The Motivational Potential of Digital Games and Gamification – The Relation between Game Elements, Experience and Behavior Change

Inaugural Dissertation
submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy to the Department of Psychology,
of the University of Basel
by
Elisa D. Mekler
from Zürich (ZH), Switzerland

Basel, 2015

THE MOTIVATIONAL POTENTIAL OF DIGITAL GAMES AND GAMIFICATION

Approved by the Department of Psychology

At the request of

Prof. Dr. Klaus Opwis (First Reviewer)
Prof. Dr. Lennart E. Nacke (Second Reviewer)

Basel, Switzerland, ___ _____ _

Prof. Dr. Roselind Lieb (Dean)

The Motivational Potential of Digital Games and Gamification – The Relation between Game Elements, Experience and Behavior Change

Elisa D. Mekler
Department of Psychology
University of Basel
Missionsstrasse 62a
4055 Basel
Switzerland
elisa.mekler@unibas.ch

Doctoral Thesis, submitted to the
Department of Psychology, University of Basel, Switzerland

Doctoral commitee:

Prof. Dr. Klaus Opwis (First Reviewer)
Prof. Dr. Lennart E. Nacke (Second Reviewer)
Prof. Dr. Rui Mata (Chairman)

Contents

Statement of Authorship

1. I, Elisa D. Mekler, hereby declare that I have written the submitted doctoral thesis "The Motivational Potential of Digital Games and Gamification – The Relation between Game Elements, Experience and Behavior Change" without any assistance from third parties not indicated.

2. I only used the resources indicated.

3. I marked all citations.

4. My cumulative dissertation is based on four manuscripts. The first, second and fourth manuscript have already been published. Manuscript 3 is in press and bound to be published in 2016. I certify here that the articles in this dissertation concern original work. Manuscripts 1, 2 and 3 presented in this doctoral thesis are largely based on my own authorship which includes concept, project management, data analysis, literature review, article writing, manuscript preparation and submission. The second authors, Julia Ayumi Bopp (manuscript 1) and Florian Brühlmann (manuscripts 2 and 3) also significantly contributed to the literature review, data analysis, article writing and manuscript preparation. Moreover, Florian Brühlmann programmed the image annotation platform described in manuscripts 2 and 3. Manuscript 4 was spearheaded by Sharon T. Steinemann, while being a master student under my supervision. I contributed to the concept, article writing and manuscript preparation, and provided recommendations throughout the project.

Abstract

Digital games have become the most popular medium of our time, having even surpassed the movie industry in terms of sales. Arguably, digital games offer some of the most enjoyable, rich and engaging experiences of all interactive products, with many players investing countless hours in gaming. Both researchers and industry professionals have taken notice of this trend and have attempted to foster games' motivational appeal to make other, 'purposeful' activities more engaging. This notion is reflected in concepts, such as gamification or serious games, which typically aim to afford gameful, that is ,'game-like' experiences. However, it is not clear what characterizes positive gaming experiences or what game aspects afford such experiences in non-game contexts. The four manuscripts which constitute this thesis therefore address three fundamental research questions: (a) What characterizes game enjoyment and how can it be assessed? (b) What game structures and psychological processes afford this experience? (c) How does this experience relate to behavior change?

Manuscript 1 consists of a systematic review of 87 quantitative studies, analyzing different operationalizations and measures of game enjoyment, its determinants, and how these were related to other player experience constructs, such as flow, presence and immersion. While the conceptualizations and measures of game enjoyment were inconsistent, enjoyment was commonly operationalized as intrinsic motivation or positive affect, as well as associated with need satisfaction. Further, enjoyment was distinguished from flow in that it may occur independently of challenge and cognitive involvement, ultimately suggesting that enjoyment may describe the valence of the player experience.

Seeing how game enjoyment was often equated with intrinsic motivation and predicted by need satisfaction, this then formed the basis for the studies described in manuscripts 2 and 3. Specifically, it has been debated amongst gamification researchers whether specific game elements may actually decrease users' intrinsic motivation. Manuscript 2 thus presents a study on the effects of three commonly employed game design elements – points, leaderboards, levels – on performance, intrinsic motivation, perceived autonomy and competence in an image annotation task. Implementation of these game elements significantly increased

performance, but did not affect perceived autonomy, competence or intrinsic motivation. This suggested that points, levels and leaderboards do not by themselves invariably decrease intrinsic motivation in non-game contexts.

The study described in manuscript 3 further explored the motivational mechanisms underlying the aforementioned game features. Based on self-determination theory, it was assumed that if perceived as informational, points, levels and leaderboards, would afford feelings of competence and hence enhance intrinsic motivation and promote performance gains. Moreover, the study also took participants' causality orientation – the extent to which individuals experience their actions as self-determined, which further influences whether they perceive feedback as informational or controlling, – as a potential moderator variable into account. Mirroring the results of manuscript 2, game elements did not significantly affect competence or intrinsic motivation, irrespective of participants' causality orientation. Similarly, participants' performance did not reflect their intrinsic motivation, as points, and especially levels and leaderboard led to a significantly higher amount of tags generated compared to the control group, suggesting that they functioned as extrinsic incentives, effective only for promoting performance quantity in the short term.

Games for change have attracted the interest of humanitarian aid organizations and researchers. However, their effectiveness to promote behavior remains unclear, and little is known about how the player experience relates to specific game properties and donating behavior. In the study outlined in manuscript 4, experimental conditions were systematically varied in their interactivity and presentation mode. After playing, watching or reading through the narrative of the game Darfur is Dying, participants had the option to donate to a charity. While interactivity increased donating by an average of 12%, presentation mode had no significant impact. Moreover, appreciation fully mediated the relationship between interactivity and donating, hinting at its relevance for the evaluation of games for change.

Taken together, the present findings show that specific game features may promote certain behaviors without necessarily affording enjoyment/intrinsic motivation. While manuscripts 2 and 3 hint at game elements functioning as extrinsic motivators, the findings of manuscript 4 emphasize the importance of appreciation and meaning.

Introduction

The field of Human-Computer Interaction (HCI) traditionally focused on the usability and utility of interactive systems and treated factors such as user enjoyment as secondary (Law & van Schaik, 2010). Yet in the last decade, research gradually moved "beyond the instrumental" (Hassenzahl & Tractinsky, 2006), and the emergence of third wave HCI stressed the importance of assuming a more comprehensive, holistic perspective on user experience (e.g., Blythe, Monk, Overbeeke, & Wright, 2004; Jordan, 2002; Shneiderman, 2004). Or as Norman put it: "Technology should bring more to people's lives than the improved performance of the tasks: it should add richness and enjoyment" (Norman, 2004, p. 111).

