D3.3 : Game Experience Questionnaire

FUGA - Fun of Gaming

028765

FUGA

The fun of gaming:

Measuring the human experience of media enjoyment

STREP / NEST-PATH

Deliverable D3.3:

GAME EXPERIENCE QUESTIONNAIRE

K. Poels, Y.A.W. de Kort & W.A. IJsselsteijn

TU/e

Title of Contract	**FUGA - The fun of gaming: Measuring the human experience of media enjoyment**
Acronym	**FUGA**
Contract Number	**NEST-PATH-028765**
Start date of the project	**01.05.2006**
Duration	**36 months, until 30.4.2009**
Reporting period	
Date of preparation	
Project coordinator	**Name: Niklas Ravaja** **Organization: Helsinki School of Economics** **Phone: +358 40 411 4565** **Fax: +358 9 431 38391** **E-Mail: niklas.ravaja@hse.fi** **Project web site: http://www.hse.fi/fuga**

Partners:	
Partner 1 (CKIR)	FI
Partner 2 (TKK)	FI
Partner 3 (HGO)	SE
Partner 4 (HMTH)	DE
Partner 5 (UKA)	DE
Partner 6 (TUE)	NL

Project funded by the European Community under the FP6
New and Emerging Science and Technology (NEST) programme

Executive summary

The current report describes the development of the Game Experience Questionnaire, a self-report measure that aims to comprehensively and reliably characterise the multifaceted experience of playing digital games. Its development and evaluation precede the testing of the other measures of game experience targeted in the FUGA project, as it will be employed for evaluating validity of these measures.

The theoretical work performed by all partners in the FUGA project (Deliverable 2.1) provided a firm and broad basis for the development of the scale. In addition, relevant scientific literature was reviewed for any conceptualisations that might complement this work and for existing scales that might serve as a starting point or inspiration for item formulation. Furthermore, empirical data was gathered during focus group interviews with different types of gamers. These interviews served as a test for comparing scientific conceptualisations and lay descriptions of first-hand experiences. They were also used as a reference guide on choice of wording in the item formulation phase.

The Game experience questionnaire is developed with a modular structure, consisting of:

1. The core questionnaire (GEQ). This is the heart of the Game Experience Questionnaire, probing multiple components of players' experience while gaming.
2. The post-game questionnaire (PGQ), probing gamers' experience after the gaming session and any after effects.
3. The social presence module (SPGQ), probing gamers' experience of and involvement with their co-player(s).

These three lists are to be administered after the gaming session has ended. Additionally, a short in-game version of the GEQ was developed, the iGEQ, for probing in-game experience multiple times during a gaming session.

A large scale survey was performed to test the long list (>100 items) and explore the factor structure of the questionnaires. Factor analyses provided a structure for the scale that made good sense in the light of theoretical considerations, with subscales that were all easy to interpret. Subsequent reliability tests resulted in the construction of robust subscales with satisfactory to high internal consistencies.

Additional explorations were subsequently performed to check sensitivity and validity of the developed measures. Statistical analyses firmly demonstrated that the GEQ and additional modules were sensitive enough to pick up differences between gamers, game types, play characteristics, and social context of play. The findings already provide new insights to scholars in the field. The Game Experience Questionnaire is now ready to be translated and used in FUGA experimentations.

Table of Contents

1. Introduction

Over recent years, digital games have rapidly gained in interest. In pop culture they rival film and television in popularity. Originally deemed frivolous, game design principles and game engines are now being re-used for purposes of education, training and therapy, spawning an area now known as 'serious games'. Meanwhile, digital games have significantly diversified in design. From classic arcade games to embodied games (e.g., the Nintendo Wii), mixed reality games, and pervasive games. Research on digital games has been thematically and disciplinary diverse, perhaps born from the multifaceted nature of digital games themselves. Scholars from a variety of perspectives, including computer science, visual design, film and television theory, performing arts, literary theory, and media psychology, to name but a few, have taken an interest in digital games. More recently, HCI as an academic discipline has turned its attention towards digital games, finding inspiration for engaging interface design, creative new metaphors for human-computer interaction, and generally focusing on the captivating experiences that can be designed through this class of computer-based technologies.

Today, the area of digital games constitutes a tremendously varied set of applications, with a wide range of associated player experiences, defying a one-size-fits-all approach to its conceptualisation and measurement. One of the main challenges facing the gaming research community is a lack of a coherent and fine-grained set of methods and tools that enable the measurement of entertainment experiences in a sensitive, reliable and valid manner. Much like the six wise (but blind) men touching the elephant, no single methodological perspective can be said to provide a comprehensive understanding of digital gaming. Following this insight, the FUGA consortium takes a critical view towards the exclusive reliance on any one single indicator for measuring game experience. In the FUGA project, we explicitly strive towards a multi-method, multi-measure approach whereby we anchor and cross-validate various measures (e.g., self-report, psychophysiological, behavioural, neural) via their simultaneous application to a certain standardized set of games, and correlating the results thus obtained. Assessing the psychometric properties (sensitivity, reliability, validity) of all measures developed in the project is one of the defining characteristics of FUGA.

We believe that a large range of measures, from reflective (subjectively controllable) to fully reflexive (uncontrollable) responses, enables a fuller characterization of the game experience than any single isolated measure, thus sensitizing us to the rich gamut of experiences associated with digital games. Moreover, limitations particular to one measure may be overcome or compensated by using corroborating evidence emerging from another measure. The combination of multiple measurement modalities can thus reduce uncertainty associated with measuring a single modality, resulting in increased validity, robustness and wider applicability of the total set of measures.

As a significant first step, we report in this deliverable on the development and validation of the Game Experience Questionnaire (GEQ). The GEQ is intended to become a freely available and widely applicable measure which allows researchers to obtain a reliable and valid indication of participants' *subjective experiences* associated with digital gameplay, both during and after a gaming episode. The GEQ covers a wide range of digital game experiences that have been identified through reviewing theoretical accounts of player experiences (for an overview see FUGA Deliverable 2.1 'Working Model of Digital Game Experience'), as well as through focus group explorations with both frequent and infrequent gamers. The factor structure of the GEQ, which has been established based on a large-scale survey, is very much in line with the theoretical constructs that have been identified in the FUGA Working Model of Digital Game Experience.

The current deliverable documents the development and evaluation of the Game Experience Questionnaire. In the remainder of this chapter, we will briefly discuss a number of relevant psychometric quality criteria that any game experience measure needs to adhere to. In addition, we will highlight some of the quality criteria that are of particular relevance in questionnaire design. Chapter 2 describes in detail the results of a number of focus group explorations of digital game experience. These focus groups provided us with important insights into lay conceptualisations of game experience, and were a valuable addition to the theoretical groundwork that was reported on in Deliverable 2.1. In an expert meeting both theoretical perspectives and focus group results were subsequently combined, providing a solid foundation for the item generation phase of the questionnaire development. Chapter 3 describes the development of the GEQ, including the general logic of the GEQ, item generation and selection, scale construction based on exploratory factor analysis, and internal consistency. Chapter 4 further explores the sensitivity of the GEQ scales to differentiate between gamers, game types, play characteristics, and variations in the social context of

play. Finally, Chapter 5 provides a brief summary and future outlook, describing the translation procedure of the GEQ, open issues, and plans for future validation.

1.1. Criteria for game experience measures

Several criteria or properties can be formulated to which a good measure of game experience should adhere. These are reliability, validity, sensitivity, robustness, non-intrusiveness, and convenience. From a psychometrical point of view, reliability, validity, and sensitivity are the three most important criteria that should be satisfied. More information on basic psychometrical properties of tests can be found in Cronbach (1990). We will briefly summarise the main criteria below.

Reliability

A measure should be consistent and stable over time, meaning that it should give comparable results if administered under comparable conditions. There are several ways to estimate reliability:

- Internal consistency reliability: the extent to which the items of a measure address the same underlying trait or characteristic. There are different ways of estimating internal consistency: Cronbach's alpha, (average) inter-item correlation, (average) item-total correlation, and split-half reliability.
- Test-retest reliability: the stability of a measure over time. Test-retest reliability can be calculated as the correlation between scores gathered at different occasions. Researchers should make sure that it can be reasonably assumed that the measured construct itself has not changed between these occasions.
- Inter-rater reliability: the degree to which different observers agree in their assessment. Inter-rater reliability can be calculated as the percentage of cases in which the observers give the same rating (for nominal measurement) or the correlation between ratings (for ordinal, interval, or ratio measurement)
- Parallel-forms: the consistency of similar measures. Parallel forms can be created by generating a large set of items addressing the same concept and randomly dividing them into two sets, administering them, and calculating the correlation. A related approach is to divide an existing measurement tool into two halves randomly, and calculate the correlation between the two scores. This is called split-half reliability.

Ideally, reliability should be assessed in several ways in order to draw firm conclusions about a measurement instrument. In reality, however, this is often not possible. Internal consistency is the most widely used form of reliability, because it can be assessed using only one measurement instrument, on one occasion, in one population. For the Game Experience Questionnaire developed in this project, we report on the internal consistency (Cronbach's alpha) of all subscales and the total questionnaire in the current deliverable (Chapter 3). The test-retest reliability of the entire battery of measures developed by FUGA (including the GEQ) will be established in WP6. Inter-rater reliability is only of relevance in cases where multiple raters are employed to score observational data. In FUGA, this is of particular relevance in the case of behavioural measures, where video observations need to be coded into uniform categories of behaviour by multiple observers.

Validity

A measure should address the intended construct. There are many different approaches to validity, the most important of which are discussed here.

- Face validity: the extent to which a measure appears to address the intended construct. This approach is based on subjective judgment, preferably of several experts.
- Content validity: the instrument is checked against the relevant content domain for the construct, to see whether the instrument is compatible with theories and addresses all relevant dimensions of the construct. For this approach, it is necessary that theories about and clear definitions of the construct and its dimensions exist.
- Criterion-related validity: comparing the measure to some other measure or criterion. Different forms of criterion-related validity are predictive validity (the extent to which a theoretically relevant criterion can be predicted), convergent validity (the degree to which the measure correlates with measures of theoretically related constructs), and discriminant

validity (the degree to which a measure is different from measures of constructs to which it is not theoretically similar).

Face validity and content validity are typically established in the construction phase of a measurement instrument. For the GEQ, they are discussed in Chapters 2 and 3. Criterion-related validity can only be established through research. It is a more objective and therefore more convincing indication of validity. Predictive validity will be established through a series of studies performed under WP7. Convergent validity is in fact at the heart of the FUGA multi-measure approach, as we are comparing and correlating the outcomes of different measurement methods throughout the project lifetime. For the GEQ, discriminant validity is established internally for the various subscales, as one would expect different scales to behave differently depending on the particulars of the game content, interface and setting.

Sensitivity

A measure should be able to distinguish between different levels of a construct with a reasonable level of detail. For example, if different levels of (a component of) game experience are expected based on different form factors, different game content or different individual characteristics, the measure should reflect this difference. For FUGA, temporal sensitivity of a measure is also of particular importance, since the range of experiences we are interested in will likely vary over time within the same gaming session. In Chapter 4 of the present deliverable, we analyse the sensitivity of the GEQ in relation to gender, play frequency, type of game, and social setting.

Robustness

A measure should be applicable across a variety of different media platforms and settings, varying in form, content, and context-of-use. The GEQ has explicitly been developed to be a widely applicable tool, offering a valid and reliable experiential assessment across gaming genres, platforms, play styles, social settings, etc.

Non-intrusiveness

A measure should be as non-intrusive and non-disruptive as possible, and not interfere with the construct that is being measured. For example, when the measurement method requires self-report during the game, this may interfere with game experiences that build on high levels of mental absorption or attentional engagement (e.g., flow, immersion). Measurement devices that are uncomfortable, restrict movement, or that are otherwise highly noticeable may also affect the experience under study. Most real-time measures will inevitably be intrusive to a certain extent (e.g., psychophysiology, fMRI). The GEQ, being a post-test self-report measure, can be assumed to be non-intrusive.

Convenience

A measure should preferably be easy to learn, easy to administer, low-cost, and portable. This will make it much easier to apply a measure outside the confines of a psychological lab, potentially adding to the measure's ecological validity. As a low-cost paper-based measure, the GEQ can be regarded as highly convenient for researchers to use.

2. Explorations of digital game experience: focus groups and expert review

In this chapter we describe the exploratory part of our conceptualization of game experience dimensions. We executed this exploratory part in two stages. First, focus groups were organised to obtain lay-conceptualisations of game experience and second, findings were discussed in an expert meeting with the aim of consolidating empirical and theoretical findings into a tentative model of game experience. We describe those two stages below.

2.1. Focus group methodology

Focus group methodology is a qualitative research tool that is frequently used in social sciences to explore people's meanings, ways of understanding, or experiences of a complex phenomenon (Lunt & Livingstone, 1996). In practice, focus group methodology typically involves a series of group interviews about a given topic or phenomenon guided by a moderator. The moderator plays a pivotal role. Not only does he or she put forward the questions and concrete topics that need to be discussed, he or she also ensures that the discussion remains on the topic at hand (Lunt & Livingstone, 1996). Focus group sizes usually vary from five to eight participants (Morgan, 1998). A rule of thumb with respect to the amount of focus groups is that one should continue to run new groups until the last groups has nothing fundamentally new to add (Lunt & Livingstone, 1996).

One of the major strengths of focus group methodology is its *exploratory nature*. Focus groups enable the researcher to get to know their target audience in detail without the need for a priori assumptions or research questions. Moreover, focus groups can serve as a source of new ideas and hypotheses (Merton, 1987). Further, focus groups are very useful in providing *context and depth*. Besides observing experiences and thoughts, the moderator can probe in order to acquire relevant background information (e.g., about motivations, contexts) on these experiences and thoughts. Related to this, focus group methodology lends itself for *interpretation* of the experiences and thoughts reported by the target audience. As such, it enables researchers to get a clearer view on the *why* of behaviour (Morgan, 1998).

2.2. Focus groups about digital game experience

Focus group methodology is an interesting research tool to explore several facets of digital gaming. For example, Miller et al. (1996) applied focus groups to explore female preferences for specific types of game designs. Their research findings report that girls placed a premium on richly textured video and audio. Girls also preferred collaborating to competing; they liked interacting with male characters, and showed a preference for simulation games. Applying in-depth interviews (i.e., a related qualitative research tool in which the participants are interviewed individually) Brown and Cairns (2004) unravelled some core aspects of the concept of game immersion.

Given the diversity of individual differences with respect to play styles (Bartle, 1996) or motivations to play games (Yee, 2002), focus groups can provide in depth, contextual, and motivational insights into the specific experiences of different types of gamers. To the best of our knowledge, the application of focus group methodology to study digital game experience is still limited.

In the FUGA project we use focus group methodology to *explore lay meanings and experiences of digital gaming*. The underlying reasons are twofold. First, in order to develop a long list of items as a starting point for our Game Experience Questionnaire, we use focus groups as one input source for item generation (next to theory – see D2.1 and expert insights – see section 4.3 of this document). Lay verbalizations can provide a rich perspective on possible game experiences. Moreover, some of them might have been overlooked by current theoretical conceptualizations. Further, lay verbalizations do also aid us in choosing the appropriate formulations of the different game experiences and, as such, ensure face validity of the different items that will be used in the Game Experience Questionnaire. An additional reason to use focus group methodology is to get more insights in the game experiences of different types of gamer like frequent gamers versus infrequent gamers.

2.2.1. Method

We organized three focus groups. Two focus groups (FG1 and FG2) included infrequent gamers (i.e., people who game at least once a month). The third focus group (FG3) consisted of frequent gamers (i.e., people who game at least once a week). Participants were undergraduates and graduated students. Their ages ranged from 19 to 28 years. Specifically, FG1 had five participants of which two were female. FG2 consisted of three male participants. FG3 had four male participants. Each focus group took about 90 minutes and participants were rewarded 10 € for their participation.

The focus groups were structured in the following way:

Introductory round: First, the moderator and the assistant moderator presented themselves and gave a brief description of the main goal of the focus groups. More concretely, they explained that the focus group was about game experience and participants could freely talk about how they experienced digital gaming. Then, participants presented themselves by briefly introducing themselves, giving their name, game frequency, and the type of games they usually played.

Individual task: We asked each participant to reflect for five minutes on what they considered as the most prominent game experiences for themselves. Participants had to write down these experiences on Post-It notes. We also asked them to indicate their most favourite game and the game they had played last. After this, all Post-It's were pasted in the middle of the table to serve as a starting point and inspiration source for the next and most crucial stage, the group discussion (see Picture 1).

Picture 1. Set up for the group discussion

Group discussion: In the group discussion participants could freely talk and interact with each other about their game experiences. The discussion was clustered around three core questions by means of a semi-structured questionnaire. Accordingly, the three core question were fixed but other, side questions could be posed along the way in order to have clarification or in-depth insights. The three core questions were: (1) On what occasions do you typically start gaming?, (2) What do you experience or feel *while* gaming?, (3) What do you experience or how do you feel *after* gaming? The moderator further probed the experiences that were reported by each participant individually. Additional Post-It's were used when new experiences were mentioned.

Group task: At the end of the group discussion participants were asked to cluster and rank all game experiences that were reported on the Post It notes depending on how central they are to gaming in general (i.e., across games). They wrote down all experiences on a large sheet of paper with the

most prominent experiences in the centre of the sheet and the less relevant experiences closer to the margins of the sheet (see Picture 2).

All focus group interviews were recorded and transcribed.

Picture 2. Group task: ranking the Post-Its on a large sheet of paper.

2.2.2. Results

In this results section we focus on the three core questions that were posed in the group discussion. We describe results for occasional and frequent gamers separately.

a) Infrequent gamers

Question 1: On what occasions do you typically start gaming?

The occasions in which the infrequent gamers typically start gaming vary considerably. A substantial amount of them reported that they started playing a game when they felt bored, when they had nothing else to do, or when they didn't feel like doing something else (e.g., studying).

> *....when I am feeling bored or when I don't feel like studying...* (Female participant, FG1, 21 years)
>
> *I game when I feel like gaming, when I don't feel like doing anything else...* (Male participant, FG2, 24 years)

Related to this, some of the infrequent gamers said that they typically started gaming upon coming home after a busy day at school or after an intensive study session.

> *If I come home after a busy day and I don't want to do anything else yet, I often play a couple of quick games before I continue with something else* (male participant, FG1, 20 years)
>
> *I sometimes play a game if I want to relax, de-stress, or divert my thoughts...* (Male participant, FG2, 28 years)

Another occasion that they put forward was more social in nature. Some of the infrequent gamers reported that they often played games when they were with friends, for example, before going out.

> *I rarely game on my own. When I game it is a social event where we sit on the couch, with beer and chips. This usually happens the hours before we go out.* (Male participant, FG2, 22 years)

One female participant reported a combination of the social and the boredom occasion.

> *When we are together with friends and we have a break or when we do not really know what to do, we sometimes play a game together.* (Female participant, FG1, 21 years)

Question 2: What do you experience or feel while gaming?

Almost all participants mentioned *fun*, *amusement*, and *relaxation* as most prominent game experiences.

Experiences during game play were often related to game immersion and flow. This means, the infrequent gamers mentioned experiences like '*loosing connection with the outside world'*, '*forgetting everything around you'*, and '*being fully occupied with the game'.* One participant explicitly linked fun with the experience of game immersion.

> *Feeling happy is linked with loosing connection with the outside world. It is a simple game and you get yourself fully drawn in.* (male participant, FG1, 22 years; talking about playing Mario Bros with his Nintendo console)

Other experiences were more closely linked to imaginative and sensory immersion. For example, '*being creative'*, '*exploring the game world'*. One of the participants linked these experiences to the fun factor:

> *I like it when you get more creative in a game. It is funny when you discover something new, something you did not expect. When you find out something that you were looking for, you feel glad.* (Female participant, FG1, 21 years)

Others said that they did not want to invest much time in exploring a game. For them, relaxation was related to simple and clear games.

> *I don't game very often so if I do I want to know directly how the game works. I don't want to invest much time in exploring the game.* (Male participant, FG2, 22 years)

Further, '*concentration'* and '*tension'* were mentioned as in-game experiences. Participants reported that these experiences were related to challenge and difficulty of the game. Some of them reported that concentration was needed in order to perform well in games, and, as the game got more difficult, more concentration was needed. However, most of them said that they typically gave up when it got too difficult.

> *If it gets difficult, you need to concentrate more. Sometimes I give up, when it really doesn't work.* (Male participant, FG1, 19 years)

Further, a lot of game experiences were connected to playing against co-located others, or playing against the computer. Experiences that are typically mentioned in this context are '*competition'* and '*enjoyment with others'*.

Participants mentioned that competition instigated feelings of '*tension'*, '*nervousness'*, and '*teasing one another'*, while at the same time, they perceived competition as '*fun'*, '*having a laugh*', and '*being connected with others'*.

> *It is always a lot of fun! For example when we play Mario Cart with four friends, there's a lot of friendly banter. It is very funny if one player gets picked on by three others. That enhances the enjoyment you have with others.* (Male participant, FG2, 22 years)

When probing what happens when they themselves were the 'victim' of this group teasing, the answer was:

> *Then I laugh about it, but of course you also feel a little bad. But they often want to get me because I am the one who tries hardest to frustrate them. So, it is always very funny...* (Male participant, FG2, 22 years)

Emotions evoked through competition with co-located people were reported as much stronger than emotions through competition with the computer, or through competition with online people. Also, participants reported that they put more effort in the game when they play against co-located friends.

Moreover, they said that they experienced more tendencies to take '*revenge'*. This was attributed to the physical presence, which enables non-verbal and verbal communication and physical contact.

> *Playing against the computer is totally different from playing against a friend who sits next to you. You can nudge him, give comments...* (Male participant, FG2, 28 years)
>
> *When you play with strangers on the Internet, you miss a part of the communication. You cannot figure out whether they play for fun or not. You cannot tease them.* (Male participant, FG2, 24 years)

Question 3: What do you experience after gaming?

With regard to experiences after gaming, results were mixed. Most of the participants said that '*time had gone by faster than expected*'. When probing on whether this led to feelings of *regret* or *satisfaction*, the answers varied according to personal and situational factors.

> *I often feel bad if I wasted my time with playing a game. However, if it is a lazy Saturday afternoon and you have nothing better to do it doesn't matter. Then I even find it useful to play a game.* (Female participant, FG1, 21 years)

Others also reported feeling bad after gaming longer than they had intended to, especially when they had other, more important things to do.

> *Sometimes when you are studying and you take a short break, you forget the time and then you feel bad that you have wasted your precious time at playing stupid games.* (Female participant, FG1, 21 years)
>
> *If the weather is nice, then I regret that I didn't spend my time outside. I find it a pity then. However, I did have fun, so it is not that bad after all.* (Male participant, FG2, 28 years)

Some of them reported that they anticipated those negative experiences. For example, one participant explicitly stated that he only quit gaming when he was in a favourable position. This way, he reported, he always has a good feeling after gaming. Another participant said that he would not start gaming when he had more urgent things to do (e.g., when he had examinations). Yet another mentioned only playing short games in order to prevent that he would spend his whole evening on playing a game.

b) Frequent gamers

Question 1: On what occasions do you typically start gaming?

There were several different occasions on which the frequent gamers start gaming. On the one hand, they reported that they often started gaming when they had nothing else to do, when looking for relaxation, or when they wanted to divert their thoughts to something else than work or school. This is very similar to the instances mentioned by infrequent gamers. On the other hand, they reported that they played games in a coordinated way, making appointments with friends, and competing with them in a team. In both cases, they played with friends, mostly online. The type of game they played differed with the occasion.

> *I game if I want to do something completely different, for example if I come home after work. Most of the time I play a couple of short First Person Shooter (FPS) games, those games you can play at any moment, against anyone. In the evenings, I play longer Real Time Strategy (RTS) games.* (Male participant, FG3, 28 years)
>
> *I play FPS games if I have nothing else to do, or World Worms Party. When I play Massive Multi-player Online Role Playing Games (MMORPG) it happens in a much more coordinated way, you really need to make appointments beforehand.* (Male participant, FG3, 23 years)

Question 2: What do you experience or feel while gaming?

Frequent gamers found it hard to report their experiences during gaming, since these depended on several factors. In general, their descriptions were richer than those of most infrequent gamers, i.e., involved more facets, e.g., competition, and control.

First, when they reported on their experiences while gaming, they often distinguished between different *types* of games.

> *FPS games are about beating the opponent and are very demanding. As such, the atmosphere and graphics are less important. With MMORPG it is all about the atmosphere and the beautiful scenes. You get yourself fully drawn in*. (male participant, FG3, 29 years)

Second, a lot of the experiences they reported involved game settings in which they cooperate in a team. As such, the social experience is very important. Experiences that emerge from this social setting are '*having fun', 'being powerful', 'having control', 'immersion', 'thrill' and 'satisfaction'*.

> *It is nice to play in a team; we often make a lot of jokes and fun together. The urge to build something evokes pleasure. The feeling of getting more and more power and more control on your environment is also part of the fun; and also that you get more status within your environment.* (Male participant, FG3, 29 years)

> *It is important that you feel that you are one team. For me realism is important so you can fully imagine yourself in the game. It causes more of a thrill. If I experience that I am really someone in the game and for my team, it gives me a feeling of satisfaction.* (Male participant, FG3, 26 years)

Third, they explicitly distinguished between experiences while playing games purely for fun and experiences while playing games in a serious game competition or online in a team.

Particularly, '*immersion'*, '*concentration'*, and '*competition'* seemed to differ between those two types of game play. Interestingly, frequent gamers reported to play console games when playing for fun and PC games when competing in more serious game competitions.

> *With MMORPG and FPS you need to sit close to the screen, they are very exacting. If you meet with friends to have some fun together, it's much nicer to lean back on the couch, the game triggers the fun, but it's also about other things then. We have a drink and we chat. The game play is purely for fun. When we play games on the PC it is much more serious, you need to be very concentrated then, and strictly focused on the game.* (Male participant, FG3, 29 years)

The frequent gamers wrote down several negative emotions as game experiences such as, *'irritation', 'disappointment', 'frustration'*, and even '*anger*'. They reported that frustration and irritation often occur when there is a lack of challenge (i.e., if the game is either too easy or too hard).

> *If a game gets too complicated I am often inclined to turn it off, to quit gaming. There has to be some challenge in the game; I don't like it if it is too easy, but if it is too complex I don't like it either. I get irritated if something doesn't work, I sometimes even get angry.* (Male participant, FG3, 23 years)

> *I often play RTS games against the same person, if we set a high difficulty setting, it gets more challenging. Of course you feel disappointed if it doesn't work and satisfaction if it does work out. I think disappointment relates to the effort you put into the game.* (Male participant, FG3, 28 years)

Also, negative emotions are stronger if the game play gets more serious.

> *Game competitions are dead serious, as serious as a soccer game Holland-Germany. You can really feel aggression, or anger. When you play for fun, it is more informal, having fun is the dominant experience.* (Male participants, FG3, 29 years)

Question 3: What do you experience after gaming?

Frequent gamers were quite unanimous with respect to their experiences after game play. In general, they did not see it as a waste of time and often had the feeling of having done something really useful. Only in very specific situations they reported the experience of disappointment or regret.

> *If I play online games I never experience it as a waste of time. When you are cooperating in a team and one member gives up, it is a pity. Then I feel disappointed.* (Male participant, FG3, 23 years)

> *Gaming is never a waste of time. Watching TV is much more a waste of time, because it is more passive, you are not involved in what happens.* (Male participant, FG3, 28 years)

> *....Only if you have been gaming for quite a long time and you did not achieve anything, I often regret having spent so much time on it. Especially when I have more urgent things to do.* (Male participant, FG3, 29 years)

2.2.3. Conclusions

On most themes we found both similarities and differences between frequent and infrequent gamers who participated in our focus groups.

First, both groups reported that they often started gaming when they had little else to do or when they wanted to relax after a busy day. However, the kind of games they played at such occasions seemed to differ. Infrequent gamers mostly chose to play single player games, whereas frequent gamers mentioned quick online FPS games.

Also, in both groups, participants referred to playing games as a way to relax and meet with friends. Here, the type of game did not seem to differ. Moreover, in both groups participants said that on such occasions gaming was a trigger for enjoyment with others and that other factors (e.g., drinking beer, eating chips, having a chat) accompanied the playing.

Notably, only frequent gamers reported coordinated, team play in digital gaming.

With respect to experiences while gaming, both groups reported experiences like fun, enjoyment, concentration, tension, trill and immersion. Nevertheless, the frequent gamers more explicitly linked the different experiences to different types of games. For example, they mentioned that for FPS games feeling immersed in the story of the game is not as important as in MMORPG's. Participants (implicitly) discerned between immersion in the game's story, the game world (i.e., via sensory stimuli), and immersion through game play. Within our focus groups, only frequent gamers referred to experiences like control and prestige.

Frequent gamers also distinguished between experiences while playing games for fun and experiences while playing games in a competition. More specifically, negative emotions were more frequent and stronger for 'serious' gaming. Infrequent gamers reported fewer negative experiences during play.

As for experiences after gaming, frequent gamers were less likely to call it a waste of time. In contrary, for them playing games was a useful activity and often replaced other leisure activities such as sports or watching television. Infrequent gamers more frequently reported regret, especially when gaming made them loose track of time and, as such, neglect more urgent activities.

At present, we have only distinguished between frequent and infrequent gamers. However, there is also considerable variation within each type. In our focus groups this was especially apparent for the infrequent gamers. Some of them only mentioned playing games as a quick activity in between other (often referred to as more useful) activities. As such, this group can be interpreted as 'incidental', or 'casual' gamers. Others reported only playing games with friends, in an informal, social setting, and might be labelled 'social gamers'. (e.g., see: http://www.gamasutra.com/php-bin/news_index.php?story=10681). For future research, it would be interesting to further explore differences in game experiences between these different types of gamers.

2.3. Expert meeting

After the focus groups described above, an expert meeting was organised, aimed at combining knowledge and insights gathered from theoretical explorations (including the FUGA consortium's deliverable D2.1) and the focus groups, and consolidating these into a tentative model (or rather: concept map) of game experience, and the components it consists of.

Five game researchers participated in this meeting. These included psychologists and experts on measurement development as well as the state of the art in theories of game experience (two females, three males). Two of them are very frequent gamers themselves, the remaining three should be labelled infrequent gamers. This provided a solid base for determining the tentative core dimensions of digital game experience.

The theoretical work performed by all partners in the FUGA project (Deliverable 2.1) provided a firm and broad basis for the development of the scale. In addition, relevant scientific literature was reviewed for any conceptualisations that might complement this work and for existing scales that might serve as a starting point or inspiration for item formulation. In spite of the recent rise of gaming

research, validated questionnaires probing game experience are scarce at this time. Commercial developers of games generally have in-house testers to perform detailed walkthroughs of the game, which provide insight into game mechanics and gameplay. Sweetser and Wyeth (2005) have provided an overview of design heuristics based on the concept of flow and their model has been validated against professional reviews in game review magazines. Although Ermi and Mayra (2005) have developed a self report questionnaire based on their SCI model, their questionnaire is not easily available and thus its validity for our purpose is hard to assess. Their model did provide input for the theoretical framework. Lastly, Ryan, Rigby, and Przybylski (2006) recently developed a new measure of need satisfaction in play, the Player Experience of Need Satisfaction (PENS), elaborated from self-determination theory (SDT, Deci & Ryan, 1985; 2000). A few relevant existing questionnaires probing related psychological concepts (e.g. PANAS, Watson, Clark & Tellegen, 1988; ITC-SOPI, Lessiter et al., 2001; and IMI, Ryan, Mims, & Koestner, 1983, IMI, 2003).

The expert meeting started with an individual and personal reflection of each researcher on what they saw as important game experiences (both in game and post-game experiences). Similar to the procedure of the focus groups, the experts wrote these down on Post-Its. Subsequently, memos were added based on theoretical considerations and the results from the focus group meetings. We then engaged in an interactive and iterative session, organising these on a whiteboard, according to centrality and similarity, thus creating a comprehensive concept map of components of game experience (see Picture 3).

Picture 3. Constructing the tentative, comprehensive model of game experience.

The most important components that resulted from this exercise were:

1. Competence: with *accomplishment*, *euphoria*, and *pride* as in game experiences, and *pride* also as post game experience.
2. Flow: with *concentration*, *being absorbed*, *flow*, *loosing track of time*, *detachment* as in game experiences, and *jetlag* as post game experience
3. Suspense: with *challenge, pressure, arousal, hoping to win, anxiety, tension* as in game experiences, and *release, feeling empty, exhausted* as post game experiences.
4. Enjoyment: with *enjoyment, pleasure, relaxation, exuberated* as in game experiences, and *energized* as post game experience.
5. Sensory immersion: with *sensory immersion and presence* as in game experiences, and *returning to the real world* as post game experience.
6. Imaginative immersion: with *absorbed in story, empathy,* and *identification with game character*.
7. Control: with *control, autonomy,* and *power*

8. Negative affect: with *shame, anger, irritation, disappointment, ignorance* as in game experiences and *guilt* as post game experience.
9. Connectedness: with *enjoyment with others, being connected with others, empathy* as in game experiences.
10. Negative affective experiences related to playing with others: *jealousy, envy, revenge, guilt.*

This provided the basis for the item generation for the long list of the Game Experience Questionnaire. Although such a list can never be exhaustive, our goal in constructing the list was comprehensiveness rather than selectiveness.

3. Development of the Game Experience Questionnaire

In the current chapter we describe the rationale for questionnaire construction and item generation, and the proceed with the methodology, statistical analyses and results of the empirical investigation into the structure of the Game Experience Questionnaire (GEQ). Lastly, all items and subscales are presented and their reliability (internal consistency) is discussed. Sensitivity and validity will be explored in the following chapter.

3.1. General logic of the GEQ development: Core and modular elements

The Game experience questionnaire is developed with a modular structure, consisting of:

1. The core questionnaire. This is the heart of the GEQ, probing multiple components of game experience. Every component should consist of 4-6 items, with high internal consistency. This questionnaire is administered immediately after a gaming session. In addition, a shorter, in-game version of the core GEQ will be developed, which can be administered multiple times during a gaming session (e.g., during short breaks). This in-game version will be modelled after the core questionnaire, but consists of only two items per component.
2. The post-game questionnaire, probing the gamers' experience after having played the game and any after effects (e.g., returning to reality, fatigue, pride, guilt).

These first two questionnaires should be applicable for any type of digital game played, in any type of setting. The remaining modules are developed for specific types of games or settings. One such module is:

3. The social presence module, probing the gamers' experience of and involvement with their co-player.

This module only applies to situations where digital gamers played with or against one or more social entities, be they co-located players, online players, or virtual players (i.e., in-game characters). In the near future, additional modules will be developed. We envisage:

4. Embodiment module: physical engagement, kinaesthetics, embodied play).
5. Pervasive game module (mixed reality, city games).

According to the present plans as stated in D3.1, these modules will not be applied within the FUGA project in the near future. We have therefore decided to first focus on the development of the core module, the post-game module, and the social presence module.

3.2. Item generation

The results of the focus groups and expert meeting were used as a starting point and reference guide for item generation. A long list of items was created with the following requirements:

1. Items should probe experience 'sec', i.e., is possible, not refer to specific game characteristics, events, or external circumstances, nor should it require respondents to attribute experiences to specific circumstances or events.

2. Wording should be as natural and intelligible as possible for all types of research participants. In formulating the items we tried to stay close to the original wording used by participants of the focus groups.

3. Qualifiers: the use of qualifiers in the items should be minimized, if not prevented.

4. Items should be formulated such that one uniform type of answering scale can be applied, in order to minimise participant burden.

A 5-point, unipolar intensity-based answering scale was developed, with points anchored at *not at all* (0), *slightly* (1), *moderately* (2), *fairly* (3), and *extremely* (4), thus qualifying the intensity of the experience described in the item.

In total, 92 items were formulated in the long list for the core GEQ. The list was constructed such that all ten components of the tentative model, listed in section 4.3. were represented by at least five or more items. Any other possibly relevant items probing additional aspects of game experience were also added, mainly based on focus group data. A few relevant existing questionnaires probing related psychological concepts (e.g. PANAS, ITC-SOPI, IMI) were also used. Twenty one items were formulated for the post-game module.

For the social presence questionnaire, Biocca et al.'s (2001) networked minds measure of social presence provided an excellent starting point for item formulation. It is based on a recently developed theory of social presence and constructed for measuring experiences in social presence technologies such as mediated collaborative work environments, mobile and wireless telecommunication, and teleconferencing interfaces (Biocca et al. 2003). Some items were used directly, others were adapted and some new items were formulated based on the focus group data. The long-list consisted of 25 items. The answering scale used was identical to that used for the GEQ and PGQ.

3.3. Research method

Design

For exploring the factor structure of our original set of items a sizeable and representative sample of gamers was needed. As the questionnaire measure is intended to be applicable across a range of gaming genres and platforms, it was important to allow participants to play a self-selected game on their own preferred gaming platform (PC, console, mobile), and in their own preferred physical setting (e.g., living room, bedroom, Internet Café, etc.). Based on these considerations we chose to use an online survey, inviting people to play their game of choice and to fill out our questionnaire immediately afterwards.

All items from the game experience questionnaire (see section 5.2), consisting of the core questionnaire (92 items), the social presence questionnaire (25 items), and the post-game questionnaire (21 items), were measured by means of a five point intensity scale with points anchored at *not at all* (0), *slightly* (1), *moderately* (2), *fairly* (3), *extremely* (4).

The full online survey was set up in four parts: one part with instructions, an introduction, the actual game experience questionnaire, and a final part with general game related and socio demographic questions. We will elaborate on those parts in the procedure section.

Participant recruitment and selection

Participants were recruited via the internet, using two different channels: Virtual Lab (vlab) and internet game forums. Both websites allowed us to track the time participants spent in filling out the questionnaire and allowed us to immediately exclude users who had already filled in the questionnaire once. On average, it took ten to 15 minutes to complete the full survey. If participants were exceptionally fast at filling out the questionnaire (i.e., eight minutes or less), we discarded them from the dataset. If they were faster than average (i.e., less than ten minutes), we checked whether they were consistent in their answers by looking at the correlation of three fun-related variables (*I enjoyed it, I thought it was fun, I liked it*). If these were not or negatively correlated, we excluded those participants from our final sample.

We launched our survey at vlab, an internet based lab connected to the Human Technology Interaction Department of Eindhoven University of Technology. An invitation to participate in a study on game experience was sent out to 495 people registered at the vlab database. A total of 262 people responded to this invitation, yielding an initial response rate of 53%. However, we excluded 50 respondents who did not fill in all game experience items or who filled in the questionnaire either too quickly or inconsistently from our main sample. Respondents who participated through vlab were paid 3 €. We further recruited participants by posting invitations at several Dutch and Belgian internet game forums (e.g., InsideGamer.nl, Minatica.be). The invitation included a link to our survey created with the online survey tool SurveyMonkey (www.surveymonkey.com). As an incentive, we raffled a PlayStation 3 among participants who took part through this online link. Two hundred and ten people reacted to this invitation. Forty two of them filled in the questionnaire too quickly, incompletely and/or inconsistently and were therefore excluded from our final sample.

Participants

Our final sample consisted of 380 participants who played a game of their own choice. This group consisted of 254 men and 120 women (sex value missing = 6), with an average age of 20.8 years (range 10 to 61 years, *SD* = 5.26 years). With respect to educational level, 5% had a low education, 13% a mid level education, and 81% was highly educated. Gaming frequency varied from daily (29%), at least weekly (38%), at least monthly (13%), at least a few times per year (12%), and hardly ever (8%). Participants that never played a digital game were excluded from the current sample.

Game characteristics

The type of games participants played, were myriad. Participants filled in the full name of the game and, with the help of a game expert, we recoded those games into 12 game genres. Participants played First Person Shooter games (22%), Role Playing games (14%), Sport games (13%), Puzzle/board/card games (11%), Action adventure games (10%), Strategy games (9%), and other genres (e.g., simulation games, fight games, children's games, music games) (11%).

Most games were played on the PC (62%) or on a console (32%). A minority (6%) used another platform (e.g., handheld console or mobile phone). Further, the majority of participants played the game inside the house: in their bedroom (43%), in the living room (29%), in their study room (18%). A minority (10%) played outside the house (e.g., public transport, car, pub).

With respect to social setting; 73% played alone, 20% played online with others, and 6% played with co-located others. From the participants who played alone, 38% played against virtual others.

Game frequency of the game they reported on ranged from less than once a month (20%), monthly (18%), weekly (41%), to daily (22%). Experience with the game varied from the first time they played that game (8%), between one week and one month of experience with that game (22%), between one month and one year of experience with that game (29%), and more than one year of experience with that game (40%). The duration of the game session varied from less than half an hour (21%), between half an hour and one hour (33%), between one hour and three hours (40%), and more than three hours (6%).

Procedure

First, as outlined above, participants were invited to take part in a study on game experience. The invitation described the purpose and the procedure of the study. More concretely, we told participants that we were interested in how people experience digital gaming and that everybody, also non frequent gamers, could participate. In the invitation we further included the instruction that before opening the link to the questionnaire they had to play a game. Participants could freely chose the game they played. However, we suggested that they would best play a game in the way they normally do (with regard to the type of game, gaming platform, physical game setting). After playing the game, participants could click on a link that guided them to the online survey. This survey started with an introduction in which we described the course of the questionnaire. Next, in the introductory part, participants had to report some details about the game they just played (e.g. name of the game, gaming device, single or multiplayer, location, experience with the game, duration of the game). Subsequently, participants were asked to report on their game experience by completing the core game experience questionnaire, the social presence questionnaire, and the post-game questionnaire. The social presence questionnaire was only shown to people who had indicated that they had played the game with other people or mediated (e.g., online) others. We ended the survey with some general questions concerning people's gaming behaviour (game frequency, average duration of a game session, how much they engage in several types of single and multiplayer game settings, type of games they play, gaming motivations, proportion of leisure time spent on gaming) and some basic socio demographic questions (sex, age, education, media behaviour).

3.4. Exploratory factor analysis and scale reliability

3.4.1. Rationale for statistical procedure

Exploratory Factor Analysis (EFA) was performed to investigate the component structure of the Game Experience Questionnaire. Ideally, a theoretical model of game experience would prescribe the structure of such a questionnaire. However, although theoretical and empirical explorations

described earlier do suggest the relevance and existence of certain components, this basis is not firm enough to warrant the use of confirmatory factor analysis. In performing EFA, important decisions have to be taken on a number of methodological issues, such as, (1) what items to include, (3) what procedure to use to fit the model to the data, (3) what number of factors to choose, and (4) what rotation method to use on the initial factor solution (Fabrigar, Wegener, MacCallum, Strahan, 1999). Unfortunately, no strict rules for making these decisions exist, and most of them are a matter of continuous debate among scholars. In our analyses, we have tried to adhere closely to suggestions and advice as formulated by Fabrigar and colleagues (1999), Tabachnick and Fidell (2001), and Costello and Osborne (2005).

Ad 1: As we reported earlier, item selection was based on thorough theoretical and empirical work. Items were subsequently formulated to match the wording used by gamers – as recorded during focus groups.

Ad 2: Principle axis factoring was selected, as this technique is best suited to explore the structure of latent variables behind a collection of items measuring a multi-dimensional concept, such as game experience, in contrast to for instance PCA which is aimed to explain a maximum proportion of variance in the data.

Ad 3: Several criteria exist for determining the optimal number of factors in factor analysis. Some have proposed the Kaiser criterion, where all factors with an eigenvalue > 1 should be selected. This criterion has met with considerable criticism. An alternative is the scree-plot criterion: when eigenvalues are plotted in order of descending values in a graph, the optimal number of factors is indicated by a sharp bend or break point in the data where the curve flattens out. In addition, several goodness-of-fit indexes have been proposed for determining the optimal number of factors, Lastly, Fabrigar et al. (1999) argue: 'the decision of how many factors to include in a model is a substantive issue as well as a statistical one' (p. 281). Theoretical and conceptual considerations of multiple test runs (i.e., with different factor solutions) are therefore a valuable and valid step in the procedure (see also Costello & Osborne, 2005). The resulting model should be interpretable and theoretically sensible.

Ad 4: The last important decision to make in the EFA procedure is what type of rotation to use. The most basic choice is between orthogonal and oblique rotations. Orthogonal rotations – e.g. varimax – constrain factors to be uncorrelated. In contrast, oblique rotations – such as, oblimin – permit correlations among factors. Varimax is probably the most well-known and widely used in psychological research. However, components of game experience are very likely to be correlated with one another. If the latent variables are, in fact, correlated, then an oblique rotation will produce a better estimate. Following comments by Fabrigar et al., Tabachnik and Fidell, and Costello and Osborne, we decided to use oblimin as a first choice, but also explored the varimax rotated factor solutions, which always resulted in very similar factor solutions.

Subsequent analyses involved scale reliability analysis, with the general aim of selecting five or six strongly loading items for every factor we included in the GEQ and resulting Cronbach's alpha scores acceptable to good, i.e. 0.7 or higher. Selection of items was based on their individual factor loadings and their correlations with the scale. Items with cross-loadings above 0.30 were not selected for inclusion in the scales.

3.4.2. GEQ – core

To analyse the structure of the core part of the GEQ, exploratory factor analysis (EFA) was performed on the collected data.

Of the items originally used in the long list, some were discarded based on their lack of variance (e.g., I felt bad, It gave me a bad mood, I found it tiresome, I felt angry, I felt ashamed). The resulting set of 82 items was used in exploratory factor analysis. Considering the number of research participants, this resulted in a item-participant ratio of about 1: 5.

The initial factor analysis revealed 14 factors with an eigenvalue over 1. On the other hand, the scree plot showed that after the first five, or perhaps eight components, eigenvalues started to decrease asymptotically. We then explored the 5, 6, 7 and 8-factor solutions in more detail. The 5-factor solution was very clear and clean, but so was the 7-factor solution, which rendered two additional factors which were hypothesised and easily interpretable. As a comparison, we ran a varimax rotation on the 7-factor solution, which was almost identical and produced the same components. As subsequent scale analyses demonstrated that these latter factors produced very reliable scales, we decided upon the 7-factor solution. Several scholars argue that specifying too few factors in a model

(i.e., underfactoring) is a more severe error than specifying too many (i.e., overfactoring) (Fabrigar et al., 1999).

The 7-factor solution explained 52% of variance in the total number of items. All factors were very clear, i.e., almost self-explanatory and most items loading on only one factor (criterion .30). The seven components are:

1. Sensory and Imaginative Immersion
2. Tension
3. Competence
4. Flow
5. Negative Affect
6. Positive affect
7. Challenge.

The factor solution and scale construction presented here provide a clear and clean structure for the Game Experience Questionnaire. The structure closely resembles the structure proposed in the tentative model, with a few exceptions. First, sensory and imaginative immersion are combined into one single scale. In addition, no subscale emerged that probes experienced control. Some of the items envisaged for this scale appear in the competence scale. Competence and control are psychologically related concepts. Also, an unforeseen component emerged probing 'tension', separate from other negative affect. The last two components listed in Section 4.3 both appear in the social presence module, with an additional third component, called 'behavioural involvement.

Scores for subscales are computed as the mean score of the items in the scales. The full questionnaire is listed in Appendix 1, in its final order.

Sensory and Imaginative Immersion

The first factor that emerged is called *Sensory and Imaginative Immersion*. Originally consisting of 23 items, subsequent scale reliability analyses resulted in the selection of the following items, with a Cronbach's alpha = 0.891, and factor loadings ranging from .587 - .847. Note: as an exception, 6 items were selected, three referring more to the sensory aspect, the other three to the more imaginative part of immersion. For the remaining factors we propose five items per scale (original Dutch items are given in parentheses).

It was aesthetically pleasing (Ik vond het aansprekend van vormgeving)

I was interested in the game's story (Ik was geboeid door het verhaal van het spel)

I felt imaginative (Ik voelde me fantasierijk)

I felt that I could explore things (Ik had het gevoel dat ik de gamewereld kon exploreren)

I found it impressive (Ik vond het indrukwekkend)

It felt like a rich experience (*Ik vond het een rijke ervaring*)

Tension

The second factor that emerged was termed *Tension*. From the twelve items loading high on this factor, five were included in the final Dutch scale, and one additional item was selected as a spare item for translation purposes. The resulting scale had an internal consistency of 0.811 (Cronbach's alpha; with extra item 0.821), and factor loadings ranging between .606 and .704.

I felt tense (Ik voelde me gespannen)

I felt restless (Ik voelde me onrustig)

I felt annoyed (Ik was geïrriteerd)

I felt frustrated (*Ik was gefrustreerd*)

I felt irritable (*Ik was prikkelbaar*)

I felt pressured (*spare item*)

Competence

The third factor was termed Competence. Sixteen items had factor loadings above .30 on this factor. Again, five were selected for the Dutch scale, plus one extra for translation purposes. The resulting Cronbach's alpha = 0.826 (with extra item 0.843), and factor loadings ranging from .474 to .755.

I felt strong (Ik voelde me zeker)

I was good at it (*Ik was er goed in*)

I felt successful (Ik voelde me succesvol)

I felt skilful (Ik voelde me vaardig)

I was fast at reaching the game's targets (Ik was snel in het bereiken van de doelen in de game)

I felt competent (*extra*)

Flow

The fourth factor that emerged was *Flow*. From the thirteen items loading high on this factor, five plus one were selected. The Flow-scale has an internal consistency score of 0.866 (alpha with extra item 0.883), and factor loadings between .614 and .821.

I forgot everything around me (*ik vergat alles om me heen*)

I felt completely absorbed (*Ik was helemaal geabsorbeerd*)

I lost track of time (Ik was mijn gevoel voor tijd kwijt)

I lost connection with the outside world (*ik was weg uit de buitenwereld*)

I was deeply concentrated in the game (Ik was ten volle geconcentreerd op de game)

I was fully occupied with the game (*extra*)

Negative Affect

The fifth factor – *Negative Affect* – was actually the most problematic one. Six items loaded high on this factor, but three of them had higher cross-loadings on other factors. In fact, all factor loadings were rather low on this factor. However, we feel that this may be an artefact, created by our instructions to participants, to first play one of their own favourite games and then fill in the questionnaire. This implies that most participants will have enjoyed their game and that they knew how to play it. In preparing for the analysis, several of the items discarded due to low variance and skewed distributions probably belonged to this group of items. As in future the questionnaire will also be used for probing people's experience of games they are not familiar with and possibly dislike, having a module measuring negative affect does seem crucial. We therefore went on to explore scale reliability using two items that were originally discarded (see starred items) and found the following set with good interpretability, theoretically sensible, and with a very acceptable Cronbach's alpha value of 0.712 (0.728 with the extra item). The reliability of this subscale will be explored further with new datasets.

I thought about other things (*Ik was met andere zaken bezig*)

I felt bored (Ik voelde me verveeld)

I was distracted (*Ik was afgeleid*)

I was bored by the story (Ik vond het verhaal van het spel saai)

I found it tiresome (*Ik vond het saai*)*

It gave me a bad mood (*extra*)*

Positive Affect

The sixth scale is termed *Positive Affect* and basically probes the fun and enjoyment of gaming. Eight items loaded on this factor, of which five plus one were selected. Cronbach's alpha for the five-item scale was 0.797 (0.838 with extra item). Factor loadings varied between .430 and .728.

I felt happy (Ik voelde me vrolijk)

I felt content (Ik voelde me tevreden)

I felt good (Ik voelde me lekker)

I enjoyed it (*Ik genoot ervan*)

I could laugh about it (*Ik kon er om lachen*)

I thought it was fun (*extra*)

Challenge

The seventh and last factor is termed *Challenge*. All six items that loaded on this factor were selected for the scale. The sixth item cross-loaded on factors one and two and was therefore not selected in the final Dutch scale, but is suggested as spare item for translation purposes. Cronbach's alpha for the five-item scale is 0.745 (0.718 for the six-item scale). Factor loadings varied between .605 and .348.

I felt challenged (Ik voelde me uitgedaagd)

I thought it was hard (*Ik vond het moeilijk*)

I had to put a lot of effort into it (*Ik moest er veel moeite in steken*)

I felt that I was learning (Ik had het gevoel dat ik aan het leren was)

I felt stimulated (Ik voelde me gestimuleerd)

I felt time pressure (*extra*)

3.4.3. In-game questionnaire (iGEQ)

For validation purposes of continuous (real-time) indicators of game experience, several partners have indicated the need for a concise 'in-game' version of the GEQ. This will be used for probing Game Experience at regular intervals during the game, and for assessing multiple short-duration phases or game events in play-back mode. For the in-game questionnaire, an identical factor structure is chosen as for the core GEQ. For every component, two items were chosen which:

1. had acceptable (or better) alpha scores, and
2. semantically represented the component's item best.

Cronbach's alpha for the two-items scales varied between 0.693 and 0.80. The in-game GEQ (short form) is reported in Appendix 2.

3.4.4. Post-game questionnaire (PGQ)

The post-game module is aimed towards measuring how people feel after they have stopped playing. Since we requested participants in our study to play a game of their own choice before filling in the questionnaire, we realise that this may have created a strong bias away from negatively termed items such as, uselessness, guilt, shame, or regret. After all, they *had* to play in order to fulfil their ' job'. In line with this consideration, we found strongly skewed data on these and related items, resulting in low variances. This artefact would result in discarding quite a high – disproportionate – number of items from the analysis. We therefore decided not to discard items prior to analysis and regard the findings on this scale as preliminary. Evidently, additional research is needed to explore and validate the structure of this module As a side note: this module is not crucial to experimentation work in the project. We will however, present the preliminary results in this report.

Exploratory Factor Analaysis (EFA) was performed on the full set of 21 items probing post-game experience. Initial PFA resulted in four factors with an eigenvalue higher than one. Although the screeplot seemed to indicate two as an optimal number of factors, we selected all four, since they represented interpretable and clear factors. Subsequent oblique rotation resulted in the following structure.

The structure of the post-game module presented here should be regarded as preliminary. Future studies will further explore and test components and subscale reliabilities. The module is reported in Appendix 3.

Negative experiences

The first factor that emerged is termed *Negative experiences*. Of the eight factors loading high on this component, six items were selected for the preliminary scale, with factor loadings ranging between .591 and .762, resulting in a scale reliability of 0.832.

I found it a waste of time (*Ik vond het zonde van de tijd*)

I felt that I could have done more useful things (*Ik vond dat ik meer nuttigs had kunnen doen*)

I felt regret (*Ik had spijt*)

I felt guilty (Ik voelde me schuldig)

I felt ashamed (*Ik schaamde me*)

I felt bad (Ik voelde me rot)

Positive Experiences

The second factor was termed *Positive experiences*. Six items of the original eight were selected to form a scale with Cronbach's alpha of 0.900, with factor loadings ranging between .678 and .854.

It felt like a victory (Ik zag het als een overwinning)

I felt proud (Ik voelde me trots)

I felt powerful (Ik voelde me machtig)

I felt satisfied (Ik voelde me voldaan)

I felt revived (Ik voelde me opgepept)

I felt energised (Ik voelde me opgeladen)

Tiredness

The third factor only consisted of two items with factor loadings above .40 (.926 and .676 respectively), and a Cronbach's alpha score of 0.764. The two items are:

I felt exhausted (Ik voelde me uitgeput)

I felt weary (Ik voelde me vermoeid)

Returning to Reality

The fourth and last factor consisted of three items with factor loadings between .404 and .702. These three items had an internal consistency (Cronbach's alpha) of 0.619. This factor, called *Returning to reality*, consisted of the following items:

I found it hard to get back to reality (Ik had moeite om terug te keren naar de realiteit)

I felt disoriented (Ik voelde me gedesorienteerd)

I had a sense that I had returned from a journey (*Ik had het gevoel dat ik van een reis terugkeerde*)

3.4.5. GEQ – social presence module

Although a number of items used in the social presence module were taken from or inspired by the Networked Minds measure of Social Presence (Biocca et al., 2001), some items were new to this scale. In addition, the application area of gaming differs significantly from the application area. We therefore decided to also perform exploratory factor analysis (EFA) on this set of items.

Of the 25 items originally used in the long list, again five had to be discarded based on their lack of variance. The resulting set of 20 items was used in exploratory factor analysis. The sample size for this analysis was N=212, which is smaller than in the previous analyses, as it only consisted of participants who indicated to have played with or against virtual, mediated, or co-located others. This resulted in a item-participant ratio of about 1: 10.

The initial factor analysis revealed five factors with an eigenvalue higher than one. However, the scree plot showed a clear bend after the third factor. The three-factor solution was very clear, and

explained 56% of variance. The solution resulted in two psychological involvement components (Empathy and Negative feelings), and a behavioural involvement component.

The resulting structure is quite different from the one in Biocca et al's networked minds measure of social presence. This is not surprising as the application domains differ considerably. For instance, the very act of playing a game together implies that players' actions are dependent on each other, even in conditions of minimal 'media richness'. Behavioural interdependence is one of the hard to reach targets in application domains such as teleconference systems.

Together these items make up the social presence module, which is only to be used in gaming situations where players played with or against others, be they virtual (e.g., in-game characters), mediated (e.g., online), or co-located. The module is reported in Appendix 4.

Psychological Involvement - Empathy

The first factor that emerged is called *Psychological Involvement - Empathy*. Six of the seven items loading on this factor were selected for the scale, with a Cronbach's alpha = 0.886, and factor loadings ranging from .683 - .826.

I empathized with the others (*Ik leefde mee met de anderen*)

When I was happy, the others were happy (*Wanneer ik blij was, waren de anderen blij*)

When the others were happy, I was happy (*Wanneer de anderen blij waren, was ik blij*)

I felt connected to the others (Ik voelde me verbonden met de anderen)

I admired the others (Ik bewonderde de anderen)

I found it enjoyable to be with the others (*Ik vond het gezellig met de anderen*)

Psychological Involvement – Negative feelings

The second component of *Psychological Involvement* is called *Negative feelings*. Again, six of the original seven items were selected for the scale, which has an internal consistency (alpha) of 0.860, and factor loadings between .511 and .855.

I was influenced by the other's moods (Ik werd beïnvloed door de stemming van de ander)

I influenced the mood of the other (Ik beinvloedde de stemming van de ander)

I felt revengeful (*Ik was wraakzuchtig*)

I felt schadenfreude (malicious delight) (*Ik had leedvermaak*)

I felt jealous of the other (*Ik was jaloers op de ander*)

Behavioural Involvement

The third scale is termed *Behavioural Involvement.* All five items that loaded on this factor were selected for the scale, with a Cronbach's alpha of 0.711, and factor loadings between .338 and .774. The items are:

My actions depended on the other's actions (Mijn handelingen hingen af van de handelingen van de ander)

The other's actions were dependent on my actions (De handelingen van de ander hingen af van mijn handelingen)

What the others did affected what I did (*Wat de ander deed beïnvloedde wat ik deed*)

What I did affected what the other did (*Wat ik deed, beïnvloedde wat de ander deed*)

The other paid close attention to me (*De ander lette op mij*)

I paid close attention to the other (*Ik lette op de ander*)

3.5. Conclusion

In this chapter the rationale for questionnaire construction and item generation were discussed. Subsequently, the methodology and statistical analyses and results of the empirical investigation into the structure of the Game Experience Questionnaire (GEQ) were described.

The Game experience questionnaire is developed with a modular structure, consisting of:

1. The core questionnaire (GEQ). This is the heart of the Game Experience Questionnaire, probing multiple components of players' experience while gaming.
2. The post-game questionnaire (PGQ), probing gamers' experience after the gaming session and any after effects.
3. The social presence module (SPGQ), probing gamers' experience of and involvement with their co-player(s).

These three lists are to be administered after the gaming session has ended. Additionally, a short in-game version of the GEQ was developed, the iGEQ, for probing in-game experience multiple times during a gaming session.

For the core GEQ, the factor solution and scale analyses provided a clear and clean structure of seven components (1) Sensory and Imaginative Immersion, (2) Tension, (3) Competence, (4) Flow, (5) Negative Affect, (6) Positive affect, and (7) Challenge. These subscales, each reliably measuring unique components of game experience, have 5 or 6 items, and a good average Cronbach's alpha score of .81.

Scale reliabilities of a short, in-game version of the GEQ for validation purposes of continuous indicators of game experience were also explored. The iGEQ's factor structure is identical to that of the core GEQ, but consists of only 2 items per scale, selected to best represent the psychological concept. Results showed that these scales were also reliable, with internal consistencies varying between satisfactory and good.

The post-game module (PGQ), aimed towards measuring how people feel after they have stopped playing, consists of four subscales. As the analysis did not unanimously and clearly indicate one optimal structure for this module, the structure presented here should be regarded as preliminary. Future studies will further explore and test components and subscale reliabilities. The next chapter will test their validity and sensitivity.

Dimensionality analysis of the social presence gaming questionnaire (SPGQ) resulted in three subscales: (1) Psychological involvement – Empathy, (2) Psychological Involvement – Negative feelings, and (3) Behavioural involvement. As gaming presents a different application domain from the domains targeted Biocca et al. (2001), the resulting structure differs from the one presented there. Reliabilities of the scales are good, average alpha .82. The scales are only to be used in gaming situations where players played with or against others, be they virtual (e.g., in-game characters), mediated (e.g., online), or co-located.

Scores for all subscales are computed as the mean score of the items in the scales. The full questionnaires are listed in Appendix 1 - 4. Scoring guidelines are reported in Appendix 5.

The results of the analyses presented in the present chapter are promising, as they have mostly rendered reliable and easily interpretable scales, matching theoretical conceptualisation and qualitative empirical insights. Sensitivity and validity of the scales constructed and discussed here will be explored in the following chapter.

4. Explorations of the GEQ: Inter-correlations and basic sensitivity

Although the structure of the GEQ presented in Chapter 3 is clear and easily interpretable, internal consistencies of all subscales are acceptable or better, and both face validity and content validity are high, additional explorations are needed to check whether the developed measure is sensitive enough to pick up differences between games, game characteristics, gamers, and game circumstances. Part of this work will be performed as an integral part of the experimental work planned in the FUGA project (see D3.1). However, the present chapter describes preliminary explorations, based on the data gathered in the survey study reported earlier. In each section below, basic sensitivity and discriminant validity of subscales in each of the GEQ modules are reported and discussed, based on background data that was gathered from respondents.

4.1. Core Game Experience Questionnaire (GEQ)

The core GEQ consists of seven subscales, each measuring unique, though sometimes related components of game experience. Basic descriptives for the subscales are reported in Table 1. Scores on most scales were moderate and ranged the full scale. The highest scores appeared on Positive Affect and Immersion. Scores for Negative affect and Tension were rather low.

Table 1. Basic descriptives for GEQ subscales.

	Mean	**SD**	**Min**	**max**
Positive affect	2.55	.73	.00	4.00
Competence	1.72	1.12	.00	4.00
Immersion	2.28	.80	.00	4.00
Flow	1.63	1.02	.00	4.00
Challenge	1.70	.79	.00	4.00
Negative affect	.57	.60	.00	3.40
Tension	.88	.77	.00	3.80

Note: scales range from 0 to 4

Correlations between the subscales are reported in Table 2. As could be expected, there are a number of significant correlations between subscales of the GEQ. The highest correlations (close to .5) exist between Positive Affect, Competence, and Immersion. Flow is correlated moderately with Competence, Challenge, and Immersion. Negative Affect is negatively correlated with all subscales but the Tension scale. This Tension scale has modest (.2 -.3) positive correlations with Challenge, Flow, and Negative Affect, and a small negative correlation with Positive Affect.

Sensitivity of Core GEQ subscales will be explored by investigating gender differences, differences between frequent and infrequent gamers, type of game played, and social setting.

Table 2. Bivariate correlation between GEQ subscales.

	Positive affect	Immersion	Competence	Flow	Challenge	Negative affect	Tension
Positive affect		.467*	.514*	.278*	.340*	-.462*	-.146*
Immersion	.467*		.519*	.488*	.472*	-.432*	.054
Competence	.514*	.519*		.437*	.395*	-.313*	.053
Flow	.278*	.488*	.437*		.454*	-.317*	.217*
Challenge	.340*	.472*	.395*	.454*		-.285*	.317*
Negative affect	-.462*	-.432*	-.313*	-.317*	-.285*		.162*
Tension	-.146*	.054	.053	.217*	.317*	.162*	

Note: N = 380; * = p< .01

Gender differences

About one-third of participants in the study's sample are female. Analyses of variance showed that female players report significantly lower scores (p<.001) on every single subscale of the GEQ, except for Negative Affect, where females score higher, and Tension, where there is no significant difference. Mean scores are reported in Table 3.

Table 3. Gender differences on GEQ subscales

	Positive affect	Immersion	Competence	Flow	Challenge	Negative affect	Tension
Female	2.32 (.67)	1.09 (.93)	1.96 (.81)	1.30 (1.03)	1.39 (.77)	0.77 (.64)	0.79 (.78)
Male	2.67 (.71)	2.02 (1.08)	2.45 (.74)	1.80 (.97)	1.84 (.76)	0.48 (.56)	0.92 (.75)
Total	2.56 (.72)	1.72 (1.12)	2.29 (.80)	1.64 (1.02)	1.70 (.79)	0.57 (.60)	0.88 (.76)

Note: N females = 120, N males = 254.

Frequent vs. infrequent gamers

We explored game experience as a function of how often participants played the game they reported on. Positive Affect, Immersion, Competence, Flow, and Challenge all increased with play frequency, as would be expected for all scales, especially the first four. Scores are visualised in Figure 1. Analyses of variance showed that these differences were significant (for Challenge, p=.005 all remaining p<.001). Subsequent contrast analyses even indicated that for Positive Affect, Immersion, Competence and Challenge, all differences between all groups were significant at the .05 level. Negative Affect showed an inverse relation with play frequency. Reported tension was not related to play frequency.

Type of game

We subsequently explored whether game experience dimensions varied as a function of the type of game participants played. For most game experience components, analyses of variance yielded significant differences (all p < .05) for Type of game played. Scores for each game type on the game experience dimensions, are visualised in Figure 2. Contrast analyses were performed to test specific differences between game types.

Positive affect showed significant differences between game types, F(5, 302)= 4.89, p<.001, Puzzle games scored significantly lower than all other types of games. The reverse was true for Negative Affect, F(5, 302)= 10.36, p<.001. Similar patterns emerged for Competence F(5, 302)= 2.59, p=.03, for which Puzzle games scored lower than FPSs and strategy games, and Flow, F(5, 302)= 2.86, p<.02.

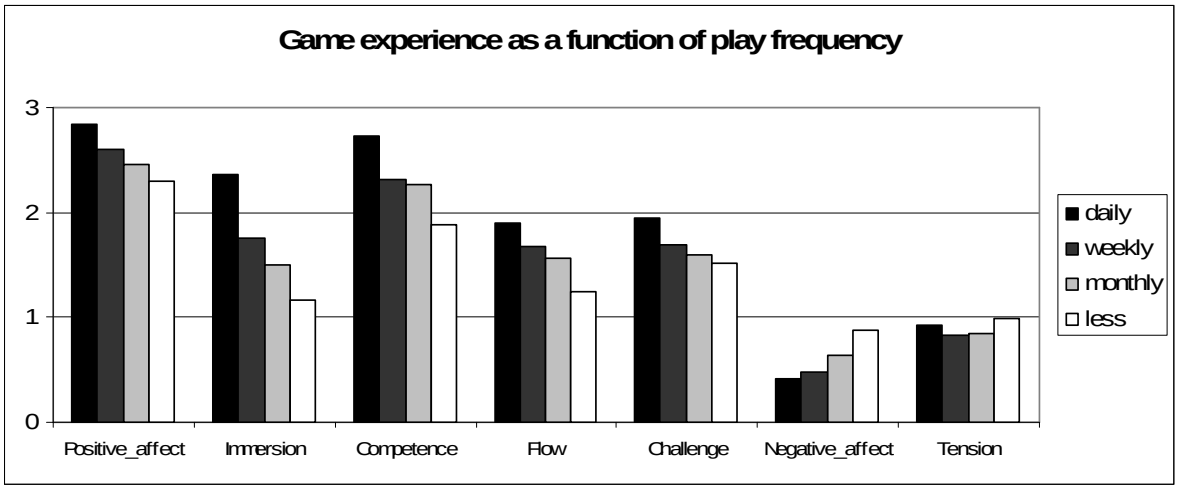

Figure 1. Game experience on GEQ subscales as a function of play frequency (this particular game)

Game type showed significant effects on Imaginative and Sensory Immersion, $F(5, 302)= 32.20$, $p<.001$. Action adventure games scored significantly higher than all other types of games. Role playing games (RPG) and first person shooter games (FPS) did not differ from each other but scored higher than sports games, strategy games and puzzle games. Subsequently, both sports and strategy games scored higher than puzzle games.

Action adventure, sports and FPS games scored highest on the Challenge dimension, $F(5, 302)= 7.00$, $p<.001$. Action adventure games differed significantly from all remaining types of games (RPG, puzzle and strategy); FPS and Sports games differed only from puzzle games.

Post hoc tests revealed no significant effects between the different game types on the Tension subscale, $F(5, 302)= 2.4$, $p=.04$.

Note: the number of participants for every type of game was rather low (N varied from 34 to 87) which could explain why some post hoc tests failed to reach significance. Nevertheless, there are some substantial differences between the different types of games that certainly deserve further exploration.

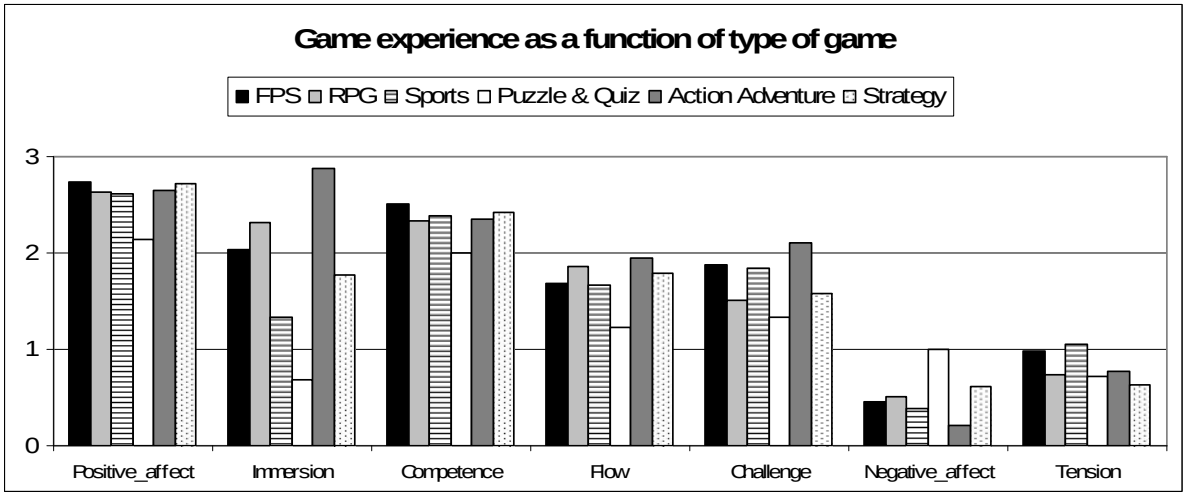

Figure 2. Game experience on GEQ subscales as a function of type of game

Social setting

To further explore the sensitivity and discriminant validity of the subscales, GEQ scores were compared for various social settings in which participants played. Based on participants' responses on two scales form the background questionnaire we constructed a categorisation of increasing social presence and relatedness. The categories describe settings in which the 'distance' (physically & psychologically) between players and co-players gets smaller. The 5 categories are:

(1) playing alone,
(2) playing with virtual others (i.e., in-game characters),
(3) playing online with unknown others,
(4) playing online with friends/family,
(5) with the co-player physically present.

The analyses are performed on a slightly smaller dataset than the original one, as the categorisation could not be unambiguously made for 27 participants. As the number of conflicting answers on the two scales used was rather high (27 out of 380), the results should be regarded with some caution. Some ambiguity may have existed in these scales, possibly weakening the reliability of this variable.

GEQ scores for the five different groups are visualised in Figure 3. All subscales except Flow and Tension showed significant differences between social settings. Moreover, score patterns for the subscales differed, indicating the distinctiveness of the components measured (i.e., discriminant validity).

Positive affect differed significantly between groups, $F(4, 348)= 7.04$, $p<.001$, with significant contrasts between the 1st and 2nd, the 3rd and 4rth, and the 4th and 5th category. Participants reported higher enjoyment with each subsequent social setting of increased social presence and relatedness.

Sensory and Imaginative Immersion showed a significant effect of social setting, $F(4, 348)= 2.36$, $p=.05$. Interestingly, contrast analyses indicated that this should be attributed solely to a significant difference between the 1st and 2nd category. In both groups, participants played alone, either without (group 1) or with (group 2) in-game characters. No significant contrasts emerged between remaining categories. Perhaps the effect is caused by the different types of games played by gamers in these two groups. For instance, participants in the first group more frequently played puzzles – which are generally experienced as less immersive – whereas participants in the second group more frequently played FPS and sports games, which are generally experienced as highly immersive (see Figure 2).

The Competence subscale differed significantly between categories, $F(4, 348)= 2.96$, $p=.02$. Contrast analyses showed that playing alone resulted in significantly, or marginally significant lower scores than all other types of social settings, where there was some (virtual, mediated, or co-located) social identity involved in the game.

No significant differences emerged on the Flow subscale.

Differences did emerge on the Challenge subscale, $F(4, 348)= 3.62$, $p=.007$. Participants playing alone reported lower challenge than those playing against virtual others (group 2), mediated friends (group 4), or co-located friends (group 5). This last group also scored significantly higher than those competing with online strangers (group 3).

Significant effects also appeared on the Negative affect subscale, $F(4, 348)= 4.31$, $p=.002$. Negative affect of those playing against virtual others was lower than of those playing alone (group 1), with online strangers (group 3) or with co-located friends (group 5). We have no explanation as to why this effect occurred.

Lastly, the Tension subscale showed no significant effects of social setting.

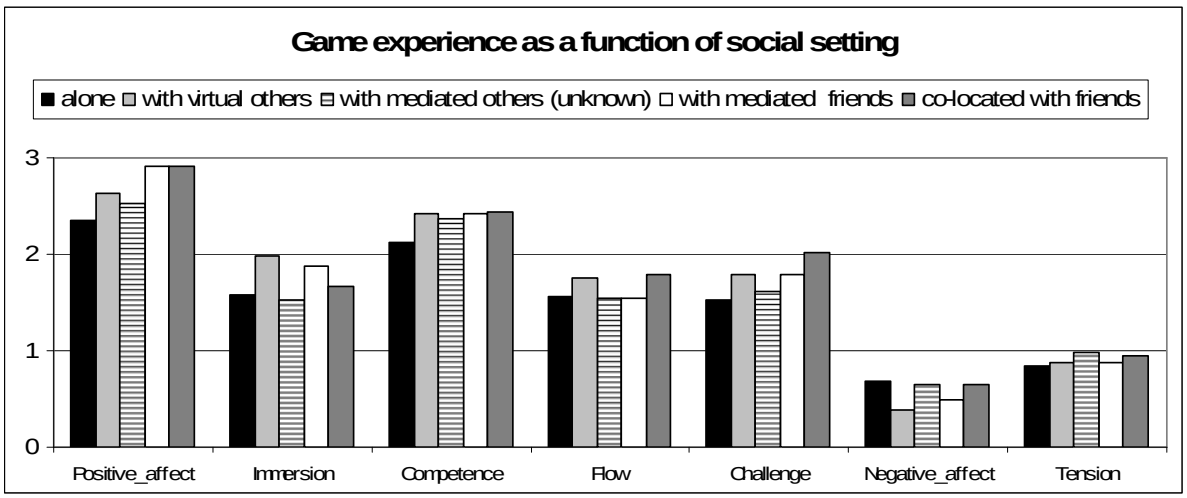

Figure 3. Game experience on GEQ subscales as a function of social setting

4.2. In-game Questionnaire (iGEQ)

The in-game experience questionnaire (iGEQ) has the exact same component structure as the core questionnaire, but only consists of two of the five or six items for every component. We repeated some of the analyses reported above and compared them, to check whether the sensitivity of this shortened version of the GEQ is still sufficient. Although, logically, the standard deviations are somewhat bigger than for the sub scales of the core questionnaire, the results are very similar.

Gender differences

Female players report significantly lower scores ($p<.001$) on every single subscale of the iGEQ, except for Negative Affect, where females score higher, and Tension, where there is no significant difference. Results are reported in Table 4.

Table 4. Gender differences on iGEQ subscales

	Positive affect	Immersion	Competence	Flow	Challenge	Negative affect	Tension
Female	2.39 (.75)	1.08 (1.15)	1.74 (.99)	1.20 (1.04)	1.80 (.96)	0.64 (.72)	0.68 (.79)
Male	2.77 (.78)	2.17 (1.32)	2.30 (1.32)	1.64 (1.09)	2.44 (1.00)	0.35 (.64)	0.83 (.90)
Total	2.65 (.79)	1.81 (1.37)	2.12 (1.00)	1.49 (1.09)	2.24 (1.03)	0.44 (.68)	0.78 (.87)

Note: N females = 120, N males = 254.

Frequent vs. infrequent gamers

As with the core GEQ, game experience was explored as a function of how often participants played the game they reported on. Again, Positive Affect, Immersion, Competence, Flow, and Challenge all increased with play frequency, as is indicated in Figure 4. Negative Affect showed an inversed relation with play frequency. Analyses of variance showed that these differences were significant (all $p<.01$). Reported tension was not related to play frequency.

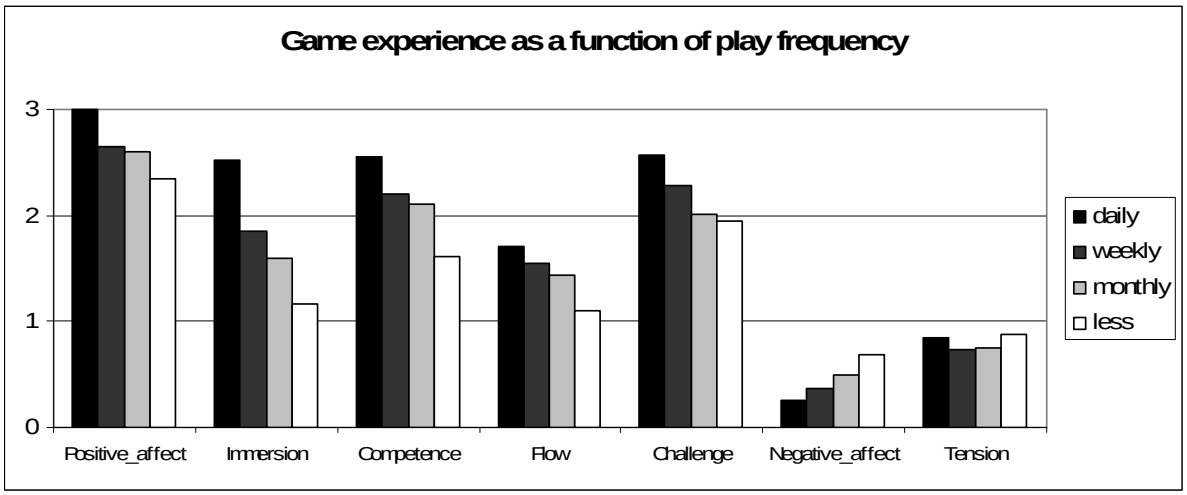

Figure 4. Game experience on iGEQ subscales as a function of play frequency (this particular game)

Type of game

We subsequently explored whether game experience dimensions varied as a function of the type of game participants played. For most game experience components, analyses of variance yielded significant differences (all $p < .05$) for Type of game played, identical to results on the GEQ. Scores for each game type on the game experience dimensions, are visualised in Figure 5.

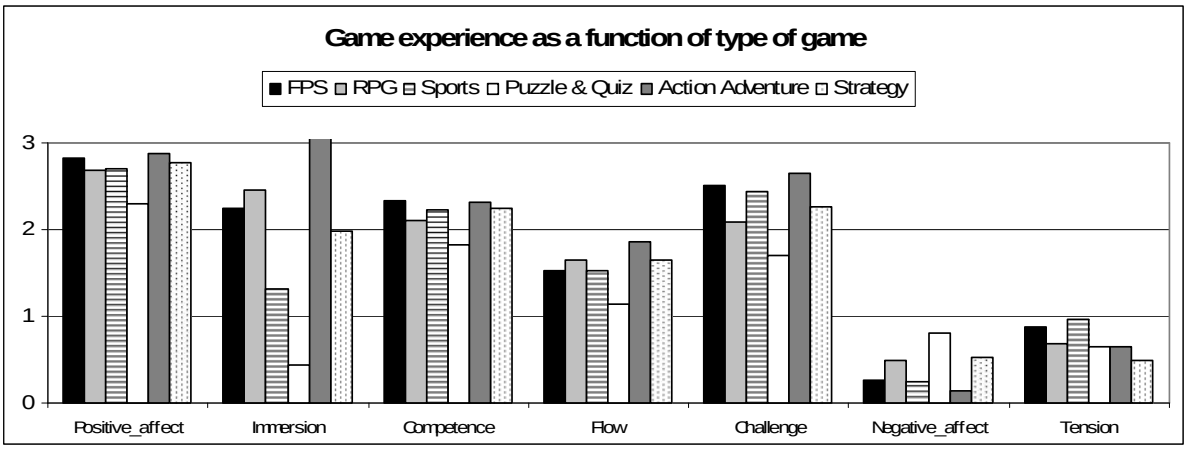

Figure 5. Game experience on iGEQ subscales as a function of type of game

Social setting

As with GEQ scores, iGEQ scores were compared for various social settings. The scores are visualised in Figure 6. Again response patterns were very similar to those reported for the GEQ scores. All subscales except Flow and Tension showed significant differences between social settings.

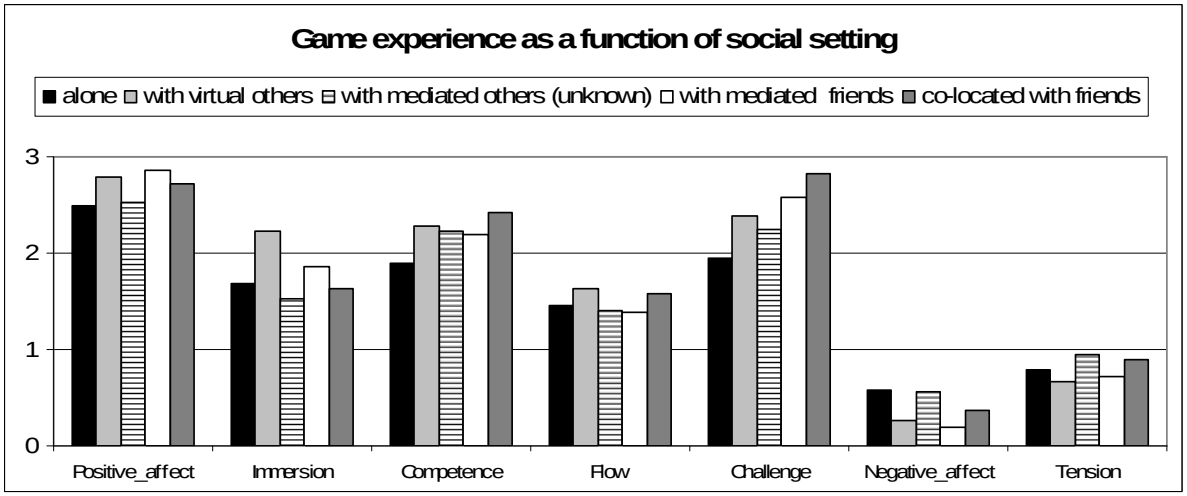

Figure 6. Game experience on iGEQ subscales as a function of social setting

4.3. Post-game Questionnaire (PGQ)

The basic descriptives for the post-game module are reported in Table 5. The scores for Positive experience are moderate, the scores on all remaining scales are low. Correlations between the subscales are modest, see Table 6.

Table 5. Descriptive statistics for Post-game subscales.

	Mean	SD	min	max
Positive experience	1.58	1.04	0.0	4.0
Negative experience	.51	.62	0.0	3.3
Tiredness	.40	.70	0.0	4.0
Return to reality	.34	.60	0.0	4.0

Note: scales range from 0 to 4

Table 6. Bivariate correlations between Post-game subscales.

	Positive experience	Negative experience	Tiredness	Return to reality
Positive experience		-.248*	.065	.339*
Negative experience	-.248*		.338*	.198*
Tiredness	.065	.338*		.331*
Return to reality	.339*	.198*	.331*	

Note: N = 380; * = p< .001

Bivariate correlations were explored between the subscales of the Core GEQ, measuring experience while gaming, and the PGQ, measuring the players' experience after having stopped gaming. The results are reported in Table 7. A number of interesting findings emerged. For one, it is clear that, although both questionnaires were employed after the game, subscales of the two lists do measure different experiences. This indicates that participants are indeed able to discern between the two phases. Second, a positive post-game experience correlates stronger with in-game experiences of competence and immersion (.64 and .61) than with positive affect (.51). It also shows high correlations with in-game experiences of flow (.49) and challenge (.52). Lastly, results indicated that returning to reality after gaming is harder after having experienced high levels of flow (.46) and immersion (.41), which is in line with theoretical readings on these psychological concepts.

Table 7. Bivariate correlation between Core GEQ and post-game (PGQ) subscales.

	Positive affect	Immersion	Competence	Flow	Challenge	Negative affect	Tension
Negative experience	-.40***	-.25***	-.23***	-.06	-.20***	.61***	.25***
Positive experience	.51***	.61***	.64***	.49***	.52***	-.30***	.17***
Tiredness	-.14**	.05	.01	.22***	.16**	.20***	.37***
Returning to reality	.10	.41***	.22***	.46***	.26***	.00	.17***

Note: N = 375; * p< .05, **p<.01, ***p<.001

Gender differences

Several significant gender differences were found. They are reported in Table 8. In line with the findings on the Core GEQ, Female participants scored lower on Positive experience and Return to reality, and higher on Negative experience. There was no significant difference on the Tiredness scale.

Table 8. Gender differences on Post-game subscales

	Positive experience	Negative experience	Tiredness	Return to reality
Female	1.14 (.89)	0.66 (.62)	0.36 (.70)	0.16 (.33)
Male	1.79 (1.04)	0.66 (.62)	0.41 (.70)	0.41 (.63)
Total	1.58 (1.04)	0.51 (.62)	0.40 (.70)	0.33 (.58)

Note: N females = 120, N males = 254.

Frequent vs. infrequent gamers

Post-game experience was explored as a function of how often participants played the game they reported on. In line with expectations, positive experience increased with play frequency, F(3, 371)=

13.90, p<.001, while negative experience decreased, F(3, 371)= 5.68, p=.001. Returning to reality was hardest for frequent players, F(3, 371)= 7.10, p<.001. The tiredness scale did not show significant effects, F<1, NS.

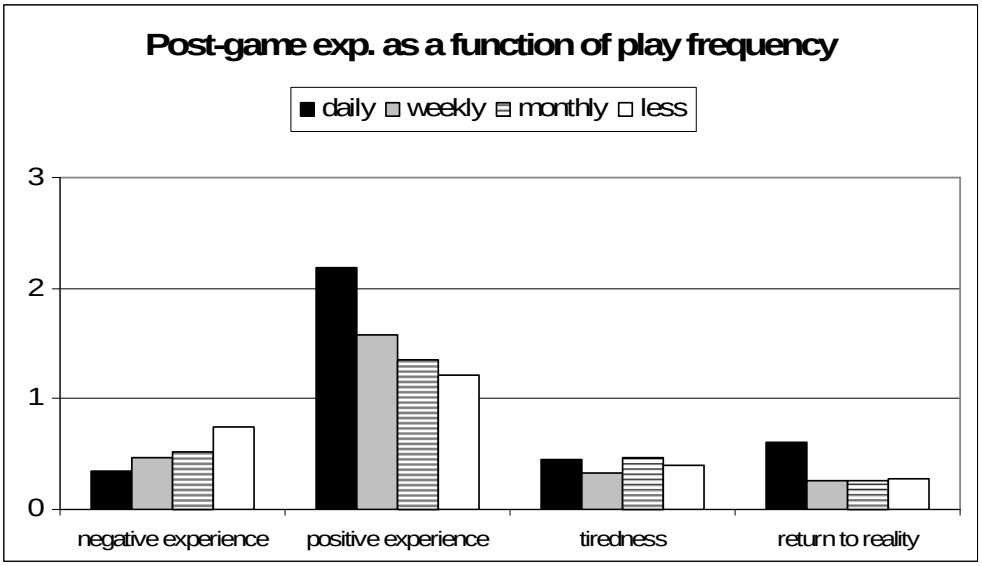

Figure 7. PGQ subscales for different play frequencies

Type of game

PGQ subscales were explored for different types of games (see Figure 8). Negative experience after playing differed significantly between games, F(5, 299)= 3.93, p=.002. The most negative experiences (though still very low) were reported for puzzles and quizzes, least negative for Action adventure games. Positive experience, F(5, 299)= 4.62, p<.001, was highest for FPS, lowest for puzzles. The remaining two scales did not show significant differences between games.

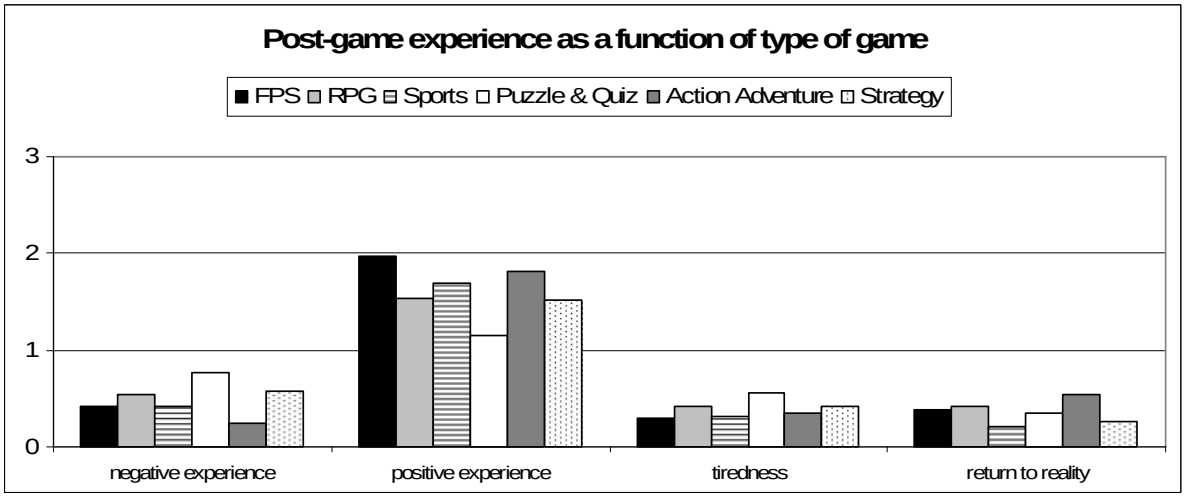

Figure 8. PGQ subscales for different types of games

Social setting

PGQ subscales also showed significant differences between different social settings (Figure 9). Players who had played alone reported the most negative experience, F(4, 346)= 3.08, p=.016. Players who played with co-located other(s) reported the highest positive experience, F(4, 346)= 6.70, p<.001.

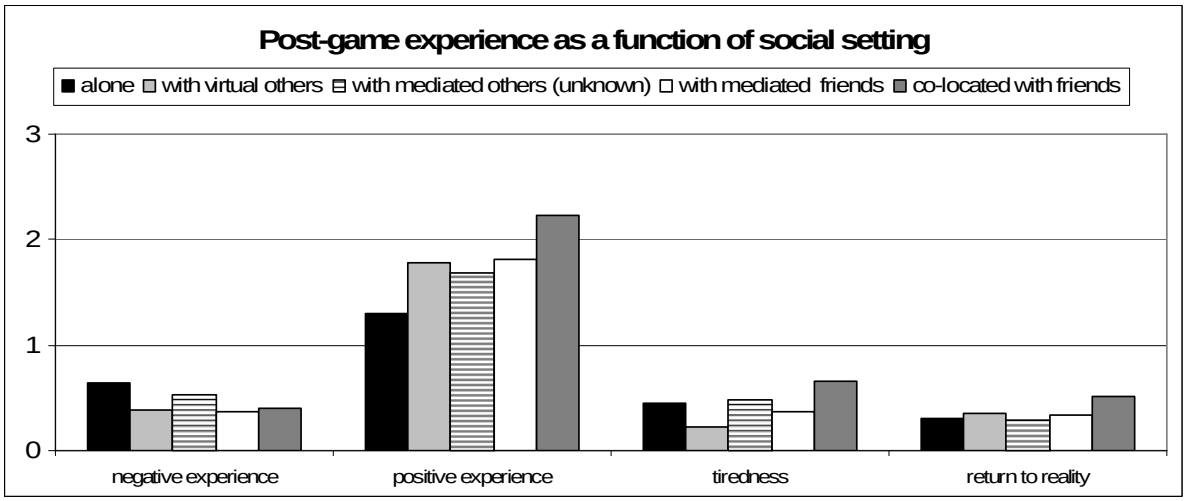

Figure 9. PGQ subscales for different social settings

Play duration

Lastly, post-game experiences were explored as a function of play duration (how long participants had gamed before filling in the questionnaire, Figure 10). Perhaps surprisingly, Tiredness did not show significant differences between groups. Perhaps this indicates that games do not cause fatigue, or that the tiredness scale is not sensitive enough. Note: This scale did not result in any significant differences on the ANOVAs performed; the scale did show a few modest but significant correlations with other scales.

Furthermore, positive post-game experience was higher for participants who had played longer, $F(3, 371)= 24.67$, $p<.001$, and their scores were lower on negative post-game experience, $F(3, 371)= 16.48$, $p<.001$. Returning to reality was harder, the longer participants had played, $F(3, 371)= 4.24$, $p=.006$.

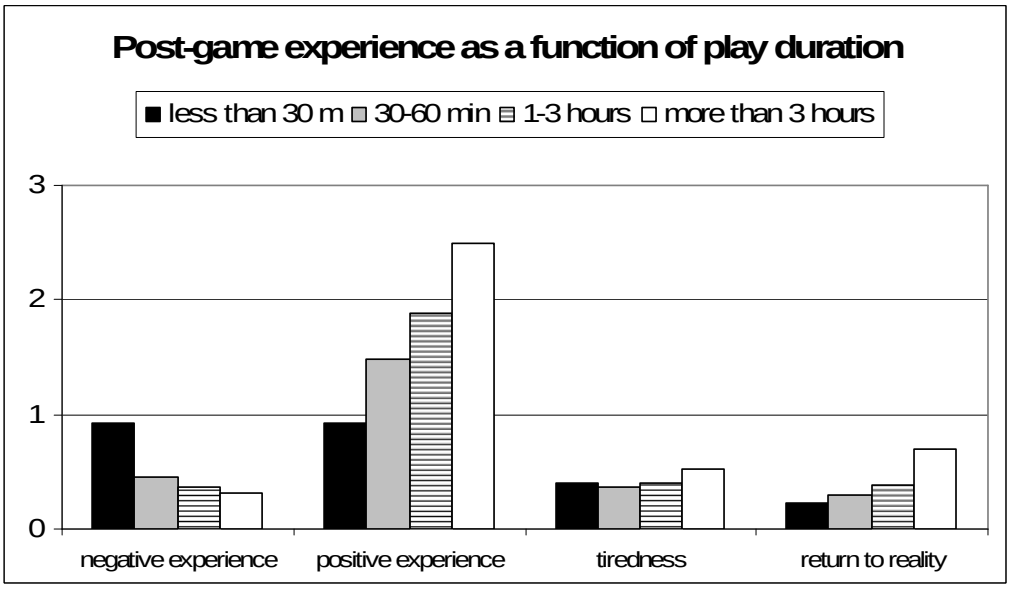

Figure 10. PGQ subscales as a function of play duration

4.4. Social Presence in Gaming Questionnaire (SPGQ)

The Social Presence module consists of three subscales. Empathy and Negative feelings measure components of psychological involvement, the third scale measures Behavioural Involvement. These scales have only been filled in by participants who played against a social entity (virtual, mediated, or co-located, N=185). Basic descriptives are given in Table 9.

Table 9. Descriptive statistics for Social Presence subscales

	Mean	SD	min	max
Empathy	1.41	1.03	.00	4.00
Negative feelings	.88	.73	.00	3.20
Behavioural	2.12	.94	.00	4.00

Note: scales range from 0 to 5.

There were significant correlations between the subscales. The correlation between Empathy and Negative feelings was strongest, r= 0.45, p<.001. Note that this correlation is positive, indicating that participants reporting more positive feelings towards their co-players also reported more negative feelings. This illustrates that both measures are indicators of psychological involvement: the more 'socially present' the other person is, the stronger their mutual influence on each other's feelings, both positively and negatively toned. The correlations of these scales with behavioural involvement were fairly low: r=.20 for empathy, and r=.28 for negative feelings.

Bivariate correlations were explored between the subscales of the Core GEQ and social presence subscales. The results are reported in Table 10. Overall, correlations were modest, but still a number of interesting findings emerged. All the competence and challenge subscales of the GEQ showed the strongest relationships with all three subscales of social presence. Game experiences of competence and challenge are associated with higher psychological involvement between co-players, both as regards positive and negative feelings, and higher behavioural involvement. In addition, negative feelings towards the co-player (jealousy, revenge, influencing each others' moods) were most strongly related to reported tension, whereas positive feelings towards the co-player (empathy) was correlated highest with reported positive affect and immersion.

Table 10. Bivariate correlation between Core GEQ and Social presence subscales.

	Positive affect	Immersion	Competence	Flow	Challenge	Negative affect	Tension
Empathy	.31***	.28***	.23***	.10	.22***	.00	.10
Negative feelings	.13*	.16*	.22***	.24***	.30***	.14*	.36***
Behavioural involvement	.16*	.10	.26***	.18**	.27***	-.03	.14*

Note: N = 224; * p< .05, **p<.01, ***p<.001

Social setting

The social presence subscales were explored for various social settings. The results are shown in Figure 11. Significant differences were found for both psychological involvement scales. Empathy increased with increasing social presence and relatedness, F(3, 181)= 21.83, p<.001. Contrast analysis showed that scores increased significantly (p<.001) with each subsequent category. Although modest for all groups, negative feelings also differed significantly between settings, F(3,181)= 5.38, p=.001. The contrast between the 1st and 2nd category (virtual vs. mediated strangers) did not reach significance, but all remaining contrasts did, indicating higher scores for negative feelings with increasing social presence and relatedness. Behavioural involvement was equal for all social settings: even for settings with low social presence and relatedness, players' behaviour is influenced by the other social entity's behaviour. This makes a lot of sense in view of the present application area: in a game, the player's and opponent's actions are interdependent and make up the very core of the activity.

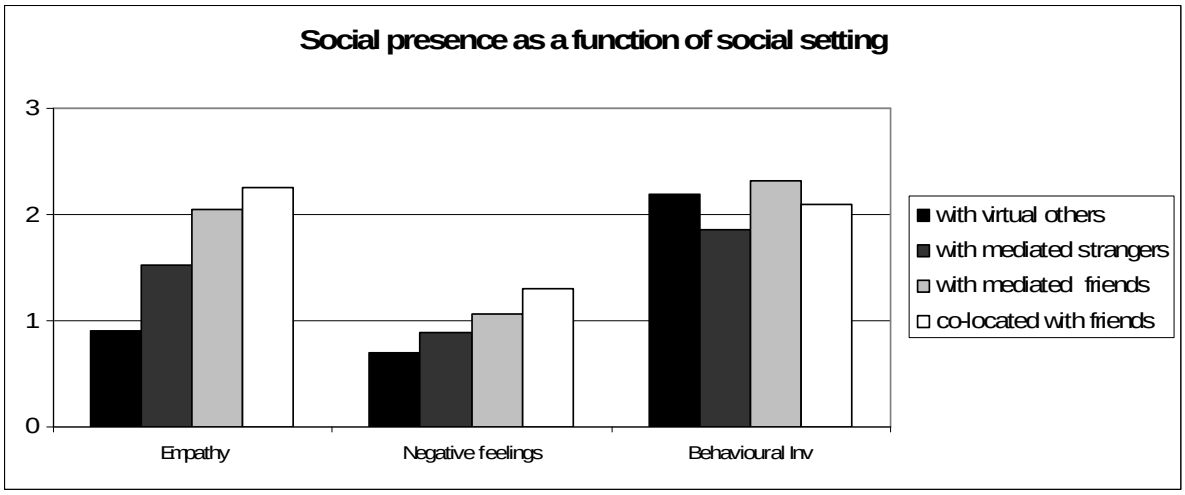

Figure 11. Social presence subscales for different social settings

4.5. Conclusion

The present chapter presented statistical explorations to check sensitivity and validity of the newly developed self-report measures. For each scale, average scores and response ranges were calculated, and bivariate correlations reported. Subsequently, response patterns were explored for different types of gamers (male vs. female, and frequent vs. infrequent gamers), different game types, and social settings.

Average scores on subscales of the core GEQ were moderate (as planned) and ranged the full scale. Exceptions to this are negative affect and tension, for which the average scores were somewhat low. The highest scores appeared on positive affect and immersion. Analyses of bivariate correlations indicated that, although related, the seven subscales do measure different components of experience. Discriminant validity was further demonstrated in subsequent analyses of variance (ANOVAs), where components each showed unique response patterns for gamer characteristics, game types and social settings.

Male players reported more game enjoyment overall than female players, as indicated by higher scores for positive affect, competence, flow, immersion, and challenge, and lower scores for negative affect. Similarly, frequent players reported higher enjoyment than infrequent players, demonstrated in almost linear relationships between scores on GEQ scales and play frequency. Significant differences were also found for game type, based on post-hoc categorisation. Playing puzzles and quizzes was generally reported as less involving and enjoying than other types of games. Action adventure games scored particularly high on immersion, followed by FPS and role playing games. Lastly, analyses of various social settings showed that with increasing social presence and relatedness to others both positive affect and challenge increased. These findings are in line with a-priori expectations and earlier findings and thus corroborate validity of the measures.

Although standard deviations are somewhat bigger for the iGEQ than for GEQ subscales, statistical analyses rendered very similar results. In spite of its conciseness, sensitivity and validity of this measure are still highly satisfactory.

For PGQ scales the average score for the positive experience subscale was moderate, the scores on all remaining scales were low. Bivariate correlations between GEQ and PGQ subscales clearly indicated that both scales measure different experiences. High scores on positive post-game experience correlated with in-game positive affect, but even stronger with competence and immersion. Feeling good afterwards thus mainly depends on whether the experience was rich and whether players felt competent in what they were doing. Flow and immersion experienced during the game were also related to experiencing the game a positive experience afterwards. Returning to reality after gaming was harder after having experienced high levels of flow and immersion.

Female participants scored lower on positive experience and returning to reality, and higher on negative experience, which corroborates the gender differences found on the GEQ. Similarly, positive experience and difficulty with returning to reality increased with play frequency and play duration, while negative experience decreased, in line with GEQ findings of higher enjoyment and involvement.

Experience after playing was most positive after FPSs, most negative after puzzles and quizzes. Furthermore, results indicated that players who had played alone reported the most negative experiences, whereas those playing with co-located other(s) reported the highest positive experiences.

SPGQ subscales showed low to moderate average scores. The two psychological involvement scales, measuring the experience and contagion of positive (empathy) and negative feelings, were positively correlated, illustrating that both measures are indicators of social presence: the more 'socially present' the other person is, the stronger their mutual influence on each other's feelings, both positively and negatively toned.

In line with expectations, psychological involvement (both empathy and negative feelings) increased with social presence and relatedness. Behavioural involvement was equal for all social settings: even for settings with low social presence and relatedness, players' behaviour is influenced by the other social entity's behaviour. This result may well be unique to the present application area: in a game, the player's and opponent's actions are interdependent and make up the very core of the activity. This is not necessarily the case in other social presence technology applications.

The SPGQ experiences – psychological involvement between co-players, both as regards positive and negative feelings, and higher behavioural involvement – were associated with higher GEQ experiences of competence and challenge. In addition, negative feelings towards the co-player (jealousy, revenge, influencing each others' moods) were most strongly related to reported tension, whereas positive feelings towards the co-player (empathy) was correlated highest with reported positive affect and immersion.

Summing up these findings, statistical analyses firmly demonstrated that the GEQ and additional modules were sensitive enough to pick up differences between gamers, game types, play characteristics, and social context of play. Moreover, each subscale's unique value (discriminant validity) has been demonstrated by different response patterns to variations in these background variables.

5. Discussion and conclusion

5.1. Questionnaire development

The general aim in developing the GEQ was to construct a reliable, valid and sensitive, comprehensive measure of game experience. The theoretical work performed by all partners in the FUGA project (Deliverable 2.1) provided a firm and broad basis for the development of the Game Experience Questionnaire. Additional relevant scientific literature was reviewed for any conceptualisations that might complement this work and for existing scales that might serve as a starting point or inspiration for item formulation. Furthermore, empirical data was gathered during focus group interviews with different types of gamers, discussed in Chapter 2. These interviews served as a test for comparing scientific conceptualisations and lay descriptions of first-hand experiences. They were also used as a reference guide on choice of wording in the item formulation phase.

The Game Experience Questionnaire was developed with a modular structure, consisting of:

4. The core questionnaire (GEQ). This is the heart of the Game Experience Questionnaire, probing multiple components of players' experience while gaming.
5. The post-game questionnaire (PGQ), probing gamers' experience after the gaming session and any after effects.
6. The social presence module (SPGQ), probing gamers' experience of and involvement with their co-player(s).

These three lists are to be administered after the gaming session has ended. Additionally, a short in-game version of the GEQ was developed, the iGEQ, for probing in-game experience multiple times during a gaming session.

A large scale survey was performed to test the long list (>100 items) and explore the factor structures of the questionnaires. Factor analyses provided structures for the scales that made good sense in the light of theoretical considerations, with subscales that were all easy to interpret. Subsequent reliability tests resulted in the construction of reliable subscales with satisfactory to high internal consistencies. Statistical analyses of differences between gamers (gender, play frequency), all demonstrated the sensitivity of the scales, as well as each subscale's unique value (discriminant validity).

The full questionnaire is listed in Appendixes 1-4. Scoring guidelines for all scales are listed in Appendix 5.

5.2. Assessment of measurement criteria for the GEQ

In Chapter one, a list of criteria for (self-report) measures of game experience was discussed. According to this overview, the GEQ should preferably be tested on – and meet – criteria for reliability, validity, sensitivity, robustness, non-intrusiveness, and convenience.

For the core GEQ, iGEQ, PGQ, and SPGQ, we reported on the internal consistency of all subscales. Results showed that all measures are reliable, with satisfactory to excellent Cronbach's alpha values. The test-retest reliability of the entire battery of measures developed by FUGA (including the GEQ) will be established in WP6. Some limitations were noted for the PGQ, for which item variabilities were suboptimal. Although even these scales performed well on the remaining criteria below (except for the tiredness scale), more research is needed before this scale's final structure can be established.

Face validity and content validity were established in the construction phase of the GEQ, through the explicit use of the theoretical framework developed under FUGA, as well as recent conceptualisations in scientific literature, and the exploration of lay conceptualisations collected through focus group interviews. Predictive validity will be established through a series of studies performed under WP7. Convergent validity is in fact at the heart of the FUGA multi-measure approach, as we are comparing and correlating the outcomes of different measurement methods throughout the project lifetime. For the GEQ and its additional modules, discriminant validity was established internally for the various subscales, as different scales showed different response patterns to variations in player type, game content, and setting.

In Chapter 4 of the present deliverable, we analysed the sensitivity of the GEQ in relation to background variables such as, gender, play frequency, type of game, and social setting. The subscales of the GEQ and modules were able to significantly distinguish between different levels of these background variables. The single exception to this is the tiredness subscale of the PGQ.

The GEQ has explicitly been developed to probe experience 'sec', i.e., without reference to specific game or interface characteristics. This will aid the tool's robustness, i.e., the wide applicability of the questionnaire, as was demonstrated in the large scale survey, which assessed game experience of players of highly diverse backgrounds, playing a multitude of games, on various platforms in different social and physical contexts.

The GEQ, PGQ, and SPGQ, being post-test self-report measures, can be assumed to be non-intrusive. The intrusiveness of the in-game version of the GEQ, the iGEQ, is minimised through its limited number of items. Moreover, researchers can decide to discard more items from this list when they are targeting specific components of experience. Each component is assessed with only two items, which can be administered extremely quickly.

Lastly, as a low-cost paper-based measure, the GEQ can be regarded as highly convenient for researchers to use. It is easy to learn and easy to administer. The fact that the questionnaire will be translated into English, Finnish, Swedish and German, and be made available to all interested parties only adds to this effect. More importantly, convenience for research participants was also taken into account, by formulating brief items using simple, natural, and common language, and the consistent use of one type of answering scale.

5.3. Further steps in the development of the GEQ

The questionnaire has been provided to project partners for translation to Finnish, Swedish, and German. An English version of the GEQ is currently being tested by TU/e.

Translation instructions for partners were given as follows:

1. Two translators are needed to translate and back-translate the questionnaires to the new language.
2. Translator 1 translates the English items into the new language. The aim is to stay as close as possible to the meaning of the original English questionnaire. However, the translated item should feel natural and logical. The translator should be proficient and fluent in English and a native speaker in the new language.
3. A second translator translates the new items back into the English language. Translator 2 is proficient and fluent in the new language and an English native speaker.
4. The new English translation is compared to the original English version. The meaning, direction and intensity of the items should be identical. Where differences occur, the two translators discuss the discrepancy and decide on a translation that unambiguously probes what the original English item is meant to probe.
5. Please do not forget to also carefully translate the instructions and answering categories.
6. For your convenience, a second file (excel) is distributed, which contains the English items structured according to the components they belong to, as well as the Dutch versions of the items.

After the translated scale has been used for the first time, statistical analysis will be performed to confirm factor structure and internal consistencies of the scales. The extra items provided for translations of the GEQ scale allow for some flexibility in adopting or discarding single items from subscales in order to construct reliable scales.

As discussed above, the PGQ structure will be further explored in future empirical work. It should be noted that this scale is not crucial to experimentations within the FUGA project, as it is focused on developing continuous measurements for in-game experience. Also, in spite of its restrictions, three of its subscales did demonstrate reliability, validity, and sensitivity in their present form.

Empirical studies planned in the FUGA (see Deliverable 3.1) will employ the GEQ. This will provide additional data for further validating the questionnaire.

References

Bartle, R.A. (1996). *Hearts, clubs, diamonds and spades: players who suit MUDs*. http://www.mud.co.uk/richard/hcds.html.

Biocca, F, Harms, C., Burgoon, J., (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence*, 12, 456–480.

Biocca, F., Harms, C., Gregg, J. (2001). *The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity*. E. Lansing, MI: Media Interface and Network Design (M.I.N.D.) Lab.

Brown, E. & Cairns, P. (2004). A grounded investigation of game immersion. *ACM CHI 2004*, 1297-1300.

Costello, A. B. & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10* (7). Available online: http://pareonline.net/getvn.asp?v=10&n=7

Cronbach, L. J. (1990) Essentials of psychological testing. 5th edition. New York: Harper Collins.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self determination in human behavior.* New York: Plenum.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 227–268.

Ermi, L. & Mäyrä, F. (2005). Fundamental components of the gameplay experience: Analysing immersion. In: S. de Castell & J. Jenson (eds.), *Changing Views: Worlds in Play.* Selected papers of the 2005 Digital Games Research Association's Second International Conference.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Intrinsic Motivation Inventory, http://www.psych.rochester.edu/SDT/measures/word/IMIfull.doc (retrieved on 01-10-2003).

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC–Sense of Presence Inventory. *Presence, 10*, 282-297.

Lunt, P. & Livingstone, S. (1996). Rethinking the focus group in media and communications research. *Journal of Communication*, 46(2), 79-98.

Merton, R.K. (1987). The focused interview and focus groups: continuities and discontinuities. *Public Opinion Quarterly*, 51, 550-556.

Miller, L. et al. (1996). Female participants' Preferences in Software Design: Insights from a Focus Group.*Interpersonal Computing and Technology*, 4(2), 27-36.

Morgan, D. (1998). The *focus group guidebook. Focus group kit 1*. Thousand Oaks, CA: Sage.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion, 30*, 347–363.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A Review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology, 45*, 736–750.

Sweetser, P & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *ACM Computers in Entertainment, 3* (3), 1-24.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics*. Boston: Allyn and Bacon.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of Positive and Negative affect: the PANAS Scales. *Journal of Personality and Social Psychology, 47*, 1063–1070.

Yee, N. (2002). *Facets: 5 motivation factors for why people play MMORPG's*. http://www.nickyee.com/facets/home.html.

Appendix 1: Game Experience Questionnaire (GEQ)

Please indicate how you felt while playing the game for each of the items,
on the following scale:

not at all	slightly	moderately	fairly	extremely
0	1	2	3	4
< >	< >	< >	< >	< >

1 I felt content
2 I felt skilful
3 I was interested in the game's story
4 I could laugh about it
5 I felt completely absorbed
6 I felt happy
7 I felt tense
8 I felt that I was learning
9 I felt restless
10 I thought about other things
11 I found it tiresome
12 I felt strong
13 I thought it was hard
14 It was aesthetically pleasing
15 I forgot everything around me
16 I felt good
17 I was good at it
18 I felt bored
19 I felt successful
20 I felt imaginative
21 I felt that I could explore things
22 I enjoyed it
23 I was fast at reaching the game's targets
24 I felt annoyed
25 I was distracted
26 I felt stimulated
27 I felt irritable
28 I lost track of time
29 I felt challenged
30 I found it impressive
31 I was deeply concentrated in the game
32 I felt frustrated

33 It felt like a rich experience
34 I lost connection with the outside world
35 I was bored by the story
36 I had to put a lot of effort into it
37 I felt time pressure
38 It gave me a bad mood
39 I felt pressured
40 I was fully occupied with the game
41 I thought it was fun
42 I felt competent

Appendix 2: In-game Questionnaire (iGEQ)

Please indicate how you felt while playing the game for each of the items,
on the following scale:

not at all	slightly	moderately	fairly	extremely
0	1	2	3	4
< >	< >	< >	< >	< >

1	I was interested in the game's story	GEQ Core – 3
2	I felt successful	GEQ Core – 19
3	I felt bored	GEQ Core – 18
4	I found it impressive	GEQ Core – 30
5	I forgot everything around me	GEQ Core – 15
6	I felt frustrated	GEQ Core – 32
7	I found it tiresome	GEQ Core – 11
8	I felt irritable	GEQ Core – 27
9	I felt skilful	GEQ Core – 2
10	I felt completely absorbed	GEQ Core – 5
11	I felt content	GEQ Core – 1
12	I felt challenged	GEQ Core – 29
13	I felt stimulated	GEQ Core – 26
14	I felt good	GEQ Core – 16

Appendix 3: Post-game experience questionnaire (PGQ)

Please indicate how you felt after you finished playing the game for each of the items, on the following scale:

not at all	slightly	moderately	fairly	Extremely
0	1	2	3	4
< >	< >	< >	< >	< >

1 I felt revived
2 I felt bad
3 I found it hard to get back to reality
4 I felt guilty
5 It felt like a victory
6 I found it a waste of time
7 I felt energised
8 I felt satisfied
9 I felt disoriented
10 I felt exhausted
11 I felt that I could have done more useful things
12 I felt powerful
13 I felt weary
14 I felt regret
15 I felt ashamed
16 I felt proud
17 I had a sense that I had returned from a journey

Appendix 4: Social Presence Gaming Questionnaire (SPGQ)

Please indicate how you felt while playing the game for each of the items,
on the following scale:

not at all	slightly	moderately	fairly	extremely
0	1	2	3	4
< >	< >	< >	< >	< >

1 I empathized with the other(s)
2 My actions depended on the other(s) actions
3 The other's actions were dependent on my actions
4 I felt connected to the other(s)
5 The other(s) paid close attention to me
6 I paid close attention to the other(s)
7 I felt jealous about the other(s)
8 I found it enjoyable to be with the other(s)
9 When I was happy, the other(s) was(were) happy
10 When the other(s) was(were) happy, I was happy
11 I influenced the mood of the other(s)
12 I was influenced by the other(s) moods
13 I admired the other(s)
14 What the other(s) did affected what I did
15 What I did affected what the other(s) did
16 I felt revengeful
17 I felt schadenfreude (malicious delight)

Appendix 5: Scoring guidelines

Scoring guidelines GEQ

The Core GEQ Module consists of seven components, the items for each are listed below.

Most components have 5 items (except the SII which has 6). We have added one extra item for each of these, to create some flexibility should a translated item not work. Please also translate and use these items in your first studies, until the translated scales have been tested.

Component scores are computed as the average value of its items.

Competence: Items 2, 12, 17, 19, and 23; 42 is a spare item for translation..

Sensory and Imaginative Immersion: Items 3, 14, 20, 21, 30, and 33.

Flow: Items 5, 15, 28, 31, and 34; 40 is a spare item for translation.

Tension: Items 7, 9, 24, 27, and 32; 39 is a spare item for translation.

Challenge: Items 8, 13, 26, 29, and 36; 37 is a spare item for translation.

Negative affect: Items 10, 11, 18, 25, and 35; 38 is a spare item for translation.

Positive affect: Items 1, 4, 6, 16, and 22; 41 is a spare item for translation.

Scoring guidelines iGEQ

The In-game questionnaire consists of seven components, identical to the core Module. However, only two items are used for every component. The items for each are listed below.

Component scores are computed as the average value of its items.

Competence: Items 2 and 9.

Sensory and Imaginative Immersion: Items 1 and 4.

Flow: Items 5 and 10.

Tension: Items 6 and 8.

Challenge: Items 12 and 13.

Negative affect: Items 3 and 7.

Positive affect: Items 11 and 14.

Scoring guidelines PGQ

The post-game Module consists of four components, the items for each are listed below.

Component scores are computed as the average value of its items.

Positive Experience: Items 1, 5, 7, 8, 12, 16.

Negative experience: Items 2, 4, 6, 11, 14, 15.

Tiredness: Items 10, 13.

Returning to Reality: Items 3, 9, 17.

Scoring guidelines SPGQ

The Social Presence Module consists of three components, the items for each are listed below.

Component scores are computed as the average value of its items.

Psychological Involvement – Empathy: Items 1, 4, 8, 9, 10, and 13.

Psychological Involvement – Negative Feelings: Items 7, 11, 12, 16, and 17.

Behavioural Involvement: Items 2, 3, 5, 6, 14, and 15.