Reality Substitution in Video Gamers:
Exploring the Reality Substitution Model and its Ability to Predict Health and Behavioural Outcomes.

(2015)

Nico V. Rosenkrantz, BSc. (Student number: 3680460)
University Utrecht, Faculty of Social and Behavioural Sciences
Social Psychology
n.v.rosenkrantz@students.uu.nl

Master Thesis
Mentor: Ron Dotsch, PhD.

Table of Contents.

Opening Statement.

This thesis is part of the master program of social psychology at the University Utrecht. My interest in video games expressed itself at an early age, as more and more households came to own a dedicated video game console. The fact that I could now dedicate my time and effort into understanding the science behind video games in the context of psychology was an opportunity too big to pass up. I have even been able to incorporate elements of not just social psychology, but labour and organisational psychology as well. It has resulted in one of the most informative and gratifying experiences I have enjoyed so far. I am grateful to my mentor Dr. Dotsch, whose advice, input and support has been invaluable for this entire thesis. My family have been incredibly supportive as well, for which I am incredibly grateful. My wish is that this thesis will contribute to the understanding of video games and the formation of a unified theory of them (a long shot but one should dream), and also to the betterment of the experiences people get from playing future generations of video games.

Nico Rosenkrantz (Student number: 3680460)
University Utrecht (2015)

Abstract.

To predict the nature (in terms of functional or dysfunctional) of the outcomes of playing video games, Rosenkrantz (2015), conceived the reality substitution model (RSM). Here, we report the two following studies.

Study 1 concerned constructing of a valid measure for reality substitution (the RSQ-18). The results indicated that all three scales of the intensity dimensions (immersion, empathy and satisfaction) had good internal reliabilities (>.7; though immersion was at .616), factor loadings (except for immersion: 47.472%), test-retest reliability (>.7) and convergent validity (GEQ, Gamer Experience Questionnaire) as well as divergent validity (SCL-90-R, Symptom Checklist 90 Revised). The results almost fully supported the internal validity of reality substitution according to the RSM, though various additional interaction effects were found.

Study 2 concerned the testing of the predictive validity of this measure (and by extension part of the Reality Substitution Model). The results indicated that playing intensely and frequently and for long periods of time (as predicted by the RSM and measured by the RSQ-18), does not increase dysfunctional outcomes. The primary (and often sole) predictor of outcomes (PTM, Prosocial Tendencies Measure; UBOS, Utrecht Burn-out Scale; BPAQ Short Form, Buss and Perry Aggression Questionnaire Short Form) was the intensity dimension. The UWES (Utrecht Work Engagement Scale) was unrelated to the RSQ-18 and its scales. The RSQ-18 is not yet able to predict the nature of video game outcomes, though its intensity dimension lends itself well for future study.

In closing, we discuss several suggestions for development of the RSM and the RSQ-18.

Introduction.

Video games play an increasing role in our lives. According to the ESA (2015), the Entertainment Software Association, an estimated 51% of U.S. households owns at least one dedicated game console. Video games are an industry of their own, worth over 22.4 billion U.S. dollars in the U.S.A. alone in 2014. This far surpasses the movie and music industry (ESA, 2015). About 42% of Americans plays regularly (3 hours or more per week). Playing video games is not just for the young. The average age of gamers is 35 years old (26% < 18 years old; 30% = 18 - 35 years old; 27% > 50 years old), 44% is female, and they have been playing on average for 13 years. Video games provide us with rich and engaging worlds (Snead, 2014). Moreover, as technology advances and computational power grows, video games stand to become even better in the future. Already the gamer is no longer a passive observer but is actively shaping the story. This makes video games an engaging medium that can elicit an array of cognitions, feelings and behaviours (Snead, 2014; Rosenkrantz, 2015; i.e. BioWare Corp., 2012; CD Projekt RED, 2015). However, for something that has such a presence in our lives, video games and their influence on us are poorly understood (Rosenkrantz, 2015).

Literature on the effects of video gaming on the gamer is incoherent and sometimes contradictory. For instance, violent video games increase aggression, reduce empathy and prosocial behaviour and desensitize people to violence (Anderson et al., 2010; Kirsh & Mounts, 2007; Anderson & Bushman, 2001). Furthermore, video games provide substantial screen time that is associated with inactivity and the development of obesity (Lanningham-Foster et al., 2006). However, a meta-analytic review by Ferguson and Kilburn (2009) points out that publication bias is a problem for studies of aggressive behaviour, as are methodological problems (such as the use of poor aggression measures) and inflated effect sizes. Once corrected, these studies provide little support for the hypothesis that violence in media such as video games is associated with increased aggression. Primack et al. (2012), reviewed articles about video games improving health-related outcomes. They concluded that, by making use of video games, psychological therapy outcomes improved by as much as 69%, physical therapy outcomes by 59%, physical outcomes by 50%, clinician skill outcomes by 46%, health education outcomes and pain distraction outcomes by 42% and disease self-management outcomes by 37%. Video games also have positive associations to enhanced visual acuity and attention (Green & Bavelier, 2003; Bavelier, 2012), increased creativity (Jackson et al., 2012) and improved quick decision-making without a loss of accuracy (Dye, Green & Bavelier, 2009). This shows that video games can lead to functional, as well as dysfunctional outcomes. Yet the mechanism behind this still awaits an explanation, backed up by evidence. One possible explanation comes from the Reality Substitution Model (RSM; Rosenkrantz, 2015).

The reality substitution model.

Short-term and long-term reality substitution. According to the RSM, the difference in the outcome of a video game - i.e. functional or dysfunctional outcome - originates from reality substitution (Figure 1; Rosenkrantz, 2015). Specifically the type of reality substitution the gamer experiences. Reality substitution is the use of *virtual* (video game) *features* to satisfy *real needs* of the gamer. By immersing oneself into the *video game's (virtual) reality* as a substitute for *actual reality*, a gamer experiences reality substitution. As long as the gamer is motivated to play, reality substitution will always occur when playing video games, albeit to varying degrees. The stronger the reality substitution, the more intensely a gamer is playing. Gamers experience *long-term reality substitution*, when they play not only intensely but also frequently and for extended periods. This leads to

dysfunctional outcomes. Gamers who play conservatively, experience *short-term reality substitution*. This leads to functional outcomes. According to the RSM, we can determine whether a gamer is at risk of dysfunctional outcomes by looking at how intensely, how frequently and how long the gamer plays.

The three stages of reality substitution. Reality substitution, and by extension its intensity, is determined by its three stages (Figure 1). The first stage is immersion. The gamers focus their attention (mental resources) on what happens in the virtual world of the video game. Specifically on the avatar or character they play, which is their representation in the video game. With the bulk of the mental resources dedicated to the virtual world, there are less mental resources available for reality. The result is similar to when you enjoy a good movie and no longer notice other people eating popcorn. Immersion as related to reality substitution is different from immersion into virtual reality. With reality substitution, it is the character in the game that experiences the virtual world and then the gamer 'shares' in these experiences (via empathy), supported by mental resources such as one's full attention. By displacing the perception of one's physical self, immersion into virtual reality is achieved (Freina & Ott, 2015). For instance via surround-sound and 3D-screens whereby the gamer experiences the virtual world (via suspension of disbelief) from a first-person perspective. Once mental resources become available, the second stage of reality substitution can occur.

In the second stage, the gamer empathises with one's character (with what they think, feel and do) in order to satisfy their own needs (Rosenkrantz, 2015). Ahn, Le and Bailenson (2013), found that embodied experiences of someone who is colour-blind by using immersive virtual environment technology, led to increased self-other merging in participants. This in turn led to increased helping behaviour towards people who are colour-blind. The empathy-stage of reality substitution is comparable to self-other merging. With the difference that, instead of shared identities, it is about shared experiences such as emotions and cognitions. Additionally, self-other merging mediates the relationship between perspective taking and changes in self-concept (Goldstein & Cialdini, 2007). Taking someone's perspective resulted in the participant to perceive his/her self to be similar to the person whose perspective he or she took. This principle is also applicable to the empathy-stage. Insecure individuals can increase their self-esteem, by playing with characters that do have high self-esteem. The video game becomes a medium through which gamers can take an experience from the virtual world to satisfy their needs in the real world. Thus, empathy leads to satisfaction.

Thirdly, the satisfaction stage. The goal of video gaming is to achieve a state of fulfilment. Gamers attain this state through empathy with their character. As they become empathic towards their characters, gamers anthropomorphize the characters (or non-biological agents) they play. This leads to more interpersonal closeness between gamer and agent (Müller, Van Baaren, Van Someren & Dijksterhuis, 2014), which in turn may lead to the gamer becoming more emotionally invested in the video game. According to the RSM, an outcome as need satisfaction inherently conditions the gamer to play video games more often and for longer periods (Rosenkrantz, 2015). This is reminiscent of Thorndike's law of effect (1898). The goal of need satisfaction naturally arises from the RSM, as it is rooted in the assumption that unsatisfied needs drive all human beings (Redmond, 2009). In other words, people play a video game because it satisfies a need for them, which is not adequately, met otherwise (Rosenkrantz, 2015).

Matching needs with video game features. Rosenkrantz (2015) proposes that reality substitution stems from the features of the video game fulfilling the needs of gamer (Figure 1). For example, multiplayer features can fulfil a need for affiliation, whereas features that allow you to edit

your character may fulfil a need for self-esteem or changeable difficulty settings may fulfil a need for achievement. Other means than video games can satisfy needs. If you are in need of relaxation, a day at the beach may be equally satisfying. The advantage of video games however is that they are accessible from the comfort of our own home, where we also have other vital resources (food and drink, shelter, warmth etc.) to our disposal. Rosenkrantz (2015) propose that in the case of a match between a need and a feature of the video game, reality substitution will occur. If there is no need to fulfil, or the video game does not offer features matching with the need, the gamer will discontinue playing as soon as suitable. Rosenkrantz (2015) gives a more comprehensive explanation of the process preceding reality substitution and the RSM as a whole. For this study, this information will suffice, as its focus lies on predicting the outcomes of reality substitution.

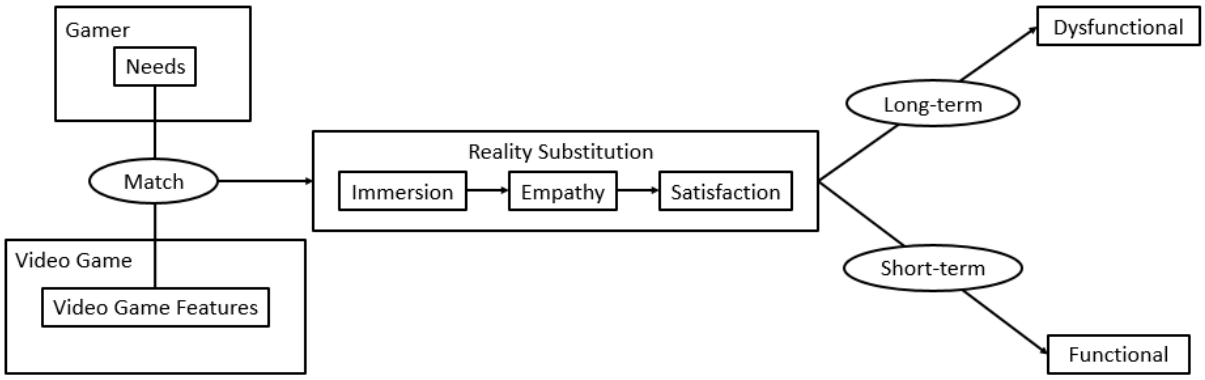

Figure 1. The Reality Substitution-Model: the match between the needs of the gamer and the features of a video game leads to reality substitution. Short-term reality substitution has functional outcomes, whereas long-term reality substitution leads to dysfunctional outcomes (Rosenkrantz, 2015).

The present study. In this paper, our focus is on finding an explanation for the contradictions in the literature concerning video games mentioned earlier. Hence, we must find a theoretical mechanism that relates and connect functional and dysfunctional outcomes. The RSM proposes that reality substitution is the underlying cause for the effects of video games on health and behaviour. However, there is currently no measure for reality substitution. Thus, in Study 1 we attempted to construct a valid measure for reality substitution using a sample of gamers. Study 2 concerned the testing of the predictive validity of this measure in regards to functional, dysfunctional, health and behavioural outcomes. Both studies have a non-experimental correlational design.

Study 1: Construction and Validation of the Reality Substitution Questionnaire.

Introduction Study 1. Study 1 was about designing a valid measure for reality substitution. This would require, according to the RSM, a scale to measure immersion, empathy and satisfaction. In addition, items to measure the frequency and duration of play were also necessary. We subsequently tested this questionnaire for its internal consistency and test-retest reliability and its internal, convergent and divergent validity.

Hypotheses. In Study 1, we tested hypothesis 1, 2 and 3. Hypothesis 1 concerned the internal validity of reality substitution. Hypothesis 1^A was that immersion would explain a significant amount of variance of empathy. Hypothesis 1^B was that empathy would explain a significant amount of variance of satisfaction. Hypothesis 1^C was that satisfaction would explain a significant amount of variance on the frequency and duration of play. Hypothesis 2 was that the Reality Substitution

Questionnaire (RSQ) has a moderate correlation with the Gamer Experience Questionnaire (GEQ). This questionnaire measures engagement into the video game with dimensions such as flow, presence and immersion. Because these dimensions relate to reality substitution (convergent validity), we would expect there to be a moderate correlation. Hypothesis 3 was that the RSQ correlates weakly with the two subscales of the SCL-90-R (Symptom Checklist 90 Revised) paranoid ideation and psychotism. As neither causes reality substitution (Rosenkrantz, 2015), we would expect the correlation with reality substitution to be weak (divergent validity).

Methods Study 1.

Item-generation. We developed three measures: the immersion scale, the empathy scale and the satisfaction scale (Table 1). As immersion measure, we adapted the Absorption Scale of the validated Job Engagement Scale developed by Schaufeli and Bakker (2004) and the immersion component used by Yee (2007). We also included adapted items of the Presence Scale of the PENS Scale (Ryan, Rigby & Przybylski, 2006). To clarify, the PENS Model (Player Experience of Need Satisfaction; Rigby & Ryan, 2007) uses the self-determination theory (Deci & Ryan, 2000) to predict positive experiential and commercial outcomes from video games. To measure empathy, we used the validated Toronto Empathy Questionnaire (Spreng, McKinnon, Mar and Levine, 2009). Additionally, we added items regarding the satisfaction of playing. For these, we looked at several items of the Intrinsic Motivation Inventory (IMI; Ryan, Mims & Koestner, 1983), also used by Ryan, Rigby and Przybylski (2006) in their development of the PENS Scale. As for the response scale, we opted for a 7-point Likert Scale as several source-questionnaire used this scale too. We measured frequency and duration of gameplay by asking the gamers for the average number of days they play per week, the number of days they played in the last 28 days and the average duration per session of play in hours (Appendix 3). We added the items for these two dimensions to the demographics-section (Appendix 4). We adjusted some of the items of these questionnaires to make sure they suited the situation (playing video games). For instance, an item such as “I enjoy my job”, would become “I enjoy my video games”. We also added a number of items to make sure the scales completely covered the definition of reality substitution. In addition, we would now have options should some items require deletion, for instance in the event that some items score poorly (i.e. low or negative item-intercorrelations). A copy of the entire series of questionnaires in English as well as Dutch is included in Appendices 2 to 4.

Table 1. *Reality Substitution, its Dimensions, its Scales and the Respective Source Material to be used for the Item-generation in Study 1.*

Construct	Dimension	Scale	Original Questionnaire and Authors
Reality Substitution (RSQ-18)	Intensity	Immersion	Job Engagement Scale: Absorption Scale *Schaufeli and Bakker (2004)*
			Motivations of Play Inventory: Immersion Component *Yee (2007)*
			Player Experience of Needs Satisfaction Scale: Physical/Emotional/Narrative Presence Scale *Ryan, Rigby and Przybylski (2006)*
		Empathy	Toronto Empathy Questionnaire *Spreng, McKinnon, Mar and Levine (2009)*
		Satisfaction	Intrinsic Motivation Inventory *Ryan, Mims and Koestner (1983)*
	Duration		-
	Frequency		-

Validation and materials. The first step to validate the RSQ was to assess the reliability. Item-remainder correlations showed whether poor items were present in the questionnaire (we desired at least .20, preferably around .40). To test the internal consistency we computed the Cronbach's alpha (at least .7, preferably between .80 and .90). These criteria were based on the research from Van Den Brink and Mellenbergh (1998). We deleted items that did not meet the criteria, or whose removal would increase reliability. We assessed test-retest reliability by asking participants to fill out the questionnaire again one week after the initial data collection (from June 1st to June 5th).

We tested the number of dimensions (the underlying internal structure) of reality substitution with a principal components analysis. We wanted to check whether the structure of reality substitution, measures by our questionnaire, matches the structure as proposed by the RSM. The RSM proposes that reality substitution is multidimensional, with the items loading high on their respective dimension (factor). It could also be that the items would load on a single dimension, thus showing reality substitution to be unidimensional. We also executed the same analysis on the individual scales, to make sure they measured only one factor. The exceptions being the items for the duration and frequency of play, containing merely one and two items respectively.

To test for convergent validity, participants also filled out the GEQ (Brockmyer et al., 2009). As mentioned earlier, this questionnaire measures engagement into the video game with dimensions such as flow, presence and immersion (19 items, 3-point Likert Scale).Brockmyer and colleagues (2009) have validated it. Our questionnaire should correlate (Spearman correlation) moderately with the GEQ, indicating convergent validity.

To test for divergent validity, participants also filled out two scales of the Symptomatic Checklist-90 Revised (SCL-90-R; Derogatis, 1992); the paranoia ideation-subscale and the psychotism-subscale (16 items, 5-point Likert Scale; validated by Prinz and colleagues in 2013). As reality substitution is supposedly unrelated to both constructs, we expect the Spearman correlation to reality substitution to be low, indicating divergent validity.

To test for predictive validity, we also presented four additional questionnaires to the participants. We did this to collect data for Study 2. We will elaborate on this later, when reporting on Study 2.

Sample. Recruitment was done primarily online via social media, such as dedicated video gaming forums, via convenience sampling. This expanded the reach of these studies far beyond the confines of the social network of the researchers. This in turn allowed for better generalisability when it came to age, gender and nationality.

We were only interested in dedicated gamers, not people who merely play the occasional Angry Birds on the train home. Therefore, sampled participants were required to play video games regularly (at least once a month) and have done so in the last four weeks, including the day they took part in this study. This same sample was used in Study 2 as well.

Procedure. The procedure of completing the aforementioned questionnaires was largely self-directed. Between May 8th and May 25th, participants were able to fill in the questionnaire. Potential participants accessed the questionnaires via the link that we posted on various websites for dedicated to gamers, such as EA's BioWare forum. Participants learned beforehand that the study would be about their experiences with video games. There would be no right or wrong answers, neither any trick questions. They could fill out the questionnaire in Dutch as well as English, in approximately 20 minutes time. We informed them about their rights with respect to voluntary

participation, anonymity and confidentiality. Lastly, we gave participants our contact information (email) and thanked them for their participation.

At the start of the questionnaire, we asked participants whether they were gamers as per our definition. If a participant would not meet the requirements, we would terminate their participation to save their time and effort. Inclusion of their data would otherwise pollute our sample. We gathered various demographic data, such as age, gender, average duration per session of play, frequency of play, last time played--etc. After the demographics (13 open-ended questions or multiple-choice items), the RSQ would follow (42 Items, 7-point Likert Scale, randomised and with various items mirrored). After that would follow the GEQ, the SCL-90-R, the UWES, the UBOS, the PTM and lastly the BPAQ Short Form. Again, we will elaborate more on the last four questionnaires in Study 2. The total number of items a participant would fill out was 155, indicating with every item to which degree the item applied to them. At the end of this series of questionnaires, we would thank the participant again and informed them about the test-retest study. If they wanted to participate, they could leave their email address at this page. Participants received an email with therein a link to the test-retest questionnaire (Appendix 4), where the procedure was largely the same as before.

Results Study 1. In the span of a little over two weeks, 334 participants started filling out the online questionnaire. We analysed the data of only the 164 participant who met the criteria for being a "gamer" and who finished all questionnaires (40 female; average age 27.3; Table 2). We divided scores (1 to 28) on the item of the frequency scale regarding the number of days played in the last 28 days by 4, to decrease the range of scores, making them comparable in range to the 7-point Likert scale of almost all other measures.

Table 2. *Distribution of Sample as per Gender and Age.*

	N (Study 1)	**Age (Study 1) *M/SD***	***N* Retest**	**Age Retest *(M/SD)***
Male	124	27.7 - 8.6	29	28.1 - 7.7
Female	40	26.1 - 9.0	8	28.5 - 8.6
Total	164	27.3 - 8.7	37	28.2 - 7.8

Reliability of the scales. We started with the reliability of the scales individually, reported in Table 3. Both the scales for empathy and satisfaction show good reliability (Cronbach's alpha > .7). However, immersion is barely adequate, primarily due to low item-intercorrelations and thus diffuse and/or ambiguous items. The test-retest indicated good results for all scales as well as their combined form (Table 3). The resulting RSQ-18 (15 items for intensity, two for frequency and an item for duration) has a Cronbach's alpha of .740.

Table 3. *The Cronbach's Alpha's for the New Scales for Immersion, Empathy and Satisfaction and the dimension of intensity as well as the Test-Retest Reliability Score for Each of Them.*

Scale/Dimension	**Number of Items**	***Cronbach's Alpha***	**Remaining Items**	**Pearson Test-Retest**
Immersion	4	.616	13, 15, 30, 33	.790 (p <.001)
Empathy	8	.875	3, 9, 14, 16, 18, 24, 27, 32	.892 (p <.001)
Satisfaction	3	.746	12, 19, 21	.784 (p <.001)
Intensity	15	.850	All the above	.885 (p <.001)

Number of dimensions. The next step was to inspect the number of dimensions of reality substitution via principal components analysis. We wanted to check whether the structure of reality

substitution, measures by our questionnaire, matches the structure as proposed by the RSM (multidimensional). Examination of the data indicated that not all of the 18 items were perfectly normally distributed, but given the robust nature of factor analysis, we did not consider these deviations to be problematic (Field, 2009). The KMO measure of sampling adequacy (.806) and Bartlett's test of sphericity (.000) indicated the data was fit for factor analysis. The communalities of the intensity-items were somewhat low however, ranging between .75 and .44 (duration and frequency items did have communalities of between .8 and .9). The RSQ-18 (the remaining 18 items for intensity, frequency and duration together) showed multidimensionality (5 factors; Table 4) and explained more than 50% of the variance. Furthermore, the RSQ-18 has less than 50% nonredundant residuals arisen from the reproduced correlations (item-intercorrelations based on the number of extracted factors) as we would desire (Field, 2009). Each item loads only strongly on their respective factor (i.e. scale). Note that we supressed factor loadings below .43 in Table 4, as they are not significant. This number is an approximation based on the recommendations from Stevens (2002) for a sample size of 100 (.512) and 200 (.364). When applying the same procedure to the intensity scales separately, they all showed only one component, as we would expect. However, they (in particular immersion) did still show low communalities (below .6) and large percentages of nonredundant residuals with absolute values greater than 0.05 that arise from the reproduced correlations. Nevertheless, they did pass the KMO measure of sampling adequacy and Bartlett's test of sphericity and managed to explain at least 50% of the variance with only one component (except for immersion: 47.472% where we would have liked to see at least 50%; Field, 2009). The resulting RSQ-18 is included in Appendix 2 and 3.

Table 4. *The Factor Loadings and the Explained Variance for the RSQ-18, after Varimax Rotation.*

Dimension	Scale	Item	Component				
			1	2	3	4	5
Intensity	**Empathy**	3	.765				
		9	.733				
		14	.652				
		16	.690				
		18	.710				
		24	.700				
		27	.712				
		32	.828				
	Immersion	13				.586	
		15				.494	
		30				.731	
		33				.791	
	Satisfaction	12		.684			
		19		.822			
		21		.825			
Duration		Duration					.890
Frequency		Days per Week			.921		
		Days last 4 Weeks			.926		
% Variance Explained			24.771	12.839	11.378	11.138	6.133
% Total Variance Explained					66.259		

Note. Rotation converged in six iterations. There are 58 (37%) nonredundant residuals with absolute values greater than 0.05 that arise from the reproduced correlations.

Mean and standard deviation of scores. After having checked the reliability and number of dimensions, we calculated the mean and standard deviation of the scores on various scales of the RSQ-18 and the other measures used in Study 1 (GEQ and SCL-90-R). They can be found in Table 5.

Table 5. *The Mean and Standard Deviation of the Scores for Study 1 on the Scales and Dimensions of the RSQ-18, the RSQ-18 itself, the GEQ, and the SCL-90-R.*

Questionnaire	Statistic	
	M	***SD***
Immersion	4.12	1.19
Empathy	4.09	1.30
Satisfaction	6.01	.94
Frequency	5.15	1.99
Duration	3.68	1.47
Intensity	4.48	.95
RSQ-18	4.51	.88
GEQ	1.77	.33
SCL-90-R	1.66	.60

Internal validity. Hypothesis $1^{(A\text{-}B\text{-}C)}$ concerned the internal validity of reality substitution and required several multiple regressions (forced entry method). Prior to the multiple regression analyses, we checked the assumptions. We checked for univariate outliers amongst the variables by looking at the boxplots. An outlier was defined as a score 1.5 - 3.0 box lengths above or below the box boundaries (Allen & Bennett, 2010), whereas extreme scores are scores above or below more than 3.0 box lengths beyond the box boundaries. Though outliers were detected, all plots of the standard residuals indicated normality, so they were not of concern (Field, 2009). Variables that showed extreme scores (average duration and the variable for the interaction effect between intensity and duration) benefitted of natural log transformations. Only one extreme score remained on the duration scale. The participant, solely, scored in the lowest category of this (skewed) scale. Since this was a valid score, we did not remove this score from the sample. No transformation of any kind however (log, inverse, square root), was successful at redistributing the average score on the frequency scale, the average score on the duration scale and the average score on the satisfaction scale. These three (and several interaction effect-variables), violate the normality-assumption of the predictors. However, the normal probability plot of standardised residuals as well as the scatterplot of standardized residuals against standardized predicted values after every regression met the assumptions of normality, linearity and homoscedasticity of residuals.

Good Durbin-Watson values indicated uncorrelated residuals. The Mahalanobis distance did not at any point exceed the critical χ^2 for $df = k$ limit (number of predictor variables in the model) at $\alpha = .001$ (Allen & Bennett, 2010), so multivariate outliers were not a concern. Lastly, we centered the predictors and used these to compute the interaction effect variables. This way, we avoided multicollinearity--a threat to the validity of our results--as the interaction effects would otherwise have strongly correlated to the individual predictors.

We started by looking at the simple effect between the scales (Table 6). We calculated these by determining the Spearman correlations (since data is not truly of interval scale and skewed), between the various scales and dimensions of the RSQ-18. These are equal to the β's that would have arisen from linear regressions, showing in effect the simple effects between these measures. We confirmed Hypothesis 1^{A}, that immersion would explain a significant amount of variance of the

empathy scale (Table 6). Hypothesis 1^B was that empathy would explain a significant amount of variance of satisfaction. This hypothesis was also confirmed. The relation between immersion and satisfaction appears to be fully mediated (no interaction effect either) by empathy (Table 7). Hypothesis 1^C was that satisfaction would explain a significant amount of variance on frequency and duration. Satisfaction does explain a significant amount of variance of the frequency scale (as does empathy), but does not do so with the duration scale. There are two interaction effects present for duration however, one between immersion and empathy (negative) and another between immersion and satisfaction (positive). Only empathy shows a significant correlation with the duration scale (Table 6). The intensity dimension significantly correlates to all scales and dimensions.

Table 6. *Results of Spearman Correlations between the Dimensions and Scales of the RSQ-18.*

Scale/Dimension	Immersion	Empathy	Satisfaction	Frequency	Duration	Intensity
Immersion	X					
Empathy	.390***	X				
Satisfaction	.152*	.323***	X			
Frequency	-.005	.191*	.354***	X		
Duration	.065	.170*	.091	.152	X	
Intensity	.627***	.934***	.461***	.208**	.170*	X

Note. * = significant at an α of .05; ** = significant at an α of .01; *** = significant at an α <.001.

Table 7. *Results of the Regression Analyses used to inspect the Internal Validity of the Intensity Dimension.*

Intensity Scale(s)		**Multiple Regression Statistics**						
Predictor	**Dependent**	***R²***	***Adjusted R²***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***B***	***p***
Immersion	**Satisfaction**	.114	.098	6.871 (3, 160)	.000	.041	.065	.052	.529
Empathy						.226	.059	.312	.000
Immersion × Empathy						-.025	.041	-.046	.538
RSQ-18 Scale(s)		**Multiple Regression Statistics**						
Predictor	**Dependent**	***R²***	***Adjusted R²***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***B***	***p***
Immersion	**Frequency**	.187	.150	5.110 (7, 156)	.000	-.175	.146	-.104	.233
Empathy						.275	.134	.179	.041
Satisfaction						.904	.185	.427	.000
Im × Em						-.001	.102	-.001	.990
Im × Sa						.348	.185	.189	.062
Em × Sa						-.291	.164	-.187	.078
Im × Em × Sa						-.213	.116	-.204	.069
RSQ-18 Scale(s)		**Multiple Regression Statistics**						
Predictor	**Dependent**	***R²***	***Adjusted R²***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***B***	***p***
Immersion	**Duration**	.082	.041	1.999 (7, 156)	.059	-.011	.029	-.034	.710
Empathy						.037	.026	.130	.162
Satisfaction						.026	.036	.066	.480
Im × Em						-.043	.020	-.198	.036
Im × Sa						.073	.036	.215	.045
Em × Sa						.027	.032	.096	.396
Im × Em × Sa						.020	.023	.102	.390

Note. Significant correlations (α = .05) have been accentuated. Im = Immersion, Em = Empathy and Sa = Satisfaction. "×" indicates an interaction-effect.

External validity. To test the convergent validity and divergent validity of the RSQ-18, its relation to two other questionnaires was tested (Spearman's Rho, one-tailed): the GEQ and a combination of two subscales of the SCL-90-R (paranoia and psychotism). Hypothesis 2 (the convergent validity) was confirmed, as the RSQ-18 indeed has a moderate correlation with the GEQ (r_S = .547, p = 0.01). We suspect flow, presence and immersion (GEQ) may be related to the intensity dimensions of the RSQ-18 (since the overlying theories are related), but less so to frequency and duration. Thus we will consider our finding as adequate, as there are no further objective guidelines for convergent validity correlations. Hypothesis 3, the divergent validity, was also confirmed, as the RSQ-18 indeed has a weak correlation with the two subscales of the SCL-90-R (r_S = .351, p = 0.01). This is considerably lower than the threshold of .85 suggested by Clark and Watson (1995), which would indeed suggest that two different constructs were measured. Further insight into the external validity can be gained from Table 8.

Table 8. *Results of Spearman Correlations between the Dimensions and Scales of the RSQ-18 and the GEQ and SCL-90-R.*

Questionnaire	*GEQ*	*SCL-90-R*
Immersion	.614***	.249***
Empathy	.459***	.362***
Satisfaction	.202**	.153
Frequency	.107	.013
Duration	.034	.004
Intensity	.584***	.390***
RSQ-18	.547***	.351***

Note. * = significant at an α of .05; ** = significant at an α of .01; *** = significant at an α <.001.

Discussion Study 1.

RSQ-18. Our study into the construction and validation yielded a questionnaire to measure reality substitution. However, the immersion scale would benefit from improvements to its internal consistency. Nonetheless, the test-retest reliability was good for all newly developed scales (immersion, empathy and satisfaction). We confirmed hypotheses $1^{(A, B, C)}$ regarding the internal validity in its entirety. The RSQ-18 correlated moderately with the GEQ, indicating convergent validity, which confirmed hypothesis 2. Finally, the RSQ-18 correlated weakly with the SCL-90-R, indicating divergent validity, which confirmed hypothesis 3.

Intensity scales. The analyses regarding the internal validity of reality substitution indicated that immersion positively relates to empathy. This indicates that people who are more immersed (have mental resources dedicated to the video game) empathize more strongly with their character in the game. To empathize, you need mental resources such as attention as empathy is not only an affective but also a cognitive process (Pijnenborg, Spikman, Jeronimus & Aleman, 2013). Consequently, the gamer cannot dedicate these resources to other things (such as monitoring and reacting to reality) as mental resources are limited (Baumeister, Bratslavsky, Muraven & Tice, 1998). The immersion scale measures the lack of attention to reality. If there is a decreased amount of attention to reality, the focus must be on the video game. The more attention available for gaming, the better the gamer can empathize with his character. The better a gamer empathizes, the more satisfying playing the video game is.

Relations between dimensions. The more satisfaction a gamer experiences from playing the video game, the greater the frequency with which a gamer plays. This is reminiscent of the law of effect (Thorndike, 1898): the satisfying experience of playing video games (the reward), increases the likelihood that the gamer will exhibit the preceding behaviour (playing video games) again. It would appear that experiencing empathy also adds to this frequency. Perhaps because empathy as we measured it has some inherent pleasurable qualities to it, making it enticing to do more.

Concerning the duration of play: though there were no significant main effects found, there were two interaction effects present (Table 7). A negative interaction effect between immersion and empathy, yet a positive interaction effect between immersion and satisfaction. This appears contradictory. We suspect that it immersion may be putting a strain on the gamer, reminiscent of the depletion theory (Baumeister, Bratslavsky, Muraven & Tice, 1998), which states that mental resources come from a limited pool that can be used up. Empathy may be straining on the gamer as well. Williams (1989) reports a positive correlation between emotional empathy and emotional exhaustion. Satisfaction however, less related to immersion than empathy (Table 6), is strong enough as a predictor to overcome the negative effect of immersion on duration. The result being that its subsequent interaction-effect is positive.

Limitations and future directions. When interpreting these results we should make a few remarks. Note that the immersion scale does not have an optimal reliability. It would benefit from additional study to increase its item-intercorrelations and Cronbach's alpha. In addition, in particular the scores on the duration scale are skewed. As gamers had longer average sessions of play than anticipated, they all scored in the highest categories. It is prudent that scientists use large response scales when measuring the duration of play. This also applies to the frequency scale. Many gamers played more than half of the days of the week, skewing the scores. However, given the use of convenience sampling, primarily on hardcore gamer-dedicated websites, this comes as no surprise. Lastly, given the fact that the results are correlational in nature, we cannot draw any conclusions as to causation between these variables. Longitudinal studies, in which the exposure to various kinds of video games is measured, would allow us to infer causal relationships. The general discussion will elaborate on several other remarks concerning this paper as whole.

Study 2: Testing the Predictive Validity of the RSQ-18: Short-term vs. Long-term Reality Substitution

Introduction Study 2. Having inspected the reliability, internal validity and external validity of the RSQ-18 in Study 1, Study 2 was to inspect its predictive validity. We wanted insight in how reality substitution relates to various constructs. Specifically, engagement, burnout, prosocial tendencies and aggression. Furthermore, we have established in Study 1 that reality substitution is multidimensional. Subsequently the way the dimensions interact had to be tested. We wanted to know whether the dimensions would have an interactive or additive effect on the outcomes.

During the procedure for Study 1, as described earlier, we also presented the necessary questionnaires for Study 2 to the participants (four of them in total). These were to be used to inspect the predictive validity of the RSQ-18 for functional as well as dysfunctional health and behavioural outcomes.

Hypotheses. We confirmed reality substitution to be multidimensional in Study 1. The RSM proposes that dysfunctional outcomes would only rise to significance when someone scores high on

all three dimensions. In other words, when a gamer plays intensely, frequently and for prolonged periods. Functional outcomes should arise when a gamer plays more conservatively (e.g. only main effects). Thus, Hypothesis 4 was that each predictor (intensity, frequency and duration), would be able to explain a significant amount of variance in functional outcomes. However, no (three-way) interaction effect was expected (hypothesis 4^{A}: work/study engagement with the UWES; hypothesis 4^{B}: prosocial behaviour with the PTM). Hypothesis 5 was that only an interaction between all three dimensions would explain a significant amount of variance in dysfunctional outcomes (hypothesis 5^{A}: (study) burnout with the UBOS; hypothesis 5^{B}: aggression with the BPAQ-SF). In other words, except for the three-way interaction-effect, we expect no other main effects to be significant. Note that we expected higher scores on the dimensions to lead to stronger, positive interaction effects. This hypothesis was based on the assumption that (according to the RSM), gamers that score high on all three dimensions experience detrimental effects from playing video games.

Methods Study 2.

Sample and procedure. We did the data collection for Studies 1 and 2 simultaneously for the convenience of the participants. Therefore, the sample (*N* = 164) and the procedure for both studies were the same as presented in Study 1.

Outcome variables and materials. We employed two pairs of outcome variables. One pair measured health outcomes--work/study engagement and (study) burnout. The other measured behavioural outcomes--prosocial tendencies and aggressive behaviour. We measured work/study engagement with the Utrecht Work Engagement Scale, or UWES, constructed by Schaufeli and Bakker (2004), whereas burnout was measured with the Utrecht Burnout Scale, or UBOS, which was constructed by Schaufeli and Van Dierendonck (2000). We measured prosocial tendencies with the Prosocial Tendencies Measure, or PTM, constructed by Carlo and Randall (2002). Aggression was measured with the Buss-Perry Aggression Questionnaire-Short Form, or BPAQ-SF. Originally constructed by Buss and Perry (BPAQ; 1992), Bryant and Smith (2001) then improved on its design, creating the BPAQ-SF. Table 9 presents an overview of these measures.

Table 9. *The Intended Outcome Variables used to Test the Predictive Validity of the RSQ-18.*

Outcome	Functional (short-term reality substitution)	Dysfunctional (long-term reality substitution)
Health	*Work/Study Engagement* (UWES) Schaufeli & Bakker, 2004 15 Items, 7-Point Likert Scale Validated by Schaufeli & Bakker (2004)	*(Study) Burn-out* (UBOS) Schaufeli & Van Dierendonck, 2000 15 Items, 7-Point Likert Scale Validated by Schaufeli, Martínez, Pinto, Salanova & Bakker (2002)
Behaviour	*Prosocial Behaviour* (PTM) Carlo & Randall, 2002 23 items, 5-Point Likert Scale Validated by Carlo & Randall (2002)	*Aggression* (BPAQ-SF) Buss & Perry, 1992; Bryant & Smith (2001) 12 Items, 6-Point Likert Scale Validated by Bryant and Smith (2001)

Results Study 2. We used multiple regression analyses (forced entry method), to determine the predictive validity of the RSQ-18. First, we checked the assumptions. We checked for univariate outliers amongst the variables by looking at the boxplots. Though we detected outliers, all plots of the standard residuals indicated normality so they were not of concern (Field, 2009). Variables that showed extreme scores (average duration and the variable for the interaction effect between

intensity and duration) benefitted of natural log transformations (as done in Study 1). Only one extreme score remained on the duration scale. The participant in question solely scored in the lowest category of this skewed scale. Since this score was valid, we did not remove it from the sample. No transformation of any kind however (log, inverse, square root) was successful at redistributing the average score on the frequency scale, the average score on the duration scale and the average score on the satisfaction scale. These three (and several interaction effect-variables) violate the normality-assumption of the predictors. However, the normal probability plot of standardised residuals as well as the scatterplot of standardized residuals against standardized predicted values after every regression met the assumptions of normality, linearity and homoscedasticity of residuals. As for the outcome variables, the average score on the UBOS and the average score on the BPAQ violated the assumption of normality. However, we remedied this with a natural log transformation of both variables.

Good Durbin-Watson values indicated uncorrelated residuals. The Mahalanobis distance did not at any point exceed the critical χ^2 for $df = k$ limit at $\alpha = .001$, so multivariate outliers were not a concern. Lastly, we centered the predictors and used these to compute the interaction effect variables. This way, we avoided multicollinearity, as the interaction effects would otherwise have strongly correlated to the individual predictors. The mean and standard deviation of the scores on the scales of the RSQ-18, the RSQ-18 itself and the outcome measures can be found in Table 10.

Table 10. *The Mean and Standard Deviation of the Scores on the Scales of the RSQ-18, the RSQ-18 itself and the UWES, PTM, UBOS and the BPAQ-SF.*

Questionnaire	Statistic	
	M	***SD***
Immersion	4.11	1.13
Empathy	4.09	1.22
Satisfaction	6.12	.76
Frequency	5.20	1.93
Duration	3.38	.97
Intensity	4.48	.95
RSQ-18	4.52	.79
UWES	3.81	1.24
PTM	2.43	.53
UBOS	3.34	1.08
BPAQ	2.47	.90

Functional Outcomes. The hypotheses for these two measurements were that each dimension of the RSQ would be able to explain a significant amount of variance (UWES: hypothesis 4^A; PTM: hypothesis 4^B). However, we expected no three-way interaction effect to be present. This however is not the case.

Just as with Table 6, Table 11 shows the Spearman correlations (since data is not truly of interval scale and skewed), between the UWES, the PTM and dimensions of the RSQ-18. These are equal to the *б*'s that would have arisen from linear regressions, showing in effect the simple effects between these various measures.

Looking at Tables 11 and 12, no dimension or combination of dimensions of the RSQ explains a significant amount of variance on the UWES. Thus, hypothesis 4^A was not confirmed. Note that a positive main effect for duration does almost reaches significance.

The results concerning the PTM show that individually only the RSQ-18 (positive) and the intensity (positive) and frequency (negative) dimensions show significant correlations. Of the intensity dimension, only the empathy scale shows a positive significant relation with the PTM. When using all three dimensions in the regression analysis, each explained a significant amount of variance (intensity: positive; frequency and duration: negative). We found that no interaction effect explains a significant amount of variance on the PTM (Table 12). Thus, we confirmed hypothesis 4^B. Note that only the intensity dimension explains a positive (instead of a negative) amount of variance.

Table 11. *Results of Spearman Correlations between the Predictor Dimensions and Scales of the RSQ-18 and the UWES and PTM.*

Dimension/Scale	Outcome Measure	
	UWES	**PTM**
Immersion	.009	.146
Empathy	-.034	.277***
Satisfaction	-.041	-.057
Frequency	-.118	-.166*
Duration	-.111	-.082
Intensity	-.028	.241**
RSQ-18	-.063	.173*

Note. * = significant at an α of .05; ** = significant at an α of .01; *** = significant at an α <.001.

Table 12. *The Results of the Multiple Regression Analyses with the Average Score on the UWES or the Average Score on the PTM as Dependent Variable and the Scales of the RSQ-18 as the Predictors.*

RSQ-18	Multiple Regression Statistics with the UWES as Dependent Variable							
Predictor(s)	***R^2***	***Adjusted R^2***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***β***	***p***
Intensity					.035	.106	.026	.745
Frequency					-.058	.052	-.093	.265
Duration					.525	.271	-.156	.054
In × Fr	.045	.002	1.040 (7, 156)	.406	-.051	.057	-.073	.370
In × Du					.288	.277	.083	.300
Fr × Du					-.038	.126	-.024	.763
In × Fr × Du					.054	.147	.030	.712
RSQ-18	**Multiple Regression Statistics with the PTM as Dependent Variable**							
Predictor(s)	***R^2***	***Adjusted R^2***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***β***	***p***
Intensity					.167	.043	.300	.000
Frequency					-.051	.021	-.191	.015
Duration					-.245	.109	-.170	.025
In × Fr	.158	.120	4.172 (7, 156)	.000	-.005	.023	-.017	.824
In × Du					.166	.111	.111	.138
Fr × Du					.057	.051	.085	.262
In × Fr × Du					-.088	.059	-.112	.136

Note. Significant correlations (α = .05) have been accentuated. In = Intensity, Fr = Frequency and Du = Duration. "×" indicates an interaction-effect.

***Dysfunctional Outcomes*.** The hypotheses for these two measurements was that only the interaction between all three dimensions (intensity, duration and frequency) could explain a significant amount of variance (UBOS: hypothesis 5^A; BPAQ-SF: hypothesis 5^B). Our data does not confirm our hypotheses.

Just as with Tables 6 and 11, Table 13 shows the Spearman correlations (since data is not truly of interval scale and skewed), between the UBOS, the BPAQ-SF and dimensions of the RSQ-18.

These are equal to the *β*'s that would have arisen from linear regressions, showing in effect the simple effects between these various measures.

Looking at Table 13, the RSQ-18, the intensity and duration dimensions show significant positive correlation with the UBOS. Of the intensity dimension, only the empathy scale shows a significant positive correlation with the UBOS. When using all three dimensions, only intensity shows a significant amount of variance explained on the UBOS. No significant two-way interactions effects were detected. The absence of a significant positive three-way interaction effect between all three dimensions when explaining variance on the UBOS prohibits us from confirming hypothesis 5^{A}.

As for the BPAQ-SF: only the RSQ-18 and the intensity dimensions show significant positive correlations with the BPAQ-SF. Of the intensity dimension, only the immersion and satisfaction scales show a significant positive correlation. Similar to the results on the UBOS when using all three dimensions, only intensity shows a significant amount of variance explained on the BPAQ-SF. We found a two-way interaction affect between intensity and frequency, but no three-way interaction effect. Thus, we did not confirm hypothesis 5^{B}.

Table 13. *Results of Spearman Correlations between the Predictor Dimensions and Scales of the RSQ-18 and the UBOS and BPAQ-SF.*

Dimension/Scale	Outcome Measure	
	UBOS	**BPAQ-SF**
Immersion	.088	.213**
Empathy	.204**	.124
Satisfaction	.129	.218**
Frequency	.122	.143
Duration	.154*	.018
Intensity	.194*	.198*
RSQ-18	.222**	.218**

Note. * = significant at an *α* of .05; ** = significant at an *α* of .01; *** = significant at an *α* <.001.

Table 14. *The Results of the Multiple Regression Analyses with the Average Score on the UBOS or the Average Score on the BPAQ-SF as Dependent Variable and the Scales of the RSQ-18 as the Predictors.*

RSQ-18	**Multiple Regression Statistics with the UBOS as Dependent Variable**							
Predictor(s)	***R²***	***Adjusted R²***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***β***	***p***
Intensity	.060	.018	1.420 (7, 156)	.201	.064	.028	.185	.023
Frequency					.010	.014	.058	.484
Duration					.069	.071	.077	.333
In × Fr					.009	.015	.051	.529
In × Du					.007	.073	.007	.926
Fr × Du					.017	.033	.040	.613
In × Fr × Du					-.021	.039	-.043	.584
RSQ-18	**Multiple Regression Statistics with the BPAQ-SF as Dependent Variable**							
Predictor(s)	***R²***	***Adjusted R²***	***F* (*df1*, *df2*) and *p***		***B***	***se***	***β***	***p***
Intensity	.100	.060	2.487 (7, 156)	.019	.063	.029	.170	.032
Frequency					.026	.014	.145	.074
Duration					-.041	.075	-.043	.585
In × Fr					.032	.016	.158	.046
In × Du					.043	.077	.043	.574
Fr × Du					.049	.035	.109	.162
In × Fr × Du					-.046	.041	-.087	.263

Note. Significant correlations (*α* = .05) have been accentuated. In = Intensity, Fr = Frequency and Du = Duration. "×" indicates an interaction-effect.

Discussion Study 2. This study into the predictive validity of the RSQ-18 indicated that no scale or combination of scales of the RSQ-18 explained a significant amount of variance on the UWES. However, the RSQ-18 and the intensity scale positively relate to prosocial tendencies, burnout and aggression. For the functional outcomes hypothesis 4^A (UWES) was unconfirmed but hypothesis 4^B (PTM) was confirmed. As for the dysfunctional outcomes, both Hypothesis 5^A (UBOS) and 5^B (BPAQ-SF) have not been confirmed.

Alternative explanations. The absence of significance regarding engagement (UWES) most likely traces back to the RSM with its roots in need-theories. The developers of the UWES defined engagement as "*a positive, affect-cognitive state of the highest level of satisfaction*" (Schaufeli & Bakker, 2001). High levels of satisfaction would imply few unsatisfied needs. It could be that people with less satisfied need, experience a weaker form of reality substitution (if it occurs at all). We suspect however that they simply have no desire to play video games at all, as they are already content with their current satisfiers. Another explanation could be that some aspect(s) of being engaged in your work or study protects you from the detrimental effects of (long-term) reality substitution. This would imply a missing third variable in the model.

Prosocial tendencies (PTM) significantly correlates to the empathy scale. Empathy is a strong motive in eliciting prosocial behaviour (Decety, 2011). This could explain the significant correlation we found. However, the items used for the RSQ-18 do not constitute a validated measure for empathy towards real-life people. Thus, we cannot say whether empathy towards a digital character (RSQ-18) refers to the same "empathy" as empathy towards real-life people or not. This however seems a reasonable assumption, one that perhaps needs to be tested in future studies.

Burnout shows a significant correlation with the empathy scale too. Yet, we can make no conclusions regarding causality--burnout or high levels of empathy towards one's character in the video game. We suspect that people with a burnout, have stronger needs than average, as they are not in a state of happiness nor content. The UWES essentially measures a state of being content. Schaufeli and Bakker (2004) noted that, though both are principally independent of each other, people would experience engagement and burnout as being opposite psychological states. Therefore, the UBOS essentially measures a state of discontent. Thus, increased needs that come with being discontent, may lead to the gamer empathizing more strongly with their characters in video games.

Aggression does not significantly correlate to the empathy scale. According to the meta-analysis of Vachon, Lynam and Johnson (2013), aggression and empathy are only weakly (negatively) related. That immersion does correlate with aggression, might due to decreased attention control in people who score high on aggression. In their study, Muris, Van Der Pennen, Sigmund and Mayer (2008) showed that aggression is negatively related to attention control. Perhaps people with poor attention control are at risk of "severe" immersion. They slip into the game only to be unable to take their attention of it again, akin to a sort of tunnel vision perhaps. This could then lead (probably mediated through satisfaction) to the increased frequency of play, which also correlates with aggression.

With three of our four hypotheses for Study 2 unconfirmed, we are unable to explain fully how video gaming leads to functional outcomes in one context and to dysfunctional outcomes in another. Merely playing often or playing for long periods of time is not enough to explain a significant amount of variance on the dysfunctional outcomes (Table 14). Nonetheless, these dimensions were able to explain a significant negative amount of variance on the PTM. Moreover,

since these two dimensions (frequency and duration) correlate to the intensity dimension, we would advise they remain part of the RSM. Perhaps we should combine them in a single temporal dimension. Intensity is able to explain a significant amount of variance of dysfunctional outcomes. Perhaps the key to differentiating between functional and dysfunctional outcomes lies in the interaction between the scales of the intensity dimension. This would make frequency and duration variables that moderate the strength or 'severity' of the intensity scale. In the following General Discussion, we will elaborate more on this point.

Limitations and future directions. At this time, we cannot say whether combining the frequency and duration dimensions has any value. Nonetheless, the recurring significance of the intensity scale alone is enough to warrant further study. The absence of a three-way interaction effect between the dimensions indicates that gamers, who score higher on all three dimensions, are not more at risk of dysfunctional outcomes. However, limiting the frequency and duration of play, may still be beneficial, as it may limit the detrimental effects on prosocial tendencies. This may also prevent the interaction effect between frequency and intensity on aggression. The next question would be how to limit the intensity of play, which positively relates to prosocial tendencies (functional), but also to burnout and aggression (dysfunctional). To answer that, we will need more data exploring the scales of the intensity dimension and possible third variables. As with Study 1, note that when interpreting these results the immersion scale has a less than optimal reliability. In addition, especially the scores on the duration scale and the frequency scale show skewedness. However, given the use of convenience sampling, primarily on hardcore gamer-dedicated websites, this comes as no surprise. Lastly, given the fact that the results are correlational in nature, our findings do not represent causal relationships.

General Discussion.

Summary. This study had as its objectives (1) the construction of a valid questionnaire to measure reality substitution and (2) testing of the predictive validity of this measure (and by extension part of the Reality Substitution Model). In this, we have been partially successful.

In Study 1, we confirmed that, with the exception of the immersion scale (.616), the empathy and satisfaction scales have good internal consistencies, factor loadings, and test-retest reliability. The same applies to the RSQ-18 as a whole, which our data indicated had convergent as well as divergent validity. We also checked the internal structure as proposed in the Reality Substitution Model. According to our data, immersion positively relates to empathy. Empathy positively relates to satisfaction (full mediation of immersion). Empathy and satisfaction positively relate to frequency. Finally, the interaction effect between immersion and empathy negatively relates to duration, whereas a stronger interaction effect between immersion and satisfaction positively relates to duration. Intensity positively relates to frequency and duration. Frequency does not appears to relate significantly to duration. In conclusion, we now have a validated measure for reality substitution.

In Study 2, the RSQ-18 showed that playing intensely, frequently and for long periods (three-way interaction effect) does not lead to a significant increase in dysfunctional outcomes (burnout and aggression), nor in functional outcomes (engagement and prosocial tendencies). Engagement showed no significant relation to reality substitution (or its scales). All other outcomes positively relate to the intensity dimension and reality substitution itself. However, no clear pattern was discernible that indicates when reality substitution is leading to (primarily) functional or

dysfunctional outcomes. Perhaps both occur simultaneously. For instance, our data allows for a scenario wherein reality substitution aides in overcoming burnout whilst making the gamer more aggressive. It also allows for a scenario wherein playing video games increases social tendencies, as well as aggression. Thus, the aforementioned contradictions in video game-literature remain without explanation. However, we have managed to inspect the predictive validity of the RSQ-18 with good results. Moreover, in doing so, we have connected both functional and dysfunctional outcomes to a single theoretical mechanism: reality substitution. By doing so, we have come closer to an explanation.

Concerning the Reality Substitution Model and the Reality Substitution Questionnaire. At the end of this study, it would appear that the reality substitution model is incomplete; no clear pattern came forward from comparing (the dimensions) of the RSQ-18 with the functional and dysfunctional outcomes. However, despite the less-than-optimal reliability of the immersion scale, the intensity and its relations towards our outcomes in Study 2 offered us new insights. Especially our empathy scale was not only reliable but also showed many significant correlations, suggesting it may well be of importance in any psychological model of video gaming. Future studies on the RSM and reality substitution expand on this dimension.

We also showed that playing video games intensely, frequently and for prolonged periods (three-way interaction effect) does not significantly increase the risk of dysfunctional outcomes. However, our data also shows frequency and duration to relate negatively to prosocial tendencies. This gives credence to the 79% of parents that place time limits on the video game playing of their child (ESA, 2015). Though it may not fully protect against dysfunctional outcomes, this precaution may well moderate the severity of the dysfunctional outcomes. Note however, that playing infrequently and/or shortly but nonetheless intensely, can still put a gamer at risk of dysfunctional outcomes. As for functional outcomes, the results are slightly more ambiguous, as no clear pattern between the two functional outcomes and the predictors was discernible. This raises the question whether perhaps there are one or more third-variables unaccounted for in the model. We recommend further study.

Alternative explanations. We have formulated the following alternative explanations. Whilst building on our own findings, these may explain the discrepancies between the literature that concludes playing video games leads to functional outcomes and the literature that concludes playing video games leads dysfunctional outcomes.

***Identity substitution*.** According to the RSM, empathy functions as a 'bridge' between gamer and the character in the game, over which experiences and skills are 'exchanged' (Rosenkrantz, 2015). Perhaps, the nature of the outcomes of video gaming would dependent on to whom 'this bridge leads'. In other words, with what kind of character you empathize. In that case, the likelihood of dysfunctional outcomes increases when empathizing with renegade characters. For instance, those that kill monsters 'for the fun of it'. Contrary to that, the likelihood of functional outcomes would increase when empathizing with paragon characters, who kill monsters to protect family and friends for instance. A gamer might (temporarily) 'take over' some of the (personality) traits, habits and/or cognitions their character, after frequent and/or prolonged empathy. Subsequent changes in health and behaviour would be as spillover effects from this empathy. It may lead to gamers heroically standing up to a bully three times their size (paragon), or have the same gamers firing guns in public places (renegade). This spillover may subside after a period of not gaming (at least not with

similar characters). Alternatively, it may not (for instance in the case of frequent and prolonged gaming). Then, these newly 'acquired' behaviours and cognitions may become ingrained. In this case, reality substitution has the effect of "identity substitution". In line with this alternative explanation, Yee & Bailenson (2007) already researched something called the Proteus-effect. How the character looks, influences how the gamer behaves in the virtual world. Reality (or in this context identity) substitution would be a broader opposite of that: how a character or avatar thinks, feels and behaves influences (via empathy supposedly) how the gamer thinks, feels and behaves in the real world. To test this explanation, we would need to compare gamers in behaviour, after subjecting them (frequently and prolonged) to either paragon or renegade characters in video games.

Conditioning and desensitising. A spin on identity substitution (taking over traits and characteristics of your character), is that of reality substitution may be more akin to conditioning and desensitising. In 2011, Li, Liau and Khoo showed that pathological gaming could be predicted by escapism. That people play video games because of unfulfilled needs is the basic premise of the RSM. A need of the gamer can be to escape from reality by being someone else and do things differently in the virtual world. In the video game, they may hit a bully they would avoid in real life. After all, video games offer a safe-to-fail environment. This goes beyond mere exposure, as gamers have to want and choose to exhibit the behaviour in question. This conduct may desensitise gamers over time (frequency and duration), to the particular behaviours they have their characters exhibit. It may even give rise to new habits. Habits refer to behavioural patterns enacted automatically when certain cues (e.g. the presence of a bully) are present. The gamer exhibits these behaviours again automatically because they gave good results a similar context in the past (Lally & Gardner, 2013). This conditioning and desensitising may even interact with the empathy from the RSM—stronger empathy towards the characters increases the effect from conditioning and desensitising. When the gamer then come across the real-life bully, the gamer may now be more likely to exhibit the behaviours they already exhibited in the video games. In this case, we would need to observe an increased occurrence of predetermined behaviours. Behaviours that occur in the video games they would play as well. Initially, these behaviours would come without a change in cognitions, which might follow as feedback on the execution of these 'new' behaviours.

Predispositions to psychopathology. A difference in functional and dysfunctional outcomes can originate in predispositions to psychopathology. Piaget (1932) proposed that children go through several stages in the development of their morality. For instance, if a driver swerves to avoid a dog and then knocks over a road sign, adults are likely to be less angry about the knocked over road sign, as opposed to when the driver had knocked it over just for fun. Though the outcome is the same, adults are more forgiving because of the good intention (saving the dog). Young children however do not pay attention to the intent, only to the outcome. They would still feel the driver deserved punishment for flattening the road sign. Our data suggests a weak (but significant) correlation with psychopathology such as paranoia and psychotism, as our sample did not score collectively '0'. This indicates that at least some symptoms were present in our sample. Perhaps dysfunctional outcomes are (only) more likely in an under- or abnormally developed (part of the) brain. Testing this explanation could be costly in comparison with the previously offered explanations, as it would most likely require imaging the brain of participants in some manner. Regardless, predispositions (to psychopathology) could explain the strength of certain outcomes and even underpin susceptibility to reality substitution and/or the other explanations.

Limitations. This study was the first to test some of the merits of (part of) the RSM. There is still some room for improvement, for instance its ability to differentiate between functional and dysfunctional outcomes. These current limitations create opportunity for future studies.

The sample. Consider the sample for instance. Dropout rates were high, mostly because people did not meet the criteria of being a gamer (playing regularly and recently) or because of the large amount of items that required an answered (seven questionnaires, 155 items). Perhaps results will differ when selecting gamers on the genre of video games they play most often. Inclusion of non-gamers in the sample could prove to give further insight as well. Should they score low on reality substitution, it would indicate reality substitution is unique to video gaming. Should they score high on reality substitution, it could mean that perhaps other means beyond video games can induce reality substitution. Alternatively, perhaps reality substitution is still too narrow in definition; some third variable may be missing from the model.

Moreover, the sampling method (convenience sampling) limits participants to those whom are regular visitors to blogs and forums, which not all gamers might do. Having said that, the sample used for both studies did contain participants from many, primarily Western, countries. Inclusion of children (aged below 18) and possibly elderly may be an option for future research. Regardless, increasing the heteronormativity and size of the sample, though perhaps costly, will benefit the studies into reality substitution and the RSM.

Variables and materials. Besides changes to the sample, changes to the materials used in this study could prove beneficial too. We recommend the usage of pairs of variables that are at opposing ends of a spectrum (one functional, one dysfunctional). Preferably, variables that are well researched, simply defined and easily observed/measured. We did not find another study of this design, yet it may well be crucial in explaining how one medium such as video games, can lead to two various outcomes.

In addition, we strongly recommend using large response scales for the inquiring into the frequency and duration of play. These variables showed skewedness in our analyses. Frequency of play can vary greatly between video gamers, whom also play much longer a time than expected. The skewed data could have affected the outcomes of this study, such as the apparent irrelevance of the duration scale. The simpler and more observable our variables and the more accurate our measurements, the better our results will be.

Design and the third-cause fallacy. We also have some recommendations for future research design. We suggest operationalizing reality substitution in the form of an intervention. By measuring burnout, prosocial tendencies, aggression and alike, before and after repeated reality substitution, we might learn more about causal relations of reality substitution. Increases in prosocial tendencies and decreases in aggression for instance, would make reality substitution a positive result following video gaming. The reverse would make reality substitution a negative result following video gaming.

Secondly, we recommend testing the entirety of the RSM. We are still unable to draw conclusions as to how reality substitution exactly leads to functional and dysfunctional outcomes. It is possible that the process preceding reality substitution (read Rosenkrantz, 2015) holds the key. In other words, control for needs and video game features in future reality substitution studies. This could require selecting specific groups of gamers based on the need that motivates them to play. These (as well as control groups), would then have to be matched to games with features that complement their needs (or not). In the case of a match, we would expect to see significantly higher

scores of reality substitution than when there is a mismatch. From here, we could study whether the significant amount of variance explained on outcomes (or on individual reality substitution scales) changes due to these (mis)matches. For instance, perhaps only gamers who play for the catharsis (release of aggression), see significant changes in their aggressive behaviour because of reality substitution. We are confident that these changes to the design will lead to new insights into video games and the effects they have on us.

Lastly, we would like to point out that video games may be the victim of a third-cause fallacy of societal scale. Inspection of our regression analyses indicates that in particular the intensity dimension has a prominent role in predicting outcomes of video gaming. In short, playing intensely (*A*) appears to be the predictor of outcomes such as aggression (*B*). Intensity also relates positively to frequency and duration (*C*). *A* causes *B* and *C* simultaneously. However, since the intensity dimension is less perceivable by outsiders than aggression, the frequency and duration of play, outsiders see *C* resulting in *B*. In other words, people may see frequent/long gameplay as the cause of outcomes such as aggression. The result is similar to measuring a fever and blaming it on your runny nose. In truth, a virus such as influenza caused the fever and the runny nose. Take burnout for instance. A burnout may lead to intense gameplay, which then leads to frequent and long gameplay. Playing frequently and for long periods does not necessarily lead to dysfunctional outcomes. It may in fact be the result of dysfunctional outcomes via intensity. Thus, we need to increase the amount of longitudinal studies into gaming to prevent mistakes such as the third-cause fallacy.

Future directions and societal relevance. More research on shared personality traits of gamers, possible physiological and developmental differences in the brain (those that precede gaming, not those that result from gaming) and other social-emotional or intellectual differences between gamers and non-gamers may show greater insight into reality substitution as well. Moreover, should reality substitution really do appear to be pivotal in determining outcomes of playing video games, the following step might be to inspect whether or not reality substitution is limited to video games. It may turn out that its effects may also arise from music, TV, books and alike.

The need for more research on video gaming would be to the benefit of multiple parties (King, Delfabbro & Griffiths, 2010), each with their own purpose: the psychologists, whom would learn more about the dynamic role of rapidly evolving technology as new human behaviours develop. We would also solve the contradictions we currently have in our literature regarding video games. Then there are the gamers, whom would benefit from education on potential health risks as well as benefits that can results from playing video games. Finally the gaming industry, that aims to develop ever better games that are to give a more appealing and rewarding experience leading to increased customer satisfaction and loyalty. Our studies into the RSM and the RSQ-18 have brought us one step closer to achieving these goals.

References.

Adachi, P.J., & Willoughby, T. (2013). More than just fun and games: The longitudinal relationships between strategic video games, self-reported problem solving skills, and academic grades. *Journal of Youth and Adolescence*, *42* (7), 1041-1052.

Ahn, S.J., Le, A.M.T., & Bailenson, J. (2013). The effect of embodied experiences on self-other merging, attitude, and helping behavior. *Media Psychology, 16 (1)*, 7-38.

Alderfer, C.P. (1969). An empirical test of a new theory of human needs. *Organizational Behavior and Human Performance, 4 (2)*, 142-175.

Allen, P.J., & Bennett, K. (2010). *PASW statistics by SPSS: A practical guide: Version 18.0.* South Melbourne, VIC: Cengage Learning.

Anderson, C.A., Shibuya, A., Ihori, N., Swing, E.L., Bushman, B.J., Sakamoto, A., Rothstein, H.R., & Saleem, M. (2010). Violent video game effect on aggression, empathy and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin, 136,* 151-173.

Anderson, C.A., & Bushman, B.J. (2001). Effect of violent video games on aggressive cognition, aggressive affect, physiological arousal and prosocial behavior: A meta-analytic review of the scientific literature. *Psychological Science, 12,* 353-359.

Basak, C., Boot, W.R., Voss, M.W., Kramer, A.F. (2008) Can training in a real-time strategy video game attenuate cognitive decline in older adults? *Psychological Aging, 23,* 765-777.

Bavelier, D. (Host). (19 November 2012). Daphne Bavelier: Your Brain on Video Games. TED Talks. [Webcast]. New York, NY: TED Conferences, LLC. Retrieved from: https://www.youtube.com/watch?v=FktsFcooIG8

Baumeister, R.F., Bratslavsky, E., Muraven, M., & Tice, D.M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74 (5),* 1252–1265.

Beullens, K., Roe, K., Van Den Bulck, J. (2008). Video games and adolescents' intentions to take risks in traffic. *Journal of Adolescent Health, 8 (43),* 87-90.

BioWare Corp. (2012). *Mass Effect Trilogy* [Video game]. Redwood City: Electronic Arts, Inc.

Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M., & Pidruzny, J.N. (2009). The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology, 45 (4)*, 624-634.

Borsiak, P., Bouikidis, A., Liersch, R., Russell, J.B. (2008). Cardiovascular effects in adolescents while they are playing video games: A potential health risk factor? *Psychophysiology, 45,* 327-332.

Bryant, F.B., & Smith, B.D. (2001). Refining the architecture of aggression: A measurement model for the Buss Perry Aggression Questionnaire. *Journal of Research in Personality, 35 (2)*, 138-167.

Buss, A.H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, *63 (3)*, 452.

Carlo, G., & Randall, B.A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence*, *31* (1), 31-44.

CD Projekt RED. (2015). *The Witcher 3: Wild Hunt* [Video game]. Warsaw: CD Projekt.

Clark, L.A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological assessment, 7 (3)*, 309.

Derogatis, L.R. (1992). *SCL-90-R: Administration, Scoring and Procedures Manual for the R (evised) Version and other Instruments of the Psychopathology Rating Scale Series.* Clinical Psychometric Research.

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.

Decety, J. (2011). The neuroevolution of empathy. *Annals of the New York Academy of Sciences, 1231 (1)*, 35-45.

Dye, M.W.G., Green, C.S., & Bavelier, D. (2009). Increasing speed of processing with action video games. *Current Directions in Psychological Science, 18 (6),* 321-326.

ESA. (2015). *2015 Sales, Demographic and Usage Data: Essential Facts about the Computer and Video Game Industry*. Retrieved from: http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf

Ferguson, C.J., & Kilburn, J. (2009). The public health risks of media violence: a meta-analytic review. *Journal of Pediatrics, 154,* 759-763.

Ferguson, C.J., San Miguel, C., Garza, A., & Jerabeck, J.M. (2011). A longitudinal test of video game violence influences on dating and aggression: A 3-year longitudinal study of adolescents. *Journal of Psychiatric Research,* 1-6.

Field, A. (2009). Discovering statistics using IBM SPSS statistics. London: Sage Publications.

Fish, M.T., Russoniello, C.V., & O'Brien, K. (2014). The efficacy of prescribed casual videogame play in reducing symptoms of anxiety: a randomized controlled study. *Games for Health Journal, 3 (5),* 291-295.

Floridi, L. (2005). The philosophy of presence: From epistemic failure to successful observation. *Presence: Teleoperators and Virtual Environments, 14 (6),* 656-667.

Freina, L., & Ott, M. (2015). *The 11th International Scientific Conference eLearning and software for Educatio Bucharest*.

Granic, I., Lobel, A., & Engels, R.C.M.E. (2014). The Benefits of playing video games. *American Psychologist, 69 (1),* 66-78.

Gee, J.P. (2003). *What Video Games have to Teach Us about Learning*. New York: Palgrave.

Gentile, D.A., Lynch, P.J., Linder, J.R., & Walsh, D.A. (2004). The effects of violent video game habits on adolescent hostility, aggressive behaviors and school performance. *Journal of Adolescence, 27*, 5-22.

Glass, B.D., Maddox, W.T., & Love, B.C. (2013). Real-time strategy game training: Emergence of a cognitive flexibility trait. PLoS One, 8 (8): e70350.

Goldstein, N.J., & Cialdini, R.B. (2007). The spyglass self: A model of vicarious self-perception. *Journal of Personality and Social Psychology, 92 (3)*, 402.

Green, C.S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423,* 534-537.

Griffiths, M.D., & Hunt, N. (1998). Dependence on computer games by adolescents. *Psychological Reports, 82,* 475-480.

Jackson, L.A., Witt, E.A., Games, A.I., Fitzgerald, H.E., von Eye, A., & Zhao, Y. (2012). Information technology use and creativity: Findings from the Children and Technology Project. *Computers in Human Behavior, 28,* 370-376.

Kasteleijn-Nolst Trenite, D.G., Martins da Silva, A., Ricci, S., Rubboli, G., Tassinari, C.A., Lopes, J., Bettencourt M., Oostings, J., & Segers, J.P. (2002). Video games are exciting: A European study of video game induced seizures and epilepsy. *Epileptic Disorders*, *4*, 121-128.

King, D., Delfabbro, P.H., & Griffiths, M. (2010). Video game structural characteristics: A new psychological taxonomy. *International Journal of Mental Health Addiction, 8,* 90-106.

King, D., & Delfabbro, P.H. (2013). Issues for DSM-5: Video-gaming disorder? *Australian & New Zealand Journal of Psychiatry, 47,* 20-22.

Kirsh, S.J., & Mounts, J.R. (2007). Violent video game play impacts facial emotion recognition. *Aggressive Behavior, 33,* 353-358.

Kühn, S., Lorenz, R., Banaschewski, T., Barker, G.J., Büchel, C., Conrod, P.J., Flor, H., Garavan, H., Ittermann, B., Loth, E., Mann, K., Nees, F., Artiges, E., Paus, T., Rietschel, M., Smolka, M.N., Ströhle, A., Walaszek, B., Schumann, G., Heinz, A., Gallinat, J., & The IMAGEN Consortium. (2014). Positive association of video game playing with left frontal cortical thickness in adolescents. *PLoS ONE, 9 (3)*, e91506.

Lally, P., & Gardner, B. (2013). Promoting Habit formation. *Health Psychology Review, 7,* 137-158.

Lanningham-Foster, L., Jensen, T.B., Foster, R.C., Redmond, A.B., Walker, B.A., Heinz, D., & Levine, J.A. (2006). Energy expenditure of sedentary screen time compared with active screen time for children. *Pediatrics, 118 (6)*, 1831-1835.

Li, D., Liau, A., & Khoo, A. (2011). Examining the influence of actual-ideal self-discrepancies, depression, and escapism, on pathological gaming among massively multiplayer online adolescent gamers. *Cyberpsychology, Behavior, and Social Networking, 14 (9)*, 535-539.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3 (2)*, 0.

Maslow, A. (1954). *Motivation and Personality*. New York, NY: Harper.

Müller, B.C., van Baaren, R.B., van Someren, D.H., & Dijksterhuis, A. (2014). A present for Pinocchio: On when non-biological agents become real. *Social Cognition, 32 (4)*, 381-396.

Muris, P., van der Pennen, E., Sigmond, R., & Mayer, B. (2008). Symptoms of anxiety, depression, and aggression in non-clinical children: Relationships with self-report and performance-based measures of attention and effortful control. *Child Psychiatry and Human Development, 39 (4)*, 455-467.

O'Neill, S. (2005). Presence, Place and the Virtual Spectacle. *PsychNology Journal, 3 (2)*, 149-161.

Paras, B., & Bizzocchi, J. (2005). Game, motivation, and effective learning: An integrated model for educational game design.

Piaget, J. (1932). *The Moral Judgment of the Child.* New York, NY: New York Free Press.

Pijnenborg, G.H.M., Spikman, J.M., Jeronimus, B.F., & Aleman, A. (2013). Insight in schizophrenia: Associations with empathy. *European Archives of Psychiatry and Clinical neuroscience, 263 (4),* 299-307.

Primack, B.A., Carroll, M.V., McNamara, M., Klem, M.L., King, B., Rich, M.O., Chan, C.W., & Nayak, S. (2012). Role of video games in improving health-related outcomes: A systematic review. *American Journal of Preventive Medicine, 42 (6),* 630-638.

Prinz, U., Nutzinger, D.O., Schulz, H., Petermann, F., Braukhaus, C., & Andreas, S. (2013). Comparative psychometric analyses of the SCL-90-R and its short versions in patients with affective disorders. *BMC Psychiatry, 13 (1)*, 104.

Przybylski, A.K., Rigby, C.S., & Ryan, R.M. (2010). A motivational model of video game engagement. *Review of General Psychology, 14,* 154-166.

Redmond, B.F. (2009). *Work Attitudes and Job Motivation Home* (D. Camburn, Ed.). Retrieved from: https://wikispaces.psu.edu/display/psych484/2.+need+theories

Rigby, S., & Ryan, R. (2007). *Rethinking Carrots: A New Method for Measuring What Players Find Most Rewarding and Motivating about Your Game*. Retrieved from: http://www.gamasutra.com/view/feature/130155/rethinking_carrots_a_new_method_.php?print=1

Rosenkrantz, N.V. (2015). Reality substitution: matching needs to video game features to explain the effects of video gaming on health and behaviour. *Unpublished.*

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 (4)*, 736.

Ryan, R.M., Rigby, C.S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30 (4)*, 344-360.

Schaufeli, W.B., & Van Dierendonck, D. (2000). *Utrechtse Burnout Schaal (UBOS): Testhandleiding*. Lisse: Swets Test Services

Schaufeli, W.B., & Bakker, A.B. (2001). Werk en welbevinden: Naar een positieve benadering in de Arbeids- en Gezondheidspsychologie. *Gedrag & Organisatie, 14*, 229-253.

Schaufeli, W.B., & Bakker, A.B. (2004). Bevlogenheid: een begrip gemeten. *Gedrag & Organisatie, 17*, 89-112.

Schaufeli, W.B., Martínez, I.M., Pinto, A.M., Salanova, M., & Bakker, A.B. (2002). Burnout and engagement in university students a cross-national study. *Journal of Cross-Cultural psychology, 33 (5)*, 464-481.

Sherry, J.L., Lucas, K., Greenberg, B.S., & Lachlan, K. (2006). Video Game Uses and Gratifications as Predictors of Use and Game Preference. *Playing video games: Motives, responses and consequences,* 213-224.

Slater, M. (2003). A note on presence terminology. *Presence Connect, 3 (3)*, 1-5.

Snead, J. (Director). (2014). *Video Games: The Movie* [Motion picture]. United States: Mediajuice Studios.

Spreng, R.N., McKinnon, M.C., Mar, R.A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment, 91 (1)*, 62-71.

Stevens, J.P. (2002). *Applied Multivariate Statistics for the Social Sciences* (4th ed.). Hillsdale, New Jersey: Erlbaum.

Stoffregen, T.A., Faugloire, E., Yoshida, K., Flanagan, M.B., & Merhi, O. (2008). Motion sickness and postural sway in console video games. *Human Factors, 50,* 322-331.

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements, 2 (4)*, i.

Van den Brink, W.P., & Mellenbergh, G.J. (1998). *Testleer en Testconstructie.* Meppel Boom Koninklijke Uitgevers.

Vachon, D.D., Lynam, D.R., & Johnson, J.A. (2014). The (non) relation between empathy and aggression: Surprising results from a meta-analysis. *Psychological Bulletin, 140 (3)*, 751.

Witmer, B.G., & Singer, M.F. (1994). Measuring presence in virtual environments (No. ARI-TR-1014). Army Research Institute for the Behavioral and Social Sciences Alexandria VA.

Yee, N. (2006). The Demographics, Motivations and Derived Experiences of Users of Massively-Multiuser Online Graphical Environments. *PRESENCE: Teleoperators and Virtual Environments, 15*, 309-329.

Yee, N. (2007). Motivations of Play in Online Games. *Journal of CyberPsychology and Behavior, 9*, 772-775.

Yee, N., & Bailenson, J. (2007). The Proteus Effect: The effect of transformed self-representation on behavior. *Human Communication Research, 33*, 271-90.

Waterworth, J.A. & Waterworth, E.L. (2003). The Meaning of Presence. Retrieved from: http://www8.informatik.umu.se/~jwworth/PRESENCE-meaning.htm

Williams, C. A. (1989). Empathy and burnout in male and female helping professionals. *Research in Nursing & Health, 12 (3)*, 169-178.

Zapata, A.L., Moraes, A.J., Leone, C., Doria-Filho, U., & Silva, C.A. Pain and musculoskeletal pain syndromes related to computer and video game use in adolescents. *European Journal of Pediatrics, 165,* 408-414.

Appendices.

Appendix 1.

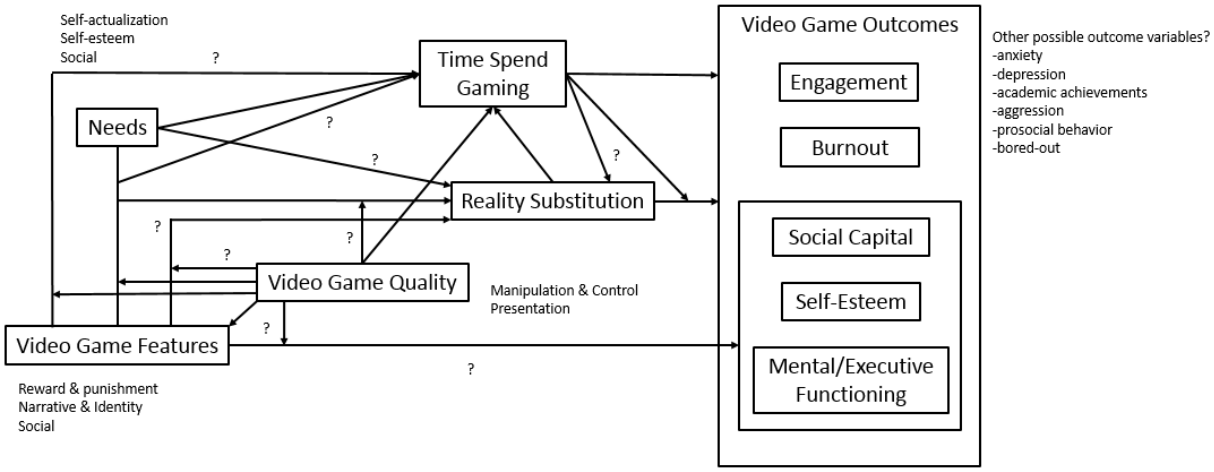

Figure 2. An early version of the reality substitution model from the developmental stages of the RSM (Rosenkrantz 2015)

Appendix 2. The final version of the intensity scale of the Reality Substitution Questionnaire-18 (RSQ-18). First in English, then in Dutch.

Instructions: these following statements concern your experiences whilst playing video games. Take your time to read these at pace. Your task will then be to indicate, per statement, to what degree they apply **to you**. There are no right or wrong answer and no trick questions.

Response Choices						
1	2	3	4	5	6	7
(Almost) Never True			Sometimes True			(Almost) Always True

#	Item	Response Choices						
3	When the actions of my character are praised, I experience positive emotions such as joy.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
9	When my character is being treated unfairly, I experience emotions such as anger or sorrow.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
12	Video games enrich my life.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
13	Whilst playing video games, I forget that I am only looking at a screen and operating a controller or keyboard.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
14	The character that I play in a video game, feels as an extension of myself.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
15	Time flies when I am playing video games.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
16	I become invested in the characters of my video games and their stories.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
18	I wonder what it would be like to stand in my character's shoes at various times during the game.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
19	I very much enjoy playing video games.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
21	I would describe playing video games as an interesting activity.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
24	I know how my character feels, even when this is not explicitly mentioned or stated.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
27	The relations that my character has in the video games (such as those with other gamers online or with NPC's, non-playable characters), feel like relations of my own.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
30	I do not notice that I am getting tired/exhausted whilst playing video games. Only when I stop or after long periods of time (hours).	**1**	**2**	**3**	**4**	**5**	**6**	**7**
32	I share the emotions that my character is experiencing.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
33	Once I am immersed into a video game, I am no more aware of what is happening around me whilst I am playing (for instance, others have trouble getting my attention).	**1**	**2**	**3**	**4**	**5**	**6**	**7**

Instructies: hier volgen een aantal stellingen betreffende uw ervaringen met videospellen en tijdens het spelen van videospellen. Lees deze op uw gemak door. Daarna is het aan u de taak om per stelling aan te geven in hoeverre deze voor u waar is. Er zijn geen foute of goede antwoorden en geen strikvragen.

Antwoordmogelijkheden						
1	2	3	4	5	6	7
(Bijna) Nooit Waar			Soms waar			(Bijna) Altijd Waar

#	Item	Antwoordmogelijkheden						
3	Wanneer de handelingen van personage worden geprezen, ervaar ik positieve emoties zoals blijdschap.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
9	Als mijn personage oneerlijk wordt behandeld, ervaar ik negatieve emoties, zoals woede of verdriet.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
12	Videospellen verrijken mijn leven.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
13	Tijdens het spelen van videospellen, vergeet ik dat ik slechts naar een scherm kijk en met een controller/toetsenbord speel.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
14	Het personage dat ik speel in een videospel, voelt als een extensie van mijzelf.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
15	Wanneer ik een videospel speel, vliegt de tijd voorbij.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
16	Ik raak betrokken bij mijn personages in een videospel en hun verhaal.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
18	Ik vraag mij af, hoe het zou zijn om in de schoenen van mijn personage te staan op verscheidene momenten tijdens het spel.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
19	Ik geniet erg van het spelen van videospellen.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
21	Ik zou het spelen van videospellen als een interessante activiteit omschrijven.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
24	Ik weet hoe mijn personage zich voelt, zelfs wanneer dat niet expliciet gezegd of vermeld wordt.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
27	De relaties die mijn personage heeft in een videospel (bijvoorbeeld met andere gamers online of met NPC's, non-playable characters), voelen als mijn relaties die van mij zijn.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
30	Ik merk niet dat ik moe word en uitgeput raak tijdens het spelen van een videospel. Slechts wanneer ik stop of na lange duur (uren).	**1**	**2**	**3**	**4**	**5**	**6**	**7**
32	Ik deel de emoties die mijn personage ervaart.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
33	Eenmaal in een videospel verzonken, heb ik niet meer door wat er om mij heen gebeurd tijdens het spelen (bijvoorbeeld dat anderen mijn aandacht niet kunnen krijgen).	**1**	**2**	**3**	**4**	**5**	**6**	**7**

Appendix 3. The items for the frequency and duration scales with the used response scales for the Reality Substitution Questionnaire-18 (RSQ-18). First in English, then in Dutch.

Scale	Item	Response Choices
Frequency	How many days per week do you play video games on average?	(Barely) None at all 1 Day per week 2 Days per week 3 Days per week 4 Days per week 5 Days per week 6 Days per week (Almost) Every day of the week
	On how many days of the past four weeks, including today, did you play video games? Choose a number ranging from 1 to 28.	[Number from 1 to 28, representing the number of days of those four weeks on which the participants played video games]
Duration	What is the average duration of your video gaming sessions?	Barely 10 minutes 1 Hour 2 Hours 3 Hours 4 Hours 5 Hours 6 Hours 7 Hours 8 Hours or longer
Schaal	**Item**	**Antwoordmogelijkheden**
Frequentie	Hoeveel dagen per week speelt u over het algemeen gemiddeld videospellen?	(Bijna) Nog geen dag per week 1 Dag per week 2 Dagen per week 3 Dagen per week 4 Dagen per week 5 Dagen per week 6 Dagen per week (Bijna) Elke dag van de week
	Hoeveel dagen van de laatste vier weken, inclusief vandaag, heeft u videospellen gespeeld? Kies een getal van 1 t/m 28.	[Nummer van 1 t/m 28, ter representatie van het aantal dagen van die vier weken dat de participant videospellen heeft gespeeld]
Duur	Hoe lang duurt een gemiddelde sessie van het spelen van videospellen voor u zo ongeveer?	Nog geen 10 minuten 1 Uur 2 Uur 3 Uur 4 Uur 5 Uur 6 Uur 7 Uur 8 Uur of langer

Appendix 4. Starting on the following page, the format of the entire series of all the original RSQ-items and other used questionnaires, including consent form and demographics, will follow. First in English, then in Dutch. Following that, the format used for the test-retest, again first in English, then in Dutch.