Arguably, digital games offer some of the most enjoyable, rich and engaging experiences of all interactive products. Their steady increase in popularity, surpassing the movie industry in terms of sales (ESA, 2015), and their growing importance as a medium for education, persuasion, and self-expression (Bogost, 2011), as well as the different cultural practices that have formed around gaming (e.g., let's play videos, modding, etc.) make them an immensely fascinating topic for the HCI community to study. Already back in the early 1980ies did HCI researchers point out that "[...] features of computer games can be incorporated into other user interfaces, [which then] can be made not only easier and more productive to use, but also more interesting, enjoyable and satisfying" (Malone, 1982, p. 68). They suggested that HCI look into games to tease out the specific "motivational techniques" (Carroll, 1982) and qualities that make games "enjoyable interfaces" (Malone, 1982), and how to design for them in non-game systems (Carroll & Thomas, 1988). Carroll and Thomas further urged that HCI researchers should also focus on fun, as not only are "experiences that are fun more attractive to people of course, but fun is also likely to have powerful influences on what people will even try to do and on how long they will persist" (Carroll & Thomas, 1988, p. 22).

With the emergence of UX research (e.g., Blythe et al., 2004; Hassenzahl & Tractinsky, 2006; Jordan, 2002), interest in digital games as an ideal for how to design interfaces that facilitate positive experiences (Calvillo Gámez, Cairns, & Cox, 2009) was also rekindled. It comes to little surprise then that in recent years practitioners have attempted to apply game design to enhance the user experience of non-game applications and services, an approach

nowadays commonly known under the monikers gamification or gameful design (Deterding, Dixon, Khaled, & Nacke, 2011).

According to Huotari and Hamari (2012), 'gameful experience' relates to an experiential condition unique to rule-based games ('ludus', cf. Caillois, 1958; Deterding et al., 2011) – in contrast to more free-form playful experiences ('paidia'). A game is (1) a rule-based formal system with (2) a variable and quantifiable outcome, where (3) different outcomes are assigned different values, (4) the player exerts effort in order to influence the outcome, (5) the player feels attached to the outcome, and (6) the consequences of the activity are optional and negotiable (Juul, 2011). There are undoubtedly countless types and genres of games, digital or not, and various definitions have attempted to describe the characteristics inherent to games. For the sake of the research outlined in this thesis however, "games are not really defined in terms of their physicality, but in terms of the experience they provide" (Calvillo Gámez et al., 2009, p. 520).

One of the defining aspects of such gameful experiences is that they are engaged with voluntarily (Deterding, 2015; Huotari & Hamari, 2012). However, other than that, a concrete definition of the term 'gamefulness' is currently lacking. Hence, it might be most promising to focus on the type of experiences most commonly associated with game play. Indeed, Huotari and Hamari (2012) suggested that if gamification is to create "gameful experiences", then the success of gamification should also be measured through the same measurement instruments as games are. Thus, in order to gain a better understanding of the elusive gameful experience – that is, how to conceptualize, operationalize and ultimately how to measure it, – it is crucial to gain insights into what actually constitutes the core player experience (PX) most associated with gameplay (Nacke & Drachen, 2011).

Typically, gamification intends to afford gameful experiences in non-game contexts with the goal of promoting user engagement and behavior change (Deterding et al., 2011; Seaborn & Fels, 2015). However, empirical evidence on the effectiveness of gamification to shape behaviour is mixed (Seaborn & Fels, 2015), and few studies have explicitly examined the underlying psychological mechanisms that may account for these findings. As illustrated in Figure 1, experience – for instance enjoyment (manuscript 1) or appreciation (manuscript 4), –

forms when the player interacts with a specific game or game elements (Engl & Nacke, 2013), which in turn influences the players' behavior. Game elements can range from very specific interface and feedback patterns, such as points, levels or leaderboards (manuscripts 2 and 3), or more fundamental aspects, such as a game's interactivity and presentation mode (manuscript 4). Apart from contextual factors (e.g., interacting with game elements at home versus while waiting for the bus), player characteristics and traits – such as a person's causality orientation (manuscript 3), – may also further determine the interplay between game elements, player experience and behavior.

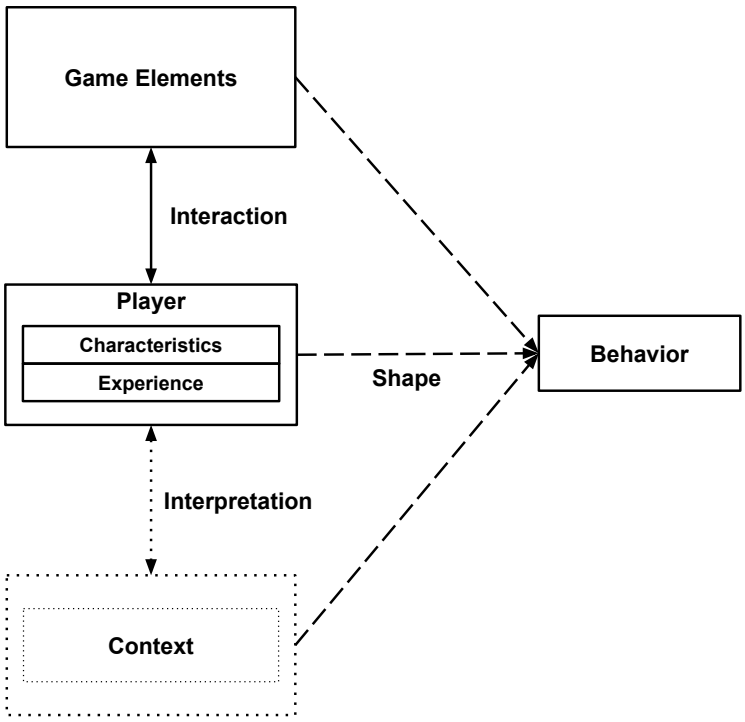

Figure 1. Framework of the relationship between game features, the player and behavior. Contextual factors are also crucial, but were not subject of the present thesis. Figure adapted from Engl and Nacke (2013).

Following from the above, this thesis covers three research questions: (a) What characterizes the player experience, specifically game enjoyment, and how can it be assessed? (b) What game structures and psychological processes afford this experience? (c) How does this experience relate to behavior change? The thesis framework is structured as follows: It

starts by providing a theoretical background consisting of two sections. The first section provides the background for the research questions on the nature of gameful experience by introducing relevant player experience concepts and frameworks and by reviewing previous empirical research on player experience. The second section elaborates on how gameful experiential qualities affect user behavior by summarizing the current state of empirical research on gamification and serious games. The theoretical background is followed by detailed summaries of the four manuscripts. Finally, the framework closes with a general discussion and conclusion of the publications by highlighting their relevance for HCI and by raising further research questions. The four manuscripts are enclosed in the appendix.

Theoretical Background

Conceptualizing and Operationalizing the Player Experience

Psychological game research has long chiefly focused on negative effects of games, such as aggressiveness (Anderson et al., 2010), addiction, and later on, on the potential benefits of gameplay, such as improved cognitive functioning (Bavelier, Green, Pouget, & Schrater, 2012). Comparatively little research has examined the player-game interaction itself, such as motives for gameplay or game properties that make for a positive gaming experience. Arguably, people play games for the experience they create (Calvillo Gámez et al., 2009; Lazzaro, 2009). However, the experience of being entertained through games is not well understood yet from a psychological perspective (Vorderer, Klimmt, & Ritterfeld, 2004; Wyeth, Johnson, & Sweetser, 2012) and there is no clear consensus on what exactly constitutes a positive player experience.

Various concepts have been proposed to describe and assess the player experience (Boyle, Connolly, Hainey, & Boyle, 2012; Caroux, Isbister, Le Bigot, & Vibert, 2015), such as immersion, flow or presence (Nacke & Lindley, 2008). The proliferation of sometimes vague, sometimes overlapping constructs, has made it difficult to decide upon which psychological concepts are best suited to study the player experience (Bernhaupt, 2010), which also poses an issue for studying gameful experiences and gamification.

Meanwhile, enjoyment has been characterized as the core experience of all entertainment media (Vorderer et al., 2004) and games in particular (Nacke & Drachen, 2011; Sweetser & Wyeth, 2005). It has been identified as a key motive for why people play games (refer to Boyle et al., 2012, for an overview) and is among the most sought after game experiences (Komulainen, Takatalo, Lehtonen, & Nyman, 2008). However, while "enjoyment" is a widely used label, it remains a largely elusive, fuzzy and overdetermined construct (Carroll & Thomas, 1988; Deterding, 2015; Nacke & Drachen, 2011; Sweetser & Wyeth, 2005). Already in 1988, Carroll and Thomas (1988) speculated that 'fun' is empirically underexplored, because it is subjective and thus difficult to measure.

Multiple models have since attempted to explain the subjective experience of enjoyment in games – often in terms of other constructs, – or to deconstruct it into subcomponents. To

further complicate matters, many models use the same terms with different meanings. For instance, game enjoyment has often been equated with the experience of flow (Sherry, 2004; Sweetser & Wyeth, 2005; Weber, Tamborini, Westcott-Baker, & Kantor, 2009). Flow describes the subjective experience of engaging in challenging, yet manageable activities (e.g., digital games), further characterized by complete cognitive absorption, time distortion and enjoyment (Nakamura & Csikszentmihalyi, 2002). But some PX scholars have argued that flow is too restrictive, as it is only concerned with extreme experiences (Jennett et al., 2008; Takatalo, Häkkinen, Kaistinen, & Nyman, 2010) and may not adequately cover more casual experiences of enjoyment, such as gaming on the phone while waiting for the bus. In fact, Nakamura and Csikszentmihalyi (2002) have shown that enjoyment may occur independently of the flow experience. Moreover, flow experiences may not be appropriate nor desirable for certain gamified services, for instance, when designing an application for young and/or inexperienced drivers to learn how to drive safely (Fitz-Walter, Wyeth, Tjondronegoro, & Scott-Parker, 2013).

Others consider flow and enjoyment distinct concepts. The Game Experience Questionnaire (GEQ), for instance, defines enjoyment as a multi-dimensional construct, made up of challenge, competence, (minimal) frustration and positive affect, whereas flow constitutes part of the involvement construct, which also includes immersion and (lack of) boredom (Gajadhar, Nap, de Kort, & IJsselsteijn, 2010). In contrast, Jennett et al. (2008) define immersion as the key component of good game experiences. However, in one experiment involving a non-game task they found that high immersion could also be accompanied by anxiety and negative affect, which makes it difficult to infer whether the task was not enjoyable or whether participants simply felt tense due to the task's increasing challenge. In short, these inconsistent, often overlapping conceptualizations make it difficult to form a common definition of game enjoyment, which poses serious challenges to effectively measuring the construct, thereby hampering our understanding of game(ful) experiences.

In order to explore question (a) 'What characterizes game enjoyment and how can it be assessed?', manuscript 1 describes a systematic literature review, which aimed at providing an overview of current operationalizations and determinants of game enjoyment, as well as the

methodological challenges that may arise from them. Some of the findings presented in manuscript 1 then form the basis for the studies described in manuscripts 2, 3 and 4, which all explored the role of enjoyment in relation to behavior change in 'purposeful' contexts resorting to game design elements.

Effectiveness of Shaping Behavior Through Game Design

Two prominent approaches have emerged to shape behavior through game design – serious games and gamification/gameful design. The distinction between an assortment of game elements and a full-fledged games is subjective, but both have in common that they aim to leverage the appeal of digital games to shape behavior and attitudes in non-game contexts elements (Deterding et al., 2011).

Industry publications promise "revolutionary" gains in customer and employee motivation through gamification (Paharia, 2013; Zichermann & Cunningham, 2011). And while most empirical studies yielded at least partially positive results when it comes to shaping user behavior through the implementation of game elements (refer to Seaborn & Fels, 2015, for an overview), some scholars have cautioned against the over-reliance on such elements, as they may diminish users' intrinsic interest and hence lead them to stop engaging with the application or service altogether (Deterding et al., 2011; Koivisto & Hamari, 2014; Seaborn & Fels, 2015).

However, the majority of currently available gamification literature focuses predominantly on studying the effectiveness of game design elements in promoting certain behavioral outcomes (Seaborn & Fels, 2015), largely ignoring the underlying psychological and experiential mechanisms that may actually account for these effects (Antin & Churchill, 2011; Deterding, 2015), (but refer to Hanus & Fox, 2015; Lieberoth, 2015; Mekler, Brühlmann, Opwis, & Tuch, 2013b; Mekler, Brühlmann, Tuch, & Opwis, 2015, for notable exceptions). To further complicate matters, existing empirical research suffers from several methodological drawbacks (Seaborn & Fels, 2015): Firstly, most studies do not report effect sizes, and some do not even conduct any statistical analyses. Secondly, comparative experiments are rare, as most studies do not employ control conditions, making it difficult to

draw any causal inferences from the results. Also, few studies base their studies on existing theories or concepts, or use already validated measuring instruments. Issues already prevalent in research on game enjoyment (Mekler, Bopp, Tuch, & Opwis, 2014).

Finally, most empirical studies on gamification and serious games investigate the impact of multiple game elements (e.g., Peng, Lee, & Heeter, 2010; Peng, Lin, Pfeiffer, & Winn, 2012), making it difficult to pinpoint how and to what extent these game elements contribute to users' experience and behavior (Seaborn & Fels, 2015). Moreover, most pattern-based approaches to gamification (e.g., Francisco-Aparicio, Gutiérrez-Vela, Isla-Montes, & Sanchez, 2013; Zichermann & Cunningham, 2011) offer little guidance in deciding which and whether game elements are suitable for a given context, or how they should be implemented (Deterding, 2015). Studying the effects of individual game elements on users' experience, as well as whether and how this in turns affects their behavior, thus contributes to gamification and serious game research by providing a more nuanced understanding of how particular game elements function in a given context, and may potentially benefit designers, as it allows for more informed decisions on how and under what circumstances game elements should or should not be implemented (Seaborn & Fels, 2015).

In order to explore questions (b) 'What game structures and psychological processes afford this experience?' and (c) 'How does this experience relate to behavior change?', manuscripts 2, 3 and 4 all describe experimental studies that isolated the effects of different game properties. This made it possible to trace back to what extent each game component affected (the interplay between) experience and behavior. Moreover, all studies employed validated, theory-based instruments, in addition to measures of behavioral outcomes.

Summary of the Manuscripts

The following manuscripts constitute this thesis. The first, second and fourth manuscript have already been published, whereas manuscript 3 is in press.

1. Mekler, E. D., Bopp, J. A., Tuch, A. N. & Opwis, K. (2014). A Systematic Review of Quantitative Studies on the Enjoyment of Digital Entertainment Games. *CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 927-936. http://dx.doi.org/10.1145/2556288.2557078

2. Mekler, E. D., Brühlmann, F., Opwis, K. & Tuch, A. N. (2013). Do Points, Levels and Leaderboards Harm Intrinsic Motivation? An Empirical Analysis of Common Gamification Elements. *Gamification 2013: Proceedings of the First International Conference on Gameful Design, Research, and Applications*, 66-73. http://dx.doi.org/10.1145/2583008.2583017

3. Mekler, E. D., Brühlmann, F., Tuch, A. N. & Opwis, K. (in press). Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance. *Computers in Human Behavior*. http://dx.doi.org/10.1016/j.chb.2015.08.048

4. Steinemann, S. T., Mekler, E. D. & Opwis, K. (2015). Increasing Donating Behavior Through a Game for Change: The Role of Interactivity and Appreciation. *CHI PLAY '15 Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. http://dx.doi.org/10.1145/2793107.2793125

The following work-in-progress and workshop publications are related to this thesis, but were omitted for the sake of brevity and focus. Some of them will however be referenced in the general discussion section.

- Bopp, J. A., Mekler, E. D., & Opwis, K. (2015). "It Was Sad But Still Good": Gratifications of Emotionally Moving Game Experiences. *CHI '15 Extended Abstracts*, 1193-1198. http://dx.doi.org/10.1145/2702613.2732852

- Forde, S. F., Mekler, E. D. & Opwis, K. (2015). Informational vs. Controlling Gamification: A Study Design. *CHI PLAY '15 Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. http://dx.doi.org/10.1145/2793107.2810297

- Mekler, E. D., & Bopp, J. A. (2015). Exploring the False Affective Dichotomy in Games – Emotions and Meta-Emotions. *CHI PLAY '15 Workshop "The False Affective Dichotomy between Positive and Negative Affect in Game Play"*. http://falsedichotomy2015.files.wordpress.com/2015/09chiplay2015dichotomy_paper_2.pdf

- Mekler, E. D., Brühlmann, F., Opwis, K. & Tuch, A. N. (2013). Disassembling Gamification – The Effects of Points and Meaning on User Motivation and Performance. *CHI '13 Extended Abstracts*, 1137-1142. http://dx.doi.org/10.1145/2468356.2468559

- Mekler, E. D., Tuch, A. N., Martig, A. L. & Opwis, K. (2014). A Diary Study Exploring Game Completion and Player Experience. *CHI PLAY '14 Proceedings of the first ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play*, 433-434. http://dx.doi.org/10.1145/2658537.2661304

- Müller, L. J., Mekler, E. D. & Opwis, K. (2015). Facets in HCI: Towards Understanding Eudaimonic UX – Preliminary Findings. *CHI '15 Extended Abstracts*, 2283-2288. http://dx.doi.org/10.1145/2702613.2732836

1. A Systematic Review of Quantitative Studies on the Enjoyment of Digital Entertainment Games

Motivation and aim of the literature review. People play games for the experience they create (Lazzaro, 2009). However, the experience of being entertained through games is not well understood yet from a psychological perspective (Deterding, 2015; Vorderer et al., 2004; Wyeth et al., 2012). Various concepts have been discussed, such as immersion, flow or presence (Boyle et al., 2012; Caroux et al., 2015), making it difficult to decide upon which psychological concepts are best suited to study the player experience (PX) (Bernhaupt, 2010). Meanwhile, enjoyment has been characterized as the core experience of all entertainment media (Vorderer et al., 2004) and games in particular (Nacke & Drachen, 2011; Sweetser & Wyeth, 2005), has previously been identified as a key motive for why people play games (refer to Boyle et al., 2012, for an overview) and is among the most sought after game experiences (Komulainen et al., 2008). Unfortunately, the definition of game enjoyment is fuzzy (Sweetser & Wyeth, 2005) and not well differentiated from other potentially related psychological components of PX, such as flow (Nacke & Drachen, 2011). This raises the issue of how the experience of playing digital games can be effectively operationalized and measured (Calvillo-Gámez, Cairns, & Romero, 2012; Wyeth et al., 2012). Moreover, many models of PX often neglect previous findings and/or are based on little empirical grounding (Nacke & Drachen, 2011; Wyeth et al., 2012). For these reasons, a common terminology for discussing and measuring PX and game enjoyment is still lacking, making it difficult to compare study findings and assess which factors contribute most to enjoyable game experiences. The goal of this systematic literature review was to provide a comprehensive overview of previous quantitative research on game enjoyment, including the different conceptualizations and operationalizations of game enjoyment currently employed, in order to gain a clearer understanding of how the experience of enjoyment when playing digital games may be adequately measured. The review also investigated how other psychological components of PX relate to game enjoyment, in order to further clarify the concept. Finally, the review identified and categorized relevant determinants of game enjoyment, thereby uncovering opportunities for further research

Method. For conducting the systematic literature review the QUOROM procedure (Moher et al., 1999) was followed, as it had already been successfully used in a previous literature review on user experience (Bargas-Avila & Hornbæk, 2011). The terms 'game' and 'enjoyment' were used to search three publication databases: ACM Digital Library, ScienceDirect and ISI Web of Knowledge. Only the term 'enjoyment' was chosen, because it has been argued that different terms, such as liking, fun or preference, which are often used synonymously for enjoyment, do not cover the exact same meaning (Blythe & Hassenzahl, 2005; Nabi & Krcmar, 2004). Exceptions were made if a study employed the Game Experience Questionnaire (Gajadhar et al., 2010) or based its questionnaire on the GameFlow model (Sweetser & Wyeth, 2005), as the original authors explicitly link these concepts to game enjoyment. The literature review deliberately exclusively included studies who examined the enjoyment of entertainment games, that is, games, whose "only" intended purpose is to generate enjoyment (cf. Nacke & Drachen, 2011; Wyeth et al., 2012), because we were interested in the 'core experience of enjoyment', undiluted by any outside purposes. A total of 87 studies from 68 papers were included in the final analysis and all studies were coded in terms of aim of the study, measures, results, measuring times, study designs, participants, game genre, duration of gameplay, and if provided, definition of enjoyment.

Results. Few of the reviewed studies attempted to define enjoyment, and the majority used self-developed, unvalidated scales. However, those that did, most often operationalized enjoyment as either intrinsic motivation – that is, pursuing an activity because it is inherently interesting, rather than due to some external outcome, – or positive affect, which was most prevalent in studies employing facial electromyography (e.g., Poels, Hoogen, Ijsselsteijn, & de Kort, 2012) or studies using the Game Experience Questionnaire (e.g., Gajadhar et al., 2010, see also Brühlmann and Schmid (2015) for an analysis of its factor structure). Moreover, game enjoyment was often associated with (e.g., Przybylski, Ryan, & Rigby, 2009) or even defined in terms of psychological need satisfaction (e.g., Tamborini, Bowman, Eden, Grizzard, & Organ, 2010). Notably, of the reviewed papers only studies published in psychology (e.g., Przybylski et al., 2009) and communication studies (e.g., Reinecke et al., 2012) journals applied the Player Experience Need Satisfaction (PENS) framework (Ryan, Rigby, &

Przybylski, 2006) and measured enjoyment with the Intrinsic Motivation Inventory. No full papers published between 2008 and 2012 at relevant HCI venues, such as CHI or International Journal of Human Computer Studies, had employed these conceptualizations. Enjoyment was also negatively correlated with feelings of guilt (e.g., Hartmann & Vorderer, 2010), and in one case even operationalized as the opposite of moral distress (Gollwitzer & Melzer, 2012). These findings are in line with the concept of serious fun proposed by Lazzaro (2009), which suggests that players experience enjoyment when a game reflects their values (e.g., the absence of guilt) and positively affects their thoughts and feelings (e.g., through need satisfaction).

Enjoyment was often related to or subsumed under other player experience constructs, such as flow, immersion or presence. Twenty-four studies measured flow in some form, with varying definitions of how enjoyment and flow relate to each other. Some used the terms "flow" and "enjoyment" interchangeably (e.g., Jin, 2011). Studies that rely on the GEQ argue that flow is a dimension of player involvement rather than enjoyment (e.g., Gajadhar et al., 2010). Lastly, some state that the two constructs share similarities and that enjoyment results from the flow experience (e.g., Weibel, Wissmath, & Mast, 2011). These inconsistent definitions stem largely from the different aspects of flow that the studies chose to focus on. Several studies associated flow with focused attention (Jin, 2011; Weber et al., 2009) and the balance of skill and challenge (Jin, 2012; Nacke & Lindley, 2008). However, Shim, Srivastava, and Hsu (2011) found that the balance of skill and challenge – arguably a defining factor of flow – did only partially account for game enjoyment on its own. Rather, this was further dependent on players' motivations for game-play. In contrast, Limperos, Schmierbach, Kegerise, and Dardis (2011) found that the experience of control – an aspect also associated with flow – was related to enjoyment, but not to other characteristics of flow. These studies suggest that enjoyment may be experienced independently from flow.

Discussion and conclusion. Based on the review of 87 recent quantitative studies, it was concluded that game enjoyment is commonly understood as the positive cognitive and affective appraisal of the game experience, associated with need satisfaction and the absence of guilt. Surprisingly, of the reviewed studies, none of the papers published at HCI venues employed the PENS framework, and only recently has player experience research started to

take up this concept (e.g., Birk & Mandryk, 2013; Johnson, Nacke, & Wyeth, 2015; Seif El-Nasr, Durga, Shiyko, & Sceppa, 2015).

Arguably, game enjoyment shares many similarities with flow. But even though enjoyment is a crucial aspect of the flow experience, they still differ in specific ways: It seems that the experience of challenge, enjoyment and deep concentration are the main characteristics of flow, while other aspects, such as a sense of control, facilitate flow and make games more enjoyable, but do not trigger flow by themselves. Put differently, flow encompasses both enjoyment and involvement, triggered by the optimal balance of challenging gameplay and player skills, but players may experience enjoyment independently of flow. Limiting game enjoyment to the experience of flow would therefore fail to account for the variety of enjoyable experiences that games may provide. Nevertheless, more empirical research is required to further probe under what circumstances game enjoyment and flow occur and whether the flow experience is characterized by particularly deep enjoyment. This conceptualization of game enjoyment has some potential theoretical implications for PX research, namely that game enjoyment may be understood as the valence of the player experience. In contrast, the intensity of the player experience may perhaps be represented by players' immersion, involvement or engagement, that is, the extent to which players' attention is held by gameplay challenges and the game environment (i.e., presence). Put differently, the more immersive a game, the more intense the player experience, eventually culminating in cognitive absorption and time distortion. Flow may thus be explained in terms of a very intense, yet positive experience (see Figure 2).

This framework supports previous conceptualizations of flow as effortless attention (Calvillo-Gámez et al., 2012) and immersion as a graded experience (Jennett et al., 2008). However, more research is needed to empirically assess its validity. For instance, whether a non-intense gaming experience may be just as enjoyable as an intense one, and how the valence and the intensity of the player experience relate to each other. Moreover, it is necessary to consider how the intensity of the player experience can be operationalized and measured, how it is affected by challenge and presence and how this relates to need satisfaction, as well as to cognitive and affective aspects of game enjoyment. Unraveling this

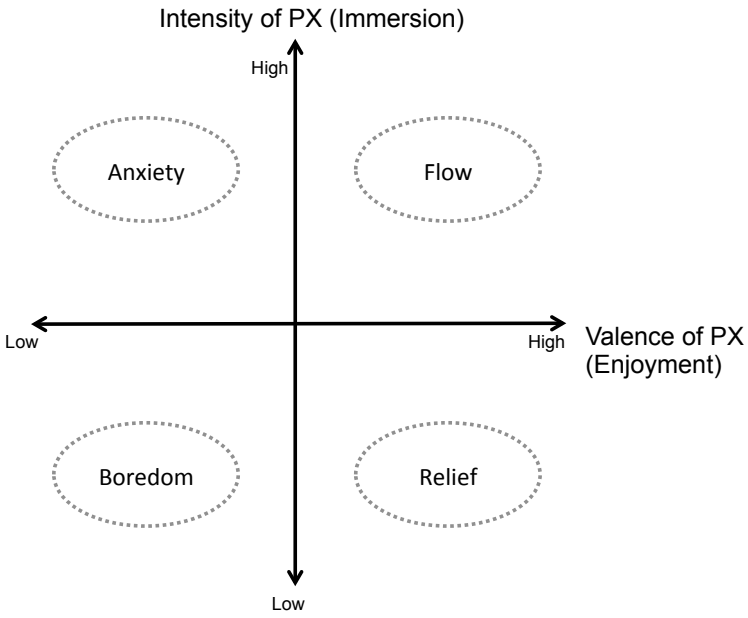

Figure 2. Valence and intensity of PX framework. Enjoyment describes the valence of the player experience, whereas immersion may denote its intensity.

issue is especially relevant in light of the increasing realism of digital games, the resurgence of VR technology, such as the Oculus Rift, and for research on the negative effects of gaming, but also for (research on) gameful design.

2. Do Points, Levels and Leaderboards Harm Intrinsic Motivation? An Empirical Analysis of Common Gamification Elements

Motivation and aim of the study. Gamification has been commonly associated with game elements, such as points, levels and leaderboards (Seaborn & Fels, 2015) – "the things that are least essential to games" (Robertson, 2010), – which has irked game designers and psychologists alike. Some have cautioned against the over-reliance on such elements, as they may diminish intrinsic interest in both game- and non-game contexts, ultimately leading users to stop interacting with the application or service altogether (Deterding et al., 2011; Hecker, 2010). Indeed, previous research from motivational psychology has found that extrinsic incentives can reduce intrinsic motivation in various contexts (see Deci, Koestner, & Ryan, 1999, for a meta-analysis). In other words, a person is no longer intrinsically drawn towards engaging in an activity, because s/he is pushed to do so through external means. This is unfortunate, as intrinsic motivation is not only associated with improved psychological well-being, but also benefits the extent and quality of effort that people put into a given task, which results in enhanced performance, creativity and learning outcomes in a variety of domains (Ryan & Deci, 2000). However, there is little empirical evidence on whether and under what circumstances these game elements may actually undermine users' intrinsic motivation (Hecker, 2010). The aim of the this study was thus to examine how three of the most commonly employed game elements – points, leaderboards, levels, – affect user behavior and whether they invariably undermine intrinsic motivation in an image annotation task.

Method. To test our hypotheses, we conducted a between-subject online experiment. The independent variable were three of the most common game elements: Points vs. leaderboard vs. levels vs. control condition. The dependent variables were user performance (amount of tags, "cheating behavior", time spent on task), intrinsic motivation and satisfaction of autonomy and competence needs. In the points condition, 100 points were awarded for every tag generated, regardless of its actual content, and the score was displayed in the upper right corner of the screen. In the leaderboard condition, participants could compare their current score to four fictitious participants in a leaderboard on the right-hand side of the screen. Participants were deliberately left unaware of the fact that fictional participants

occupied the leaderboard. This static leaderboard was implemented so that all participants had the same chance to rise in ranks, as the leaderboard positioning may have had a confounding effect on motivation otherwise (Von Ahn & Dabbish, 2008). To reach the lowest position on the leaderboard, participants had to generate at least ten tags. For each subsequent position, participants had to score 3000, 6000, and 10000 points respectively. These step sizes were chosen to allow participants to reach a reasonably high position on the leaderboard, but it was expected to be still reasonably challenging for participants to come up with more than 100 tags. In the levels condition, participants were presented with a vertical progress bar labeled with "next level" and the corresponding points necessary to reach the indicated level. Progression to the next level mirrored the leaderboard condition, albeit without the option for (seemingly) social comparison.

A total of 295 participants (191 female) took part in the experiment, where they had to generate as many tags as possible for describing 15 abstract painting. Five $50 gift coupons for an online consumer electronics retailer were raffled among all participants. The raffle was deliberately chosen as incentive, because it was assumed that it would not distort the experimental effects of game elements on intrinsic motivation, due to being a form of unexpected, task-noncontingent reward. Task-noncontingent rewards have previously been found not to affect intrinsic motivation, as receiving them does not require doing or completing the task and hence they are not perceived as controlling (Deci et al., 1999).

Results. Participants in the game element conditions generated significantly more tags than participants in the control condition ($F(3, 291) = 11.109$, $p < .001$, $\eta_p^2 = .102$). Planned contrasts showed that participants in the points condition significantly outperformed participants in the control condition ($F(1, 153) = 10.523$, $p = .001$, $\eta_p^2 = .064$), and were in turn significantly outperformed by participants in the leaderboard ($F(1, 154) = 5.23$, $p = .024$ $\eta_p^2 = .033$) and level conditions ($F(1, 151) = 3.91$, $p = .050$ $\eta_p^2 = .026$). Performance did not differ between the leaderboard and levels conditions. Additionally, a a 3x4 repeated measures ANOVA with time (3 blocks with 5 images each) as within-subject factor and condition as between-subject factor revealed a significant time x condition interaction ($F(6, 582) = 2.462$, $p = .023$, $\eta_p^2 = .025$), as well as a significant main effect of time on performance ($F(6, 582) =$

17.447, $p < .001$, $\eta_p^2 = .057$) was found. Although performance over all experimental conditions decreased over time (see Figure 3), participants' performance in the leaderboard and levels conditions declined more slowly than in the other two conditions. No significant effects of game elements on intrinsic motivation were found ($p = .499$). Participants reported similar levels of task enjoyment and interest, regardless of whether they received feedback in form of points, leaderboard, levels, or none at all.

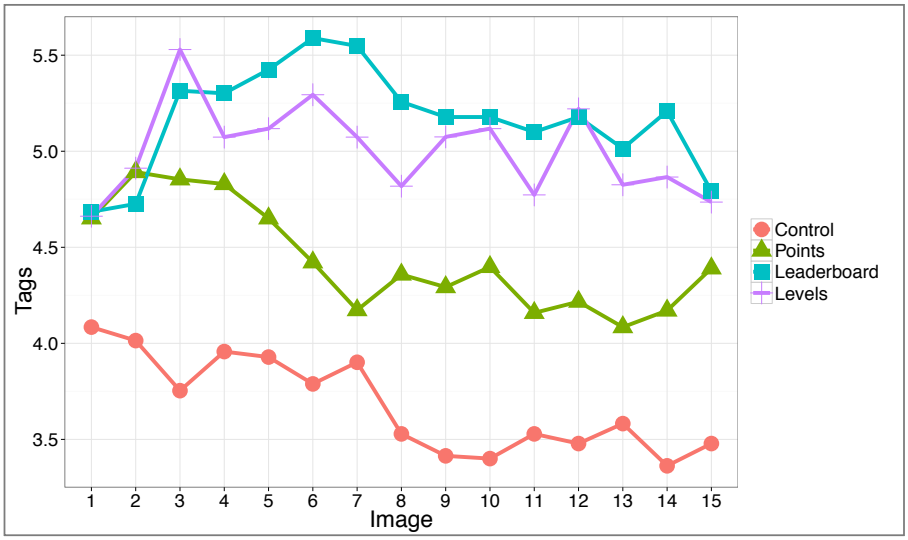

Figure 3. Average number of user-generated tags per condition over the course of 15 images.

Discussion and conclusion. This study was the first to provide empirical evidence that game elements, such as points, levels or leaderboards, need not invariably undermine intrinsic motivation. In contrast to the study by Hanus and Fox (2015), intrinsic motivation was not negatively affected by the addition of game elements. Apparently, participants did not perceive the game elements as particularly controlling, compared to the control condition. Perhaps, points, levels and leaderboards by themselves did not affect intrinsic motivation negatively, because they were not linked to other potentially pressuring external events, such as cash prizes for the best performance (Deci et al., 1999). Note that in the study of Hanus and Fox (2015) badges were incorporated in a school setting and collecting badges was a mandatory requirement for receiving course credits. In contrast, participants in the present study found themselves in a relatively laid-back setting, as they chose to participate in the study voluntarily and, in case of the leaderboard condition, competing against people whom

they did not know. Thus, it is assumed that due to the relative absence of controlling factors, participants' need for autonomy was not threatened.

While this experiment provided evidence that points, levels and leaderboards are an effective means to increase short-term performance without forcibly decreasing intrinsic motivation, the study featured two major shortcomings that warrant mention. Firstly, only measures of participants' self-reported intrinsic motivation were included. While self-reported and "free choice" measures of intrinsic motivation yielded comparable results in previous studies (Deci et al., 1999), it would have been valuable employ a behavioral "free choice" measure of intrinsic motivation by letting participants choose whether they want to continue engaging with the image annotation task, even after the conclusion of the experiment. Secondly, the task itself scored only slightly above average on intrinsic motivation. As rewards only threaten intrinsic motivation for activities that people find interesting (Deci et al., 1999), game elements may not have significantly affected intrinsic motivation, due to the task not being interesting enough.

3. Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance

Motivation and aim of the study. Satisfaction of the psychological need for competence has been found to promote intrinsic motivation and game enjoyment (cf. manuscript 1). It has been argued that video games excel at competence need satisfaction, because they offer copious amount of immediate feedback in the form of various performance indicators, including points, levels and leaderboards (Przybylski, Rigby, & Ryan, 2010). While rewards may often be perceived as controlling, and hence decrease intrinsic motivation (cf. manuscript 2; also Hanus & Fox, 2015), positive informational feedback has the potential to support competence need satisfaction and thus increase intrinsic motivation. Several studies have already discussed (Francisco-Aparicio et al., 2013) or empirically studied (Peng et al., 2012) this notion, but none have systematically examined individual game elements, in order to isolate their effects on user motivation and behavior. The aim of this experiment was therefore to study whether points, levels and leaderboards increase competence need satisfaction, which in turn was expected to increase intrinsic motivation and boost performance quantity and quality in an image annotation task. Moreover, because apart from situational factors, individual differences may also account for the differing effects of gamification (Koivisto & Hamari, 2014), participants' causality orientation, that is, to what degree they tend to experience feedback as controlling or informational, was also assessed (see Figure 4).

Method. The study employed the same study design that was described in manuscript 2. In total, 273 participants (178 female) took part in the study. Participants were randomly assigned to one of the four conditions (points, levels, leaderboards, control group with no game elements) and were asked to tag 15 abstract paintings. As it was expected that causality orientation further moderated the effects of game elements on behavior and motivation, participants were also asked to fill in the General Causality Orientation Scale (Deci & Ryan, 1985). Each vignette describes an incident and lists two ways of responding to it, whereupon participants state how likely it is that they would respond in such a way.

Results. As in the previous study, game elements significantly increased the amount of tags generated ($F(3, 265) = 10.09$, $p < .001$, $\eta_p^2 = .103$), but did not significantly affect

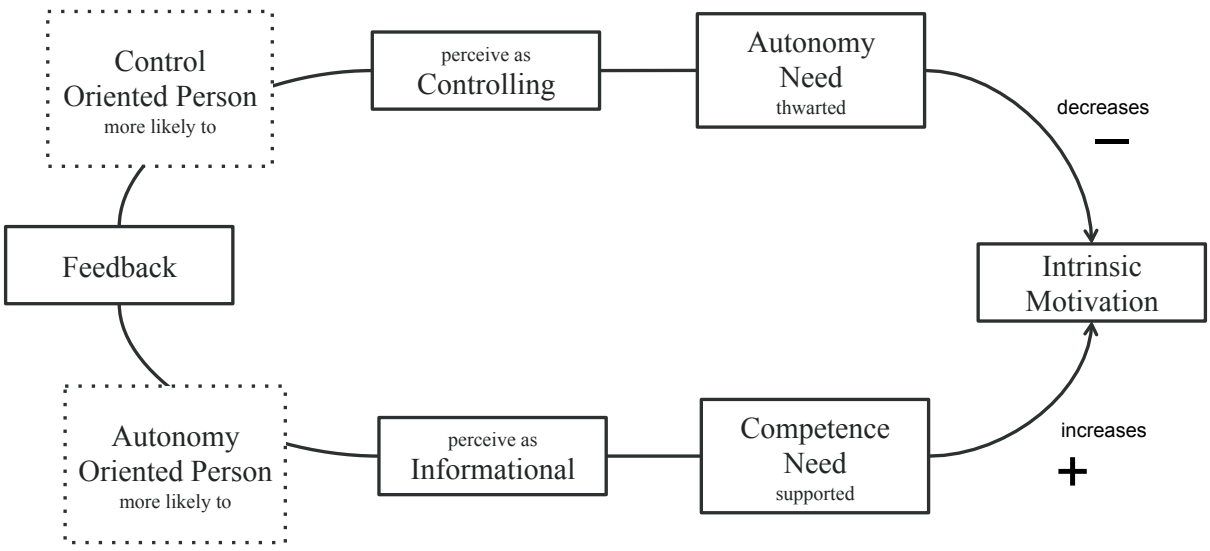

Figure 4. Feedback may be perceived as controlling or informational, thereby affecting need satisfaction and intrinsic motivation in different ways. A person's causality orientation may further moderate how feedback affects need satisfaction and intrinsic motivation.

intrinsic motivation, competence need satisfaction or tag quality. Autonomy oriented participants were significantly more intrinsically motivated ($F(7, 265) = 6.903$, $p = .009$, $\eta_p^2 = .03$) and generated significantly more tags ($F(1, 265) = 5.31$, $p = .022$, $\eta_p^2 = .02$) than control oriented participants, but causality orientation did not further moderate the effects of game elements on performance or motivation. Overall, tag quantity was negatively correlated with tag quality ($r = -.38$, $p < .001$), as well as slightly positively correlated with intrinsic motivation ($r = .14$, $p < .05$) and competence need satisfaction ($r = .18$, $p < .001$). Intrinsic motivation was positively correlated with competence need satisfaction ($r = .41$, $p < .001$).

Discussion and conclusion. The present study is one of the first to cover several aspects still underexplored in current gamification research (Seaborn & Fels, 2015). Firstly, it attempted to empirically evaluate the impact of gamification on intrinsic motivation and need satisfacition, two of the most frequently appealed to, yet seldom empirically studied constructs in gamification literature (Seaborn & Fels, 2015). It did so by employing a validated, theory-based instrument, the Intrinsic Motivation Inventory, in addition to measures of two different behavioral outcomes, tag quantity and quality. As of now, this study is also one of the first to isolate the effects of individual game elements in a comparative experiment.

While game elements did not affect tag quality, tag quantity was a significant negative predictor of tag quality in all experimental conditions. This may indicate that participants motivated to generate many tags, might have disregarded tag quality in favor of tag quantity. Notably, this also means that participants in the gamified conditions generated more tags than participants in the plain condition, but at a comparable quality. Overall, it could be argued that they performed better than participants who were not presented with any game elements, further hinting at the potential effectiveness of gamification to promote certain behaviors.

Further, while tag quantity was slightly positively correlated with intrinsic motivation, participants' reported intrinsic motivation did not reflect their performance. Interestingly, a recent meta-analysis by Cerasoli, Nicklin, and Ford (2014) found that intrinsic motivation only moderately predicted performance quantity in a variety of domains, whereas extrinsic incentives were found to be strong positive predictors. The fact that game elements did not increase competence need satisfaction or intrinsic motivation, but still improved tag quantity, suggests that in this particular study context, points, levels and leaderboards may have functioned as (effective) extrinsic incentives.

4. Increasing Donating Behavior Through a Game for Change: The Role of Interactivity and Appreciation

Motivation and aim of the study. Games for change, also known as social impact games or serious games for social change, are digital games with the purpose of not only entertaining, but reaching players and motivating them to support the social change the game is advocating. Very few studies have examined the impact of games for change on behavior-related variables. The goal of the study was twofold: Firstly, gain a better understanding of the role of interactivity by expanding upon the experiment of Peng et al. (2010). This was achieved by employing a more fine-grained experimental design, including an interactive text condition, as well as interactive resp. non-interactive text with pictures conditions, in addition to the original gameplay, gameplay video, and text conditions used in the study by Peng et al. (2010). Secondly, this study examined the interplay of experiential constructs, such as enjoyment and appreciation, and donating behavior, in order to learn more about the psychological mechanisms underlying games for change.

Method. Two hundred thirty four participants (131 female) were randomly assigned one of six variants of the narrative of Darfur is Dying, varying in terms of interactivity (interactive vs. non-interactive) and presentation mode (text only, text and pictures, animation). The interactive text, and text with pictures conditions were created by modifying the non-interactive text in Twine, an online tool for designing hypertext-based interactive fiction. Pictures were screenshots from Darfur is Dying. After reading/watching/playing through the story of a young Darfuri girl being captured by the Janjaