Review

Why do people buy virtual goods: A meta-analysis

Juho Hamari a, b, c, *, Lauri Keronen c

a Gamification Group, Laboratory of Pervasive Computing, UC Pori, Tampere University of Technology, Finland
b Gamification Group, Digital Culture, UC Pori, University of Turku, Finland
c University of Tampere, Finland

A R T I C L E  I N F O

Article history:
Received 13 July 2016
Received in revised form 19 January 2017
Accepted 22 January 2017
Available online 24 January 2017

Keywords:
Virtual goods
Virtual worlds
Online games
Social networking services
Freemium
Free-to-play

A B S T R A C T

During the last decade, virtual goods have become an important target of consumption online (especially in games, virtual worlds and social networking services) amongst physical and digital goods. In this study we investigate the question of why do people purchase virtual goods by conducting a meta-analysis (random effects model) of the existing quantitative body of literature (24 studies) on the topic. The meta-analysis revealed an important aspect of value of virtual goods: contrary to traditional goods, the reasons why people purchase virtual goods are tightly connected to the platform where they are sold in. These findings underline the significance of service design and its relationship to the formation of value of virtual goods: the value of virtual goods is context-bound, and therefore, bound to the environment where they are usable in. Most factors that were found to be significant predictors of purchase behavior (such as network effects, self-presentation, enjoyment, ease of use, flow and use of the platform) are directly related to the aspects and design of the platform beyond the general attitudes towards virtual goods themselves. Moreover, we found that enjoyment and prolonged use of the platform were more important predictors for purchases in virtual worlds than in games.

© 2017 Elsevier Ltd. All rights reserved.
1. Introduction

Virtual goods have become one of the major categories of consumption in online environments among the purchases of normal physical goods (sold on e.g. Amazon) and digital goods such as music (e.g. iTunes). Virtual goods refer to virtual objects such as items, avatar clothing, weapons, virtual furniture, currencies, characters and tokens that commonly exist solely within variety of virtual environments (Fairfield, 2005; Hamari & Lehdonvirta, 2010; Lehdonvirta, 2009) where they are usable in. Whereas digital goods such as music and photos can be duplicated, virtual goods are rivalrous implying that they can’t be copied but are rather regulated by the rules of the given virtual economy (Fairfield, 2005; Hamari & Lehdonvirta, 2010; Harvainen & Hamari, 2015; Lehdonvirta, 2009).

Virtual goods are often categorized into appearance, social and functional based goods (e.g. Lehdonvirta, 2009). For example, appearance based goods only affect the appearance of virtual character or avatar whereas functional goods can contain game items such as more powerful weapons, armor or other boosts that increase various character attributes. Virtual goods are bound by the rules of the environment where they are used. Virtual goods exist within virtual economies, such as in games and virtual worlds (Lehdonvirta & Castronova, 2014).

Selling virtual goods has recently become a default business model for games publishers and many social networking services (Alha, Koskinen, Paavilainen, Hamari, & Kinnunen, 2014; Hamari & Jarvinen, 2011; Hamari & Lehdonvirta, 2010; Hamari, 2011; Hamari, Hanner, & Koivistoinen, 2017b; Lehdonvirta & Castronova, 2014; Lehdonvirta, 2009; Nieborg, 2015). Many online games and virtual worlds allow free access to main features and instead rely on generating revenues by selling virtual goods on top of the core service. Thus, understating why people purchase virtual goods is a pertinent practical issue for the service operators. Since most of these services have no entry pricing and virtual good purchases are voluntary, it is essential to understand which factors can increase virtual good purchases. In free-to-play games, only a minute percentage of registered users purchase virtual goods, (2% according to Forbes, 2013).

The topic started to attract academic attention circa 2005 and the first quantitative studies on the motivations to purchase appeared circa 2008. Thus far, studies on topic have been interested in predicting purchase behavior with, for example, experiences related to use of the core service (Animesh, Pinsonneault, Yang, & Oh, 2011; Määttänen & Salo, 2013), attractiveness of virtual goods themselves (Kim, 2012; Kim, Gupta, & Koh, 2011; Wang & Chang, 2014), and desire for self-representation in virtual world (Kim et al., 2011, Kim, Chan, & Kankahalli, 2012). Moreover, the topic has been examined from viewpoints of several theoretical perspectives such as technology acceptance (Cha, 2011; Domina, Lee, & MacGillivray, 2012; Hamari & Keronen, 2016), theories of planned behavior and reasoned action (Gao, 2014; Kaburuan, Chen, & Jeng, 2009), expectancy-disconfirmation model (Wang & Chang, 2013; Wang & Chang, 2014), as well as transaction cost theory (Guo & Barnes, 2011; Guo & Barnes, 2012). Although this body of literature covers a variety of factors affecting purchase behavior and intention, individual studies are often naturally limited to a singular service, virtual world or a game which restricts generalization of their findings on top of limited sample sizes and inevitable measurement errors. Therefore, the understanding is in need for a comprehensive meta-analysis and synthesis of previous research findings.

Although selling virtual goods has become powerful business strategy in virtual world and online games, this commodification of games has certainly faced heavy resistance from the users and developers (Alha et al., 2014; Hamari, 2015; Hamari et al., 2017a; Kimppa, Heimo, & Harvainen, 2016; Lin & Sun, 2011). Many free-to-play game publishers encourage users to purchase functional goods for faster progression and competitive advantage in the game. However, paying for competitive advantage has been regarded as highly incompatible with the nature of games and many players perceive purchasing advantageous goods as cheating. In fact, previous studies have discussed how such purchases can decrease engagement, immersion and flow in gameplay experience (Alha et al., 2014; Bartle, 2004; Hamari & Lehdonvirta, 2010; Lin & Sun, 2011; Paavilainen, Hamari, Stenros, & Kinnunen, 2013). In case competitive balance within a game is desired, understanding why people buy goods for other reasons is essentially important. However, such problems are not as prominent in non-game virtual worlds which are often employed for more instrumental purposes than games are. Therefore, investigating whether the reasons and their effect on purchases may differ between games and virtual worlds.

Therefore, this study aims to address the question of why do people buy virtual goods (in games and virtual worlds) via conducting a meta-analysis of existing quantitative literature on the topic. More specifically, this study provides estimates for most studied direct correlations with purchasing intention for virtual goods in games and virtual worlds investigated in the literature. Since this study aims for as comprehensive as possible literature review of quantitative research and meta-analysis, we are not restricted by theoretical assumptions stemming from any particular theoretical frameworks. We examine the correlations between variables regardless of whether the analyzed studies had modelled a relationship between them in their structural models, therefore, firstly, the study has not decided before data collection which independent variables are of interest as long as a correlation between them and the dependent variable related to virtual good purchases can be established in the reviewed research papers, and secondly, the study is also able to take into account relationships between variables not disclosed as results in prior literature. As the primary objective of this study is to rigorously synthesize and therefore validate and repeat the studies done on the question of why people wish to purchase virtual goods in differing environments, the emphasis of study is to increase the validity and reliability of the empirical results on this area. Therefore, by collating quantitative studies and laying down a reliable foundation for the area of virtual goods consumption motivations, this study act as a reference point for further studies that attempt to take this vein of research further.

2. Procedure & methods

This section reports the individual phases in the analysis (see Fig. 1) conducted by the authors of the present study. We began the analysis by conducting literature searches, first for defining...
comprehensive but accurate set of keywords with exploratory searches and then performing formal search of studies. We were required to form criteria for mechanical and unambiguous rejection of unsuitable search hits for analysis. Next, we coded the relevant statistics, findings and measures from the remaining set of studies. Then, we conducted descriptive analysis by inspecting sample sizes, virtual environments and theoretical foundations in the included studies. Finally, we validated our approach for random effects model in meta-analysis by testing heterogeneity in the studies and eventually conducted actual meta-analysis. All analyses were conducted in Microsoft Excel.

2.1. Search

Following guidelines of Webster and Watson (2002) and Ellis (2010), the analysis procedure commenced with a literature search. The search procedure was undertaken in the Scopus database (February 2015) which is the largest abstract and citation database of scholarly literature (Elsevier B.V., 2014). Scopus is also the most relevant repository for studies within the disciplines where literature on why people adopt and use different technologies is being published. Among many others, Scopus also includes the AIS, ACM, IEEE and Science Direct libraries.

We began the search process by conducting a set of exploratory searches of articles for discovering and identifying the common terminology in order to determine comprehensive but accurately describing set of keywords for formal search. This was first searching with rather wide terms while ordering the results by relevance and picking up some central keywords from clearly relevant studies and then making further searches with these revealed keywords. We also inspected studies that were referring already gathered relevant studies. Repeating this process iteratively while keeping accurate terms and discarding too wide keywords eventually resulted in acceptable search string. This basically consists of two parts: 1) action of making purchases and 2) context of virtual items and typical type of games or virtual worlds that allow such purchases. This search string was targeted for meta-data (titles, abstracts and keywords) of the articles rather than entire text such purchases. This search string was targeted for meta-data (titles, abstracts and keywords) of the articles rather than entire text. This basically consists of two parts: 1) action of making purchases and 2) context of virtual items and typical type of games or virtual worlds that allow such purchases. This search string was targeted for meta-data (titles, abstracts and keywords) of the articles rather than entire text such purchases. This search string was targeted for meta-data (titles, abstracts and keywords) of the articles rather than entire text.

2.2. Inclusion criteria

We determined 5 consecutive criteria for systematic and straightforward inclusion of studies for the analysis. The whole inclusion process resulted in 20 valid published papers (including 24 studies) for further analysis by discarding 94 irrelevant or unsuitable articles from a total of 116 search hits. The search and inclusion was initially conducted by one researcher (Master-level). Both the included and omitted articles were saved and another researcher (then Postdoc/Assistant professor -level) checked the omitted articles to make sure no suitable articles were omitted. Exclusion disagreements were discussed among researchers which lead into a unanimously selected pool of papers. All aspects of papers were used in determination of fit (e.g. abstract, title, full content).

First, the studies were inspected as to whether they were duplicates. Eight research articles were omitted for sharing same results or data with more recent and extended versions of these papers that were already included.

Second, 11 search hits were omitted since they were not full papers that had been published either in peer-reviewed journals, conferences or books.

Third and the largest omission category with 59 rejections was research articles that were on a different topic than the purchasing behavior of virtual items. In this category, most frequent topic was concerning interactive virtual product experience (e.g. inspecting virtual model of real item or wearing it on an avatar) and its effects on purchasing behavior of non-virtual items. This category also contained studies focusing on behavioral outcomes of augmented reality in purchasing decisions, effect of brand advertising in virtual worlds on real product purchase intentions and development reports of different virtual product systems in addition to less frequent various topics.

As a fourth step, the remaining entries were inspected for whether they included a quantitative empirical study. On this basis, 6 entries were omitted since the meta-analysis required actual measures from empirical studies with a similar research problem.

Finally, and most pertinent to the meta-analysis calculations, the studies had to report correlations between their variables. Otherwise, no restrictions on independent variables were set. A total of 12 studies were omitted for not reporting a correlation matrix; thus the data had no missing data in the final data set. Beyond these steps, further quality criteria were not needed.

2.3. Coding

Before the actual analysis, the data is required to be in comparable format. When analyzed studies reported correlation matrices, it did not require any extensive modifications in this study. However, four studies instead reported squared correlations which required square rooting in our data. While this was straightforward process, squared correlations should be interpreted with caution since possibly negative correlations gain positive sign when squared. Therefore, we contacted the authors of these papers in order to guarantee that the correlations had been positive.

We identified three different types of variables related to measuring purchasing behavior of virtual items: intention to purchase, actual purchasing behavior and loyalty. Clear majority of the studies were interested in predicting general purchasing intention which was simply coded as “purchase intention”. Unfortunately, the research concerning loyalty and actual purchase behavior was so rare and scattered that we could not analyze these variables comprehensively in our meta-analysis. Therefore, this review was...
forced to focus on purchasing intention.

In addition, we inspected all independent variables of all studies on whether they measured what the variable name indicated. As an example, many studies measured “enjoyment”, however, some studies measured the enjoyment of using the core service while others were interested in the enjoyment of using virtual goods and even some investigated the enjoyment of shopping activity itself. We combined or separated these cases accordingly.

The literature search revealed three studies that contained several subgroup analyses. Fortunately, these studies collected their subgroup questionnaire data separately from independent subject groups instead of splitting data afterwards. Therefore, all subgroup correlations could be treated as independent findings in this meta-analysis.

2.4. Meta-analytic approach

Reviewing published research can be divided into two overall approaches: 1) traditional qualitative method (also known as the narrative method) in which the conclusions of reviewed studies are practically summarized using words, and 2) meta-analysis which is a mathematical and quantitative approach, and where the effect sizes of the reviewed studies are combined using calculations (Ellis, 2010).

The narrative approach has been found to be insufficient when synthesizing findings from contradictory results, especially for a large number of studies (Hunter & Schmidt, 2004), whereas the meta-analytic approach provides more comprehensive results with estimates for effect size, different metrics for reliability, and information about different kinds of bias. Moreover, unlike the narrative approach, meta-analysis does not suffer from increased complexity in interpreting large amounts of studies. Instead, meta-analysis addresses the discrete limitations of individual studies and settles conflicting findings (Paré, Trudel, Jaana, & Kitsiou, 2015). As the limitations of traditional narrative review are acknowledged, it is reasonable to employ a meta-analysis in this study.

2.4.1. Meta-analytic calculation model

More specifically, meta-analysis is a mathematical and statistical method for combining the results of previous studies that address a similar research problem (or the data/results which can be used to address a similar research problem) (Glass, 1981). There are two main approaches for mathematical meta-analysis (Ellis, 2010; Hunter & Schmidt, 2004): one developed by Hunter and Schmidt (Hunter & Schmidt, 2000; Schmidt & Hunter, 1977) and the other by Hedges et al. (Hedges & Olkin, 1985; Hedges & Vevea, 1998; Hedges, 1981, 1992).

In the approach of Hedges et al., raw correlations are z-transformed before combining the effects, and weights of n − 3 are used instead of the original sample size (n) for each study. In contrast, the method by Hunter and Schmidt uses untransformed correlations, and the original sample size of each study. However, an analysis using this approach should modify the weights to be taken into account and correct the study-specific faults such as measurement reliability. The calculation of Hedges et al.’s random effect model uses the between-studies variance (Ellis, 2010). These two approaches will likely produce slightly different mean effect sizes and intervals, but it is difficult to say which one is better overall as the differences are minor (Ellis, 2010).

For example, Field (2005) ended up with results contradictory to a similar study of Hall and Brannick (2002), even though both employed the two methods in similar conditions using Monte Carlo simulations. However, Johnson, Mullen, and Salas (1985) compared extensively different meta-analytic approaches and concluded that the Hunter & Schmidt method produces differing results and should be used with caution. Although Schmidt and Hunter (1999) later argued that this difference was caused by use of an inappropriate formula for error correction, we were more confident with the method of Hedges et al. and it was chosen as an approach for the purposes of this analysis.

Both meta-analytic calculation approaches include at least two different models, namely to account for fixed and random effects. In principle, a fixed effects model should be used when the studies share identical data collection conditions and a single value for the true effect is assumed. Thus, using a fixed effect generally produces less variance as well as tighter confidence intervals. On the other hand, a random effects model should be used when the study conditions are expected to vary, and the distribution for the true effect is assumed. Indeed, in most real-life scenarios and meta-analyses, it would be absurd to assume that identical study conditions exist between studies. Moreover, as our data clearly suggests dissimilar conditions with varying variable details as well as different cultures and demographics amongst the respondents, it is reasonable to employ a random effects meta-analysis. Therefore, we used the approach of Hedges et al., and a random effects model in our meta-analysis. The calculation formulas used in our meta-analysis are available in Appendix B.

2.4.2. Test of heterogeneity

Despite the assumption that a random effect-basis is preferred to a fixed effects model when combining the effect sizes of independent studies, we verified our model approach using tests for heterogeneity. The heterogeneity of our data was tested with Q-statistics and I²-values for every relationship that was analyzed in meta-analysis (these heterogeneity tests are available in Appendix C). The Q-statistic (Cochran, 1954) is the classical measure for heterogeneity while the I²-value represents the percent of the variance explained by the heterogeneity of the data, and the minimum of 0% indicates that all variability is instead due to sampling error within trials (Higgins & Thompson, 2002). All Q-estimates were statistically significant at p < 0.01 and each I²-value was above 80% (mostly above 90%). Thus, the random effect model is seen as a proper approach for conducting this particular meta-analysis.

2.4.3. Publication bias

To address the problem of publication bias, fail-safe N was calculated for each of the analyzed relationships. The fundamental concept is to determine the number of additional studies with zero result needed to nullify an effect. There are two main approaches for such calculations (Long, 2001): one method based on the sum of the Z scores (Rosenthal, 1979) while the other uses effect sizes (Orwin, 1983). The latter was used in this analysis since it provides more accurate results without the need for an interpretation of statistical significance testing (Long, 2001). Additionally, the method collaborates well with the classes for effect sizes featured in this study. We used the small-threshold as criterion value for fail-safe studies, and zero for the mean effect size of the fail-safe studies (Eq. (1)).

\[
\text{Fail-safe } N = \frac{k(r - 0.1)}{0.1}
\]

where k is the number of studies in the analysis, r is the mean effect size, and value 0.1 (the small effect size threshold) is the criterion value for fail-safe studies. A higher fail-safe N value implies a more reliable result in the aspect of publication bias. For interpretation of the value, the fail-safe N/k—ratio should exceed the threshold of 2.0, otherwise publication bias might pose a potential problem (Sabherwal, Jeyaraj, & Chow, 2006).
2.4.4. Moderator analysis

The purpose of our moderator analysis is to examine the difference in meta-analysis results between two different types of games. The difference between the two correlation estimates is examined by way of Q-test, which tests the homogeneity and significance of variance between groups (Borenstein, Hedges, Higgins, & Rothstein, 2009). Similar to actual meta-analysis, the test also requires some decisions regarding the calculation model to be used. First, one must choose between a fixed or random effect model, depending on how the within group estimates are to be calculated. Similar to the main meta-analysis, we had no reason to believe that even studies within the same game categories would have such identical research conditions, that a fixed effect could be assumed. Therefore, the subgroup estimates are calculated using a random effects model. As a second issue, one must decide whether to assume true between-studies variance for both subgroups or to estimate separate variances. However, a relatively low number of studies within subgroups does not allow for the separate variances for each group to be estimated with any reasonable accuracy. On the other hand, we had no reason to assume different variances for these groups, so the same within studies estimate for variance is used for both subgroups.

3. Results

3.1. Details of the reviewed studies

The inclusion process resulted in 20 research papers for further analysis and they are represented in Table 1. When counting also different sub-studies, the total number of studies is 24. The studies have been published between years 2008 and 2015. Majority of the studies are journal articles and the data contains only single conference paper. Sample sizes range from 38 to 2481 with a mean of 529 and standard deviation of 612.

The frequencies of different environment types are shown in Table 2. Twelve studies have been conducted in the context of virtual worlds and 7 studies in the context of games (SNGs: 3, MMOs: 3, and Mobile games: 1). A single did not specify the context.

Most of the studies did report their experiment concerning more than a single service or did not report actual titles of services (7). Among the reported titles, most of the studies did use Habbo Hotel (5), Second Life (3) or Cyworld (3) virtual worlds. Most frequent game title was World of Warcraft but with frequency of only 2. Rest of the reported titles were used in single studies (Table 3).

We also investigated the distribution of theoretical frameworks utilized in the body of literature. As seen in Table 4, our review reveals that majority of the studies did not specify any clear theoretical foundation or used variety of variables from different frameworks or studies. Moreover, use of specific theories and models is rather scattered and only low frequencies are detected. Nevertheless, some studies utilized technology acceptance model (Davis, Bagozzi, & Warshaw, 1989), theory of planned behavior (Ajzen, 1991) and unified theory of acceptance and use of technology (Venkatesh, Morris, Davis, & Davis, 2003).

3.2. Variables

In total, the collected data contains large number of different variables and 398 unique correlation pairs. As we were interested in factors that explain virtual goods purchases, we meta-analyzed the correlations between the purchase-related variable and any variables that were featured in at least 3 individual studies. Table 5 introduces the most frequent variables in the reviewed research literature, number of studies examining them (k) as well as a brief description for each variable. These variables are also featured in meta-analysis. Therefore, beyond this meta-analysis, there exists a long-tail of variables that have been investigated in individual studies. Naturally, such a long list is outside the scope of this meta-analysis.

3.3. Meta-analysis

3.3.1. Main findings

The results in Table 6 (also visualized in Fig. 2) show most frequently studied variables in the literature and our meta-analytically produced estimates for their correlation with Purchase Intention which were in order of magnitude: Attitude (0.662), Flow (0.482), Perceived Network Size (0.480), Self-Presentation (0.478), Subjective Norms (0.466), Social Presence (0.438), Perceived Value (0.418), Service Use Enjoyment (0.370), Service Use Intention (0.359), and Perceived Ease of Use (0.333). Every estimate in the analysis was clearly positive and statistically significant at p < 0.001 with adequate failsafe N values (See Table 6).

3.3.2. Moderating effect of service type

While games and virtual worlds offer purchasable virtual goods, they are relatively different types of environments. Whereas games are commonly competitive, rule-driven, fast-paced goal-oriented and narrative rich, virtual worlds are commonly free-form and have no clearly defined goals or game-like competition. In games, purchasing virtual goods can give unfair competitive advantage as they can make the game character stronger (Alha et al., 2014; Hamari & Lehdonvirta, 2010; Hamari, 2015; Lehdonvirta, 2009). Therefore, the motivations for purchasing virtual goods in these environments may differ. To address this assumption, we expanded the meta-analysis by investigating the differences between effect between the game and virtual world environments.

Since the number of studies become lowered due to the grouping, we reduced the required k of studies to two for each category. As a result, the comparison analysis compares five relationships between the virtual environment categories (Table 7 and Fig. 3). An asterisk denotes a p-value of lower than 0.05 and in most cases the p-value was lower than 0.01. The results showed a large difference for correlation between Service Use Intention and Purchase Intention (\( Q^* = 22.492 \)), where games had considerably lower correlation (0.211) compared to mediocre estimate of virtual worlds (0.465). Quite similarly, there was a large difference in correlation between Service Use Enjoyment and Purchase Intention (\( Q^* = 22.492 \)), where again the relationship for games (0.185) was significantly weaker than the estimate for virtual worlds (0.461). Moreover, there was slight difference between correlations for Flow and Purchase Intention (\( Q^* = 5.920 \)), where games had a lower estimate (0.437) compared to virtual worlds (0.557). However, the analysis could not detect significant difference for correlation between Subjective Norms and Purchase Intention (\( Q^* = 0.052 \)) since both categories had similar estimates (games: 0.453, worlds: 0.494). In addition, there was no noticeable difference in relationship between Attitude and Purchase Intention (\( Q^* = 0.027 \)) as both categories showed similarly high correlations (games: 0.666, worlds: 0.654).

Despite the fact that number of studies is lowered to two studies at minimum, all group estimates are statistically significant and positive. However, in relationship between Subjective Norms and Purchase Intention, correlation estimates for both service categories had wide confidence intervals (see Fig. 3) due to high variation in previous research findings. Games-group had 95% confidence interval of 0.435 whereas virtual world-category had
Although the literature showed rather varying findings on the strength of the relationship between the variables, the correlation estimates in this analysis were clearly positive in both categories. Other studies within their categories had rather unanimous results which is shown in relatively narrow confidence intervals (0.245 at most).

4. Discussion

This study investigated the question of why do people purchase virtual goods by conducting a meta-analysis of the existing quantitative body of literature. The results revealed that across the literature the following factors were positively associated with virtual good purchases as well as the most investigated factors in the body of literature investigating predictors of virtual goods purchasing: attitude, flow, network size, subjective norms, service use enjoyment, service use intention and perceived ease of use. Attitude towards virtual goods had clearly the strongest association with purchase intention. Moreover, the results showed differences in the magnitude of some purchase motivators between games and virtual worlds: in virtual worlds, service use intention and enjoyment were significantly stronger predictors for virtual good purchases than in games.

In contrast to consumer research in general, as can be seen from the set of variables examined in the literature, the research on virtual goods consumption has rather heavily focused on aspects...
related to the platform on which the virtual goods are being used (such as the enjoyment derived of the platform use or the use intentions towards the platform as well as the network size), whereas literature on consumption of goods in general is commonly focused on the aspects of the products themselves. The findings of this study and the focus in the literature strongly indicate that virtual goods inhabit a highly curious environment: virtual goods are bound by the rules of the service in which they are used, developers control the supply and value of virtual goods by controlling how new virtual goods can be spawned into existence, how they can be traded, who can own them at any particular time, their price, their rate of degradation and whether they can be traded back to ‘real money’. Ultimately, any of this does not matter unless the developers have also created an appealing and enjoyable enough environment to which users are attached to. Without users using the platform, the virtual goods remain in a limbo of virtual meaning. Customers do not choose the games they start playing based on what purchasable goods the game might have, and therefore, it creating an appealing platform that hosts the virtual goods remain an important prerequisite. Instead, users are arguably more likely to choose the core services based on their entertainment value or interestingness. Thus, in order to enable purchases, potential customers must first use the core service and enjoy it as such. Indeed, it is also the result of this study that consumers consider alternatives.

Fig. 2. Meta-analysis of correlations with purchase intention and their 95% confidence intervals.

Table 5
Most frequent variables and their brief descriptions.

<table>
<thead>
<tr>
<th>Variable</th>
<th>k</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase Intention</td>
<td>24</td>
<td>Intention to purchase virtual goods (Ajzen &amp; Fishbein, 1980).</td>
</tr>
<tr>
<td>Service Use Enjoyment</td>
<td>8</td>
<td>Extent of how enjoyable and fun using the game or virtual world it is.</td>
</tr>
<tr>
<td>Subjective Norms</td>
<td>8</td>
<td>Perceived social pressure from other people on how acceptable playing games or using virtual worlds is (Ajzen &amp; Fishbein, 1980). Also often referred as “social norms” or “social influence”.</td>
</tr>
<tr>
<td>Flow</td>
<td>6</td>
<td>Flow is a mental state where a person is fully immersed, deeply concentrated and truly enjoys when performing a certain activity (Csíkszentmihályi, 1990). Flow is the optimal hedonic experience in playing games or using virtual worlds.</td>
</tr>
<tr>
<td>Attitude Toward Purchase</td>
<td>5</td>
<td>Attitude is own opinion on how positive or negative purchasing virtual goods is (Ajzen &amp; Fishbein, 1980).</td>
</tr>
<tr>
<td>Purchase Service Use Intention</td>
<td>4</td>
<td>Intention to play games or use virtual worlds (Ajzen &amp; Fishbein, 1980).</td>
</tr>
<tr>
<td>Perceived Ease of Use</td>
<td>3</td>
<td>The degree to which an individual believes that using a particular system would be free of physical and mental effort” (Davis, 1989). Especially in context of games, ease of use denotes effortless in user interface rather than difficulty level.</td>
</tr>
<tr>
<td>Perceived Network Size</td>
<td>3</td>
<td>Perception on amount of friends, peers and people around are also using the service.</td>
</tr>
<tr>
<td>Perceived Value</td>
<td>3</td>
<td>Perceived ratio between value and price or virtual goods in which valuable but cheap goods become desirable whereas expensive items make users consider alternatives.</td>
</tr>
<tr>
<td>Self-Presentation</td>
<td>3</td>
<td>Desire for expressing oneself in virtual world by character customization such as wearing stylish clothing and accessories on avatar.</td>
</tr>
<tr>
<td>Social Presence</td>
<td>3</td>
<td>Sense of real human contact and sociability in virtual world.</td>
</tr>
</tbody>
</table>

k – number of studies examining the variable.

Table 6
Results of the meta-analysis. k – number of studies, ∑ n – cumulative sample size, r – correlation coefficient, lower and higher bounds of 95% confidence interval, Z – z-score for correlation estimate, p – statistical significance of estimate, fs N – failsafe N.

<table>
<thead>
<tr>
<th>Variables</th>
<th>k</th>
<th>∑ n</th>
<th>r</th>
<th>95% Conf. Int.</th>
<th>Z</th>
<th>p</th>
<th>fs N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Use Enjoyment × Purchase Intention</td>
<td>8</td>
<td>8045</td>
<td>0.370</td>
<td>0.275 – 0.459</td>
<td>7.144</td>
<td>0.000</td>
<td>22</td>
</tr>
<tr>
<td>Subjective Norms × Purchase Intention</td>
<td>8</td>
<td>3868</td>
<td>0.466</td>
<td>0.294 – 0.608</td>
<td>4.915</td>
<td>0.000</td>
<td>30</td>
</tr>
<tr>
<td>Flow × Purchase Intention</td>
<td>6</td>
<td>2043</td>
<td>0.482</td>
<td>0.399 – 0.557</td>
<td>10.015</td>
<td>0.000</td>
<td>23</td>
</tr>
<tr>
<td>Attitude Toward Purchasing × Purchase Intention</td>
<td>5</td>
<td>3102</td>
<td>0.662</td>
<td>0.507 – 0.719</td>
<td>14.405</td>
<td>0.000</td>
<td>29</td>
</tr>
<tr>
<td>Service Use Intention × Purchase Intention</td>
<td>4</td>
<td>5272</td>
<td>0.359</td>
<td>0.247 – 0.461</td>
<td>5.966</td>
<td>0.000</td>
<td>11</td>
</tr>
<tr>
<td>Perceived Ease of Use × Purchase Intention</td>
<td>3</td>
<td>3837</td>
<td>0.333</td>
<td>0.216 – 0.440</td>
<td>5.360</td>
<td>0.000</td>
<td>7</td>
</tr>
<tr>
<td>Perceived Network Size × Purchase Intention</td>
<td>3</td>
<td>4751</td>
<td>0.480</td>
<td>0.401 – 0.551</td>
<td>10.469</td>
<td>0.000</td>
<td>12</td>
</tr>
<tr>
<td>Perceived Value × Purchase Intention</td>
<td>3</td>
<td>759</td>
<td>0.418</td>
<td>0.331 – 0.497</td>
<td>8.637</td>
<td>0.000</td>
<td>10</td>
</tr>
<tr>
<td>Social Presence × Purchase Intention</td>
<td>3</td>
<td>639</td>
<td>0.478</td>
<td>0.329 – 0.626</td>
<td>4.743</td>
<td>0.000</td>
<td>12</td>
</tr>
</tbody>
</table>

Fig. 2. Meta-analysis of correlations with purchase intention and their 95% confidence intervals.
on purchase behavior in games (small effect) that in virtual worlds (medium effect). At first, this finding might seem unintuitive since games, after all, are commonly strongly associated with the pursuit of enjoyment. However, recent related literature may shed light on possible explanations for this results. Prior literature examining the association between purchase behavior of virtual goods in games and the game experience (Hamari, 2015; Lin & Sun, 2011) has observed and discussed that the enjoyment (and related factors) may have a more complex, dual-directional effect on purchases in games. In order to create demand for the virtual goods in games, many game developers may intentionally seek to frustrate players by creating artificial obstacles and hindrances, and therefore, generate sales through negative enjoyment or by facilitating social stratification, and thus, creating incentives for users to stand out (Hamari & Lehdonvirta, 2010; Hamari, 2015; Hamari et al., 2017a; Lehdonvirta, 2009; Lin & Sun, 2011; Mäntymäki & Salo, 2015).

Therefore, on one hand, developers are required to make the game enjoyable enough for the players to come and stay in the game, but on the other hand, it may be in the developer’s best interest to create gaps into gameplay in order to sell them more virtual goods that address those frustrations. Indeed, our results may lend support for these prior findings; the low effect size between enjoyment and purchase intentions in games may suggest that there is a double-sided effect: on one hand, enjoyment by default increase willingness to purchase virtual goods (especially through increased playing intentions) but on the other hand, virtual goods are purchased if the game is not enjoyable enough. Relatedly, the strength of association between playing and purchase intentions also varied between games and virtual worlds in the same manner as the relationship between enjoyment and playing intentions: in virtual worlds the association was of medium strength whereas in games it the association was only of small strength.

The results show a strong positive relationship between attitude and virtual good purchases independent of service type. This relationship might be more crucial in games, which can be regarded as the more competitive setting than virtual worlds. Many free-to-play game developers utilize the competitive nature of games and offer virtual goods that have functional advantages such as more powerful weapons, better armors or items to temporary improve character statistics. Although the business strategy is currently widely used, many players are uncomfortable with paying for unfair performance boost (Lin & Sun, 2007; Lin & Sun, 2011). It remains an interesting further line of inquiry to investigate what factors may alleviate or strengthen these strong attitudes towards virtual good purchases.

Subjective norm is central variable in theory of planned behavior (Ajzen, 1991). According to this theory, human behavior is affected by social pressure on whether such behavior is accepted or not by others. Beneficial to the theory, our meta-analysis results confirm that purchase intention has significant correlation with subjective norms, indicating that when purchasing virtual goods is accepted by others, people themselves are more likely to make purchases. While the results of this study certainly confirm an effect, they also show a large variation between the studies. However this high variance could be explained with individual differences in accepting social influence. Some people might perceive opinions of other more important whereas some might perform actions without particularly considering what others think of them. Alternatively, some platforms may lack features that would facilitate social interaction, and therefore, also the level of social influence may be diminished. Despite the possible high variation in individuals, opinions of others definitely have an effect on own purchase motivation.

Flow has been characterized as an optimal hedonic experience in use of service in which the user is deeply concentrated and fully immersed in the activity (Csikszentmihalyi, 1990). As flow represents another hedonic experience in use of service, it is arguably related to enjoyment. However, the key difference between the variables is that flow can be regarded as more specific experience whereas enjoyment represents more general enjoyment in service use. Thus, we expected finding stronger association between flow and purchase intention than enjoyment had. Our meta-analysis revealed significant and relatively strong correlation between flow and purchase intention. Moreover, our results showed noticeably stronger association between flow and purchase intention than enjoyment had.

In technology acceptance model (Davis, 1989), perceived ease of use predicts the use intention of information systems. However, the
present analysis examined its relationship with purchase intention. Although direct effect of ease of use was not modelled by any of the analyzed studies, we were able to estimate its correlation with purchase intention. Our meta-analysis results showed significant relationship between the variables, indicating that higher core service usability can increase purchase intentions. However, the correlation was only mediocre in its strength and the weakest relationship in this meta-analysis. On the other hand, we believe that ease of use would have more significant role by indirectly affecting purchase intention. Thus, future meta-analyses could examine the relationship between ease of use and other variables. Moreover, they could employ structural equation modeling approach and examine the indirect relationship of ease of use and purchase intention through other variables.

4.1. Limitations and future directions

There are some limitations in this review that should be acknowledged. First, although meta-analysis can be technically conducted when a relationship has been examined by at least two studies, most analyses typically require more findings. Due to relatively low number of previous quantitative research and especially because of scattered nature of research literature, this analysis occasionally contained rather low number of studies for some of the correlation pairs. While we managed to achieve significant relationships and mostly narrow confidence intervals for our estimates, some correlation pairs would preferably need more independent findings for wider generalization. Moreover, especially in the moderator analysis the statistical power is especially low. Therefore, naturally the results should be interpreted with healthy caution. In any case, however, the current meta-analysis offers reliability over any singular studies and affords an overview to the quantitative body of literature on why people purchase virtual goods.

The quantitative literature (and therefore this meta-analysis — Table 1) on why people purchase virtual goods afford only a limited view to the entire vein of literature on the topic. Especially recent qualitative studies have attempted to more holistically identify factors that affect virtual good purchases (Hamari & Järvinen, 2011; Hamari & Lehdonvirta, 2010; Hamari et al., 2017a; Mäntymäki & Salo, 2015). If we investigate the factors identified in the qualitative literature, we can see a partial overlap between qualitative body of literature and aspects included in this present study. However, whereas the quantitative body of literature seems to have focused on more abstract psychological constructs, the qualitative body of literature has significantly more focused on the interplay of game/virtual world design and purchase practices; how the game design creates demand for virtual goods purchases (Hamari & Lehdonvirta, 2010; Hamari, 2011; Lehdonvirta, 2009; Lin & Sun, 2011; Mäntymäki & Salo, 2015; Oh & Ryu, 2007) which can be considered one of the most alluring aspects of virtual consumption when compared to consumption of material or digital goods (Fairfield, 2005; Hamari & Lehdonvirta, 2010). Such aspects include e.g. the need to purchase virtual goods in order to continue the game, to win over other players, to trade, to make friends and to personalize (See e.g. Hamari et al., 2017a; Mäntymäki & Salo, 2015). Therefore, future should investigate these practices more closely beyond just the very abstract psychometric constructs.

Another fruitful avenue to investigate why people purchase virtual goods would be in further investigating the game production process and the negotiating process for the boundary between crafting piece of art versus a finely-tuned profit-oriented platform (as many games nowadays often are). Studies such as Alha et al., 2014; Alha, Koskinen, Paavilainen, & Hamari, 2016; Alves & Roque, 2007; Hamari, 2011; Hamari & Lehdonvirta, 2010; Hamari & Järvinen, 2011; Jankowski, Brödka, & Hamari, 2016; Kimppa et al., 2016; Lin & Sun, 2011; Mäntymäki & Salo, 2015; Nieborg, 2015; Prax, 2013; Zagal, Björk, & Lewis, 2013 have paved the way on this area, however, work remains to be done on investigating how demand is created for virtual goods, what kinds of game design can induce purchases as well as how too aggressive game design for profit can turn players away.

Moreover, not all virtual goods are the same. On an abstract level they can be divided into at least two categories by their characteristics. “Functional” goods are items that improve performance or progression in the service such as more powerful weapons, unlock new content and temporary improvements to virtual character. On the other hand, “appearance”-based goods do not grant any functional benefits and instead only alter the look of the virtual character. Since there is noticeable difference between these two types of goods, the motivations for purchasing can also be expected to vary. However, only a few studies clearly reported the type of virtual goods, preventing us from analyzing differences between the attractiveness of functional and appearance based virtual goods. Future efforts could examine the motivational differences between purchasing the two types of virtual goods.

Beyond the heterogeneity of virtual goods, also the consumers who buy them can differ. So called “player types” have been one of the most notable research avenues in game research in general (See e.g. Bartle, 1996; Hamari & Keronen, 2017; Hamari & Tuunanen, 2014; Kalio, Mayrā, & Kaipainen, 2010; Yee, 2006) as well as there has been an increasing debate over gamer demographics (See e.g. Koivisto & Hamari, 2014; Mayrā et al., 2016; Williams, Yee, & Caplan, 2008). The present review made clear, however, that demographic factors or ones related to gaming orientation of consumer have not been in central focus on research investigating why people purchase virtual goods. Therefore, naturally, we recommend further literature to more strongly focus on aspects related to the consumer in virtual consumption. For example, it might be interesting to investigate, whether achievement-oriented players are more like to purchase functional items that make winning easier, or whether socially-oriented players are more likely to purchase items that boost self-presentation as well as whether immersion-oriented players seek to deepen their experience within the game by focusing on extra content and roleplay. Moreover, it would be interesting to investigate how the level of disposable income and education play are role in susceptibility to virtual consumption.

Since we aimed for as comprehensive as possible literature review of quantitative research and meta-analysis, we were not restricted by theoretical assumptions stemming from theoretical frameworks such as Technology Acceptance Model (Davis, 1989) or Theory of Planned Behavior (Ajzen, 1991). We examined the correlations between variables regardless of whether the analyzed studies had modelled a relationship between them in their structural models. Therefore, our study not only presents reliable results on the topic why people purchase virtual goods but is also able to take into account relationships between variables not disclosed as results in prior literature. Although using zero-order correlations provides the most accurate and unbiased estimates as in this review, it is not uncommon for meta-analyses to instead use standardized regression path coefficients as their metric. Future meta-analyses could respectively synthesize previous path coefficients, however, it is important to note that model structure significantly affects these estimates. Thus, such meta-analyses should not interpret relationship strengths too strictly and perhaps instead examine proportion of significant effects. In either case, using correlations provides the most reliable estimates for relationships between variables.
Finally, we encourage structural equation modeling studies to report correlation matrixes not only to allow their data to be included in future correlation meta-analyses but also to address convergent and discriminant validity. Although majority of the reviewed studies followed this practice, unfortunately many did address only other of the validities or showed no results for neither of these potential issues. We also would like to emphasize scholars not to hide measures such as non-significant path coefficients and total variance in dependent variable explained by independent variables \( R^2 \) in the research model. Unreported non-significant path coefficients are especially unfortunate from the view point of meta-analyses.

### Appendix

#### A. Complete literature search string

**TITLE-ABS-KEY** (purchase OR purchasing OR repurchase OR shopping OR buying OR “behavioral outcomes” AND virtual items” OR “virtual goods” OR “virtual products” OR “digital items” OR game items” OR “using virtual currency” OR “facebook games” OR “in game purchase” OR “virtual world application” OR “social network games” OR “free-to-play” OR “online games stores” OR “in virtual world”).

#### B. Meta-analysis calculation formulas (Borenstein et al., 2009)

1. Use \( k \) for number of studies and \( n \) for sample sizes \( n - 3 \).
2. Fisher z-transform correlations before calculation:

\[
z = \frac{1}{2} \ln \left( \frac{1 + r}{1 - r} \right).
\]

3. Estimate between-studies variance:

\[T^2 = Q \frac{df}{C},\]

where

\[Q = \sum_{i=1}^{k} n_i z_i^2 - \frac{\left( \sum_{i=1}^{k} n_i z_i \right)^2}{\sum_{i=1}^{k} n_i},\]

\[df = k - 1.\]

\[C = \sum_{i=1}^{k} n_i - \frac{\sum_{i=1}^{k} n_i^2}{\sum_{i=1}^{k} n_i} \]

4. Random effect model weight:

\[w = \frac{1}{1 + r}.\]

5. Magnitude of effect size estimate:

\[z = \frac{\sum_{i=1}^{k} w_i z_i}{\sum_{i=1}^{k} w_i}\]

6. Standard error of effect size estimate:

\[SE = \sqrt{\frac{1}{\sum_{i=1}^{k} n_i}}.\]

7. 95% confidence intervals of effect size estimate:

\[ bounds = zq \pm 1.96 \times SE.\]

8. Statistical significance:

\[Z = \frac{zq - 1}{\frac{SE}{\sqrt{2}}},\]

9. Inverse Fisher z-transform back to correlation:

\[r = \frac{\exp(2z) - 1}{\exp(2z) + 1}.\]

### References


