLINE SURVEY

An online survey was used to test the involvement questionnaire (invQ) with a larger number of subjects [11]. The invQ was implemented as an online survey to reach as many people as possible in the shortest amount of time. This also allowed testing with a more multi-cultural sample of participants

3.1 Method

3.1.1 Participants

Of the 161 people who started the online survey, 107 people finished viewing all clips and answering all questions. Of those, sixty-three were male and forty-four were female. Mean age was 31 years, with a standard deviation (SD) of 7.6. On average, participants owned 1.5 TVsets. Participants who did not own a TVset watched multimedia on their PC/notebook/Mac. Participants indicated that for 36% English was their mother tongue, and other major languages represented were Dutch (35%), German (9%), and French (6%). Also, participants indicated that most of them (35%) watched between 5-10 hours of multimedia the week before completing the online survey.

3.1.2 Audio/Video

Stimuli Nine different clips were selected to gather feedback on a wide range of audio/video content. From each clip, a one-minute fragment was selected. Table 1 details the selected fragments. One additional multimedia source was selected to serve as training for participants, i.e. to go through the questionnaire once and become acquainted with the questions.

All fragments were encoded to flash video, with their native aspect ratio of 3:4 or 16:9 kept constant. Fragments were offered in four different pseudo random orders.

3.1.3 Scales

Demographic questions. To gather information about the participants' background, a number of demographic questions were asked, such as age, gender, nationality, mother tongue, highest finished education, current occupation, the amount of TVsets they have (if none, what do they use to watch multimedia, and if one or more, where are they located), amount of hours multimedia watched last week, which browser platform they were using (Firefox, IE, etc.), which kind of screen they were using and which screen resolution.

Involvement Questionnaire. After viewing an audio/video fragment, participants were asked whether they had seen the fragment already, and if so, when (approximately) and where did they think it came from (an example answer could be '6 months

ago, Star Wars!'). Once they completed those questions, the 38 items from the invQ were offered. Each item was rated on a seven-point Likert scale, ranging from strongly disagree to strongly agree.

3.1.4 Procedure

Participants were invited to participate through email, which was distributed through several email lists. Once participants clicked on the link to the online survey, they were asked to pick a username and a password. This was necessary to be sure that each participant received a unique ID, to secure anonymity for all participants. Once they created the username and password, the next screen came up, showing instructions. After the instructions, the demographic questions were shown (along with a request to answer them). Next, instructions for filling out the questions with the seven-point Likert scale appeared on the screen. The training fragment appeared, and the respondents went through the invQ for the first time. Subsequently, the fragments were shown, and after each fragment participants were requested to fill in whether they had seen it before or not, and then the invQ. Three participants, selected randomly, received a twenty euro Amazon voucher for their participation.

Table 2. Items as used in the first and the second item pool. If the item remained the same in both item pools, the second item pool only mentions 'idem'.

First item pool	Second item pool
1. I felt like laughing out loud.	Idem
2. I felt relaxed.	Idem
3. While watching the video I was getting restless.	Idem
4. The video held my attention.	Idem
5. I enjoyed the images in the video.	This video was aesthetically appealing to me.
6. This video totally took me away to another world.	This video took me away to another world.
7. I felt disappointed when the video finished.	Idem
8. While watching the video I was shifting position often.	Idem
9. This video had me on the edge of my seat.	This video sequence had me on the edge of my seat.
10. I felt more informed after watching this video.	Idem
11. I lost track of time.	Idem
12. The video entertained me.	Idem
13. This video gave me new ideas to think about.	Idem
14. I was totally absorbed.	Idem
15. This video really fascinated me.	Idem

Table 1. Ten audio/video fragments selected for the online survey.

Fragment	Description
training	Women in mid-19 th century outfits are talking inside a house
documentary	Two ice-bears sit on drift ice while a narrator comments on their circumstances.
soap opera	Mother, daughter and two grandparents are fighting about who should pay the (grand)daughter's college tuition.
feel-good movie	The arrivals gate of an airport is shown, with a voice-over mentioning how love actually is everywhere.
science-fiction movie	An arena where robots are encircling a small amount of people is shown. Another figure demands that these people join his cause.
cartoon	A little dragon wakes up a girl and telling her that she will be late for soldier-training.
drama movie	A young man and his motorcycle are being beaten up.
news	Discussion about the legalisation of marijuana.
sports	Finale tennis game of Wimbledon 2008.
boring	Somebody is mowing their lawn.

<i>First item pool</i>	<i>Second item pool</i>	<i>First item pool</i>	<i>Second item pool</i>
16. This video made me curious to learn more.	I would like to know what happens next.	37. I felt confused after watching this video.	Idem
17. I felt amused.	Idem	38. I felt like switching channels, to see if there was something better on tv.	Idem
18. I felt excited after watching this video.	I felt excited while watching this video.	39. Watching this video taught me something new.	Not used
19. I felt like talking to the television.	Idem	40. Watching this video gave me a new understanding of the subject.	Not used
20. I felt involved.	Idem	41. I felt emotional.	Not used
21. I sympathized with what was going on.	Idem	42. I laughed out loud.	Not used
22. I would have turned the tv off if I could.	Idem	43. I enjoyed the images in the video.	Not used
23. I could relate to the situation.	Idem	44. I didn't think about anything else while watching.	Not used
24. There was an obvious plot.	There was a coherent storyline.	45. I was willing to believe in the world shown in this video.	Not used
25. I was happy when the video finished.	Idem	46. I enjoyed the graphics in the video.	Not used
26. I felt disgusted after watching this video.	I felt upset after watching this video.	47. This video helped me to learn about myself.	Not used
27. I felt embarrassed because of the way somebody acted in the video.	Idem	48. This video gave me something to talk about later.	Not used
28. Watching this video made me feel scared.	Idem	49. I smiled while watching the video.	Not used
29. I forgot where I am.	I forgot where I was.	50. I feel sleepy.	Not used
30. I felt compassion for what was going on.	Idem		
31. This video gave me interesting information.	Idem		
32. I was bored because of the video.	The content of this video does not interest me in any way.		
33. I was making predictions about what would come next.	Idem		
34. Watching this video made me feel sad.	Idem		
35. I would have liked it if the video continued.	Idem		
36. I was concerned about the situation in the video.	Idem		

3.2 Results

To determine which questions best represent the operational definition of involvement, and which questions are best eliminated, Muthen & Muthen [13] advise to determine the number of factors through the use of e.g. the eigenvalues and whether it is possible to interpret all factors. Next, the quality of the items can be assessed through the size of their factor loadings, and whether they load on one or more factors. The quality of factors can also be assessed by the amount of items loading on it, e.g. 3 to 4 items is advised. Once poor items and factors are identified, they can be eliminated and the exploratory factor analysis can be repeated to see whether the remaining items and factors are stable.

3.2.1 Discriminative Power Analysis

Discriminative power analysis was used to investigate which items differentiate between groups of participants [14], to assist with eliminating poor items. The groups of participants investigated are those with the highest and lowest scores on each item. The discriminative power (DP) was calculated per item per multimedia fragment by taking the highest 25% and the lowest 25% of the scores. A weighted mean per quartile is calculated, and the DP is created by subtracting the two means. A larger

difference between the means indicates a more effective item. The decision was made to look at the lowest and highest 25 scores.

Results showed that the following questions differentiate well between participants (for details see table 2):

- questions 3, 14, 16, 21, 25, and 33 are stable across all 9 audio/video fragments;
- questions 7, 12, 23, 30, 35, and 38 are stable across 8 audio/video fragments

Results further showed there were items that scored low and thus did not differentiate well between participants:

- questions 19 and 27 score low across 8 audio/video fragments; and
- questions 1 and 28 score low across 7 audio/video fragments.

3.2.2 Data Screening

Before starting with Exploratory Factor Analysis, data needed to be assessed for normality, since several methods for factor analysis make the assumption of normality. Normality was inspected via histograms. Few items showed a normal distribution, which means that the technique chosen for factor analysis should not rely on the assumption of normality.

Factorability and communalities were inspected across all fragments to assure that enough participants filled out the scale and that the data would be sufficiently reliable and valid. Per multimedia fragment, the ratio of participants to items was approximately 2,5 to 1. While this was on the low side, there are other factors which determine whether enough data is gathered to achieve stable and precise results [15], such as factorability and communalities. Factorability was determined through the Kaiser-Meyer-Olkin (KMO) measure. The value for KMO varies between 0 and 1.0, and it is advised to make sure that the overall KMO value is above .60 before proceeding with an EFA [16]. For the current experiment, no KMO values were below .60., hence it was possible to proceed with Exploratory Factor Analysis [16].

Average communalities were well above .5 [15] and although the standard deviations are larger than expected, the hypothesized factors are reasonably determined [15]. MacCallum et al. [15] advise a minimum of 6 or 7 items per hypothesized factor. After the elimination in section 2, each hypothesized factor between 4 to 9 items. Therefore, the decision was made to continue with the current dataset.

3.2.3 Exploratory Factor Analysis

From the available Exploratory Factor Analysis (EFA) techniques, principal factors analysis is least affected by non-normality (as opposed to principal components analysis and maximum likelihood) and was used for all reported EFAs [16]. The exploratory factor analysis was conducted for the complete data set, since the end goal was to produce a scale which is valid and reliable for multiple kinds of audio/video fragments. To determine the number of factors, the scree test and parallel analysis was used.

The scree test, developed by Cattell [16], plots the value of eigenvalues against the amount of factors. To determine the cutoff points for the factors, one looks for the point where a straight line drawn through the points changes its slope. However, this is not an exact science.

Parallel analysis was first proposed by Horn [16]. O'Connor [17] constructed a program for SPSS with which parallel analysis can be conducted. The program randomly generates a data set with the same amount of cases (N) and variables (items). Next, the program performs several factor analyses (without rotation), while keeping track of the eigenvalues. As a last step, the program averages out the eigenvalues per factor. It is advisable to retain only the factors with eigenvalues which exceed the averaged eigenvalues from the randomly generated data set. Considering that the data in the current dataset are not normally distributed, the decision was made to use permutations from the raw data, rather than completely randomly generated data (<https://people.ok.ubc.ca/briocconn/nfactors/nfactors.html>, 19.01.2010). The results of the parallel analysis found that 4 factors would provide a good explanation of the data.

To obtain a first insight into the factors, the EFA was performed with the instruction to cut off factors once the eigenvalue in the analysis before rotation is greater than 1. Considering that the factors of the involvement construct might not be orthogonal to each other, the oblique rotation method promax was used. Oblique rotation allows correlations between factors [16], and unless there are correlations above .32 (indicative of 10% or more overlap in variance among factors), oblique rotation should not be used [16].

Oblique rotation proved to be necessary, since there was least one correlation above .32 for the EFA [16]. Therefore, the decision was made to run all EFAs with principal factor analysis and promax, to maximize factor loadings and interpretation possibilities. The decision was made to count only items with a factor loading of .400 or higher

Results showed that the following items did not load on any of the 6 found factors and were therefore removed:

- question 5 (“This video was aesthetically appealing to me”); and
- question 24 (“There was a coherent storyline”).

Furthermore, question 19 (I felt like talking to the television) loads on two factors, so it was removed as well. The discriminatory power analysis also showed that the following items did not differentiate well, and were therefore removed:

- question 1 (“I felt like laughing out loud”);
- question 27 (“I felt embarrassed because of the way somebody acted in the video”); and
- question 28 (“Watching this video made me feel scared”).

A second EFA was run with the remaining items. Maintaining a balance between the number of items that are necessary to cover the whole involvement construct and the compactness of the scale was an important consideration in the process. For practical purposes, the invQ needs to be as compact as possible. A compact scale increases the possibility for future research to test a wide range of fragments without participants getting bored or tired because of a lengthy questionnaire, rather than the offered fragments. To reduce the item-pool further, communalities and Cronbach’s α from the second EFA were inspected. Results showed that the following items have a low communality, and were removed:

- question 37 (“I felt confused after watching this video”); and
- question 8 (“While watching the video I was shifting position often”).

Removing these items from factor 2 improved Cronbach’s α from .860 to .892. Factor 1 seems overdetermined with 15 questions. Therefore, the decision was made to take items out of factor 1. Comments from participants (made during the cognitive interviews or written during the online survey) lead to the removal of the following questions:

- question 4 (“The video held my attention”);
- question 11 (“I lost track of time”); and

- question 29 (“I forgot where I was”).

Participants stated that of course the video held their attention, since that was the task asked of them. Regarding question 11 and 29, the comments were that the multimedia fragments were too short to lose track of time or to forget where you are. Considering that neither question 4 nor 11 made it to the top consistently in the DP analysis, these comments are probably also valid for one-minute multimedia fragments. Question 33 (I was making predictions about what would come next) has a communality value below .3, and was removed from the questionnaire. Finally, question 20 was removed (I felt involved), because it was already covered by other, less abstract, questions (e.g. ‘This video took me away to another world.’).

A third EFA was run to determine the final structure of the questionnaire. Table 3 shows the new version of the invQ, and the factor loadings. The results from the last exploratory factor analysis shows that question 35 (I would have liked it if the video continued) loads on both factor 1 and factor 2 (see figure 2). However, this might not be the case in subsequent research, so herefore question 35 is retained. Furthermore, the decision was taken to change question 10 to “I would discuss this video with others”, and question 13 to “I could easily understand what was going on”. Question 16 was rewritten to represent curiosity more explicitly, and was changed to “I am curious about what happens next”.

Table 3: The items remaining in the invQ.

Nr.	Question	Factor	Loading
6	This video took me away to another world.	1	.653
7	I felt disappointed when the video finished	1	.655
9	This video sequence had me on the edge of my seat	1	.866
12	The video entertained me.	1	.745
14	I was totally absorbed	1	.711
15	This video really fascinated me.	1	.709
16	I am curious about what happens next.	1	.632
17	I felt amused.	1	.781
18	I felt excited while watching this video.	1	.940
35	I would have liked it if the video continued.	1/2	.541/.450
3	While watching the video I was getting restless	2	.647
22	I would have turned the tv off if I could.	2	.867
25	I was happy when the video finished.	2	.878
32	The content of this video does not interest me in any way.	2	.785
38	I felt like switching channels, to see if there was something better on tv.	2	.810
10	I would discuss this video with others.	3	.945
13	I could easily understand what was going on.	3	.843
31	This video gave me interesting information.	3	.903
2	I felt relaxed	4	-.501
26	I felt upset after watching this video.	4	.808
34	Watching this video made me feel sad.	4	.749
36	I was concerned about the situation in the video.	4	.618
21	I sympathized with what was going on.	5	.800
23	I could relate to the situation.	5	.727
30	I felt compassion for what was going on.	5	.704

4. DISCUSSION & CONCLUSION

With regard to the previously stated hypotheses, the first version of the invQ allowed a global indication of where participants fell on the involvement continuum. The scores for the invQ also differed within participants, depending on the audio/video fragment show. However, the six previously hypothesized clusters did not replicate into six factors. Therefore, the involvement with audio/video content construct needs to be updated. Almost all questions from factor 1 could be characterized via the captivated and expressions of engagement clusters. Factor 2 reflected items that were based on lack of involvement, while items loading on factor 3 reflected informative interest. Factor 4 items represented negative affect, and the items loading on factor 5 all represented empathy.

Further research is thus necessary to determine whether the current construct of involvement is stable. The validity of the invQ also needs to be further addressed. While there is a high face validity, other kinds of validity, such as discriminant validity [11], have not been tested. Additionally, behavioural measures for involvement with audio/video content can now be developed and compared with the invQ. The scope of the invQ appears to be limited to audio/video content. Possible applications would be in the area of atmosphere: does changing the light or dimensions of a room change our involvement with audio/video content? Another possibility is watching in group: so far, our research has focused on watching audio/video content alone. However, does our involvement with said audio/video content change when we watch it with a group of people? And does it matter whether this is in a cinema, or in the privacy of our own home? Also, does it matter how well we know the group of people, and how large the group is?

Perceived video quality is also an application area. Perceived video quality is influenced by numerous factors, and while several models [18-21] assume that involvement is part of this, the relation between involvement and perceived video quality has not been fully experimentally established. Additionally, considering that 3D television is supposed to enhance the viewer experience [22], it would be interesting to see whether the same content in 3D or 2D would be rated differently on the invQ.

When assuming that involvement is an important factor for a certain user experience, it might be necessary to define involvement such that it is specific enough. An overall measure for immersive tendencies [4] might not be able to capture the wanted part of the user experience under investigation. Creating a valid and reliable questionnaire is not an easy process, but the process established in this paper can hopefully contribute to the creation of well-defined constructs, and valid and reliable

measurements for these constructs.

5. ACKNOWLEDGMENTS

This study was conducted while the first author was working for Philips Research Eindhoven. Our thanks go to Don Bouwhuis for his comments.

6. REFERENCES

- [1] Van den Ende, N., Sekulovski, D., Hoonhout, J. and Meesters, L. 2008. Attributes Underlying Involvement with Video Material. In Proceedings of the 1st international conference on Designing interactive user experiences for TV and video. Silicon Valley, California, USA, 2008, ACM.
- [2] Trochim, W. M. K. 1989. An Introduction to Concept Mapping for Planning and Evaluation. *Evaluation and Program Planning*, 12, 1-16.
- [3] Zaichkowsky, J. L. 1986. Conceptualizing Involvement. *Journal of Advertising*, 15, 2, 4-14.
- [4] Witmer, B. G. and Singer, M. J. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence-Teleoperators and Virtual Environments*, 7, 3, 225-240.
- [5] Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T. and Walton, A. 2008. Measuring and Defining the Experience of Immersion in Games. *Int. J. Hum.-Comput. Stud.*, 66, 9, 641-661.
- [6] Zaichkowsky, J. L. 1985. Measuring the Involvement Construct. *The Journal of Consumer Research*, 12, 3, 341 - 352.
- [7] Celsi, R. L. and Olson, J. C. 1988. The Role of Involvement in Attention and Comprehension Processes. *The Journal of Consumer Research*, 15, 2, 210 - 224.
- [8] Klimmt, C. and Vorderer, P. 2003. Media Psychology "Is Not yet There": Introducing Theories on Media Entertainment to the Presence Debate. *Presence- Teleoperators and Virtual Environments*, 12, 4, 346-359.
- [9] Sweetser, P. and Wyeth, P. 2005. Gameflow: A Model for Evaluating Player Enjoyment in Games. *Comput. Entertain.*, 3, 3, 3-3.
- [10] Malone, T. W. 1982. Heuristics for Designing Enjoyable User Interfaces: Lessons from Computer Games. SIGCHI conference on Human Factors in Computing Systems. ACM, Gaithersburg, Maryland, USA, 63-68.
- [11] Spector, P. E. 1992. Summated Rating Scale Construction: An Introduction. Sage Publications, USA.
- [12] Loevinger, J. 1957. Objective Tests as Instruments of Psychological Theory. *Psychological Reports*, 3, 635-694.
- [13] Muthén, L. K. and Muthén, B. O. 2010. Mplus User's Guide. Sixth Edition. Muthén & Muthén, Los Angeles, CA.
- [14] Hosker, I. 2002. Social Statistics.

- Data Analysis in Social Science Explained. . Studymates Ltd., UK, Somerset.
- [15] MacCallum, R. C., Widaman, K. F., Zhang, S. B. and Hong, S. H. 1999. Sample Size in Factor Analysis. *Psychological Methods*, 4, 1, 84-99.
- [16] Tabachnick, B. G. and Fidell, L. S. 2006. *Using Multivariate Statistics*, 5th Ed. Prentice Hall
- [17] O'Connor, B. P. 2000. Spss and Sas Programs for Determining the Number of Components Using Parallel Analysis and Velicer's Map Test. *Behavior Research Methods, Instruments, & Computers*, 32, 3, 396-402.
- [18] Engeldrum, P. G. 2000. *Psychometric Scaling: A Toolkit for Imaging Systems Development*. Imcotek Press, Winchester, MA.
- [19] Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R. and Yang, Z. 2009. Quality of Experience in Distributed Interactive Multimedia Environments: Toward a Theoretical Framework. In *Proceedings of the Proceedings of the seventeen ACM international conference on Multimedia* (Beijing, China, 2009). ACM.
- [20] Pereira, F. 2005. A Triple User Characterization Model for Video Adaptation and Quality of Experience Evaluation. *IEEE 7th Workshop on Multimedia Signal Processing*, 1-4.
- [21] Reiter, U. 2009. Perceived Quality in Consumer Electronics -from Quality of Service to Quality of Experience. *IEEE 13th International Symposium on Consumer Electronics*.
- [22] Seuntjens, P. J. H. 2006. *Visual Experience of 3d Tv*. PhD dissertation, Eindhoven University of Technology, Eindhoven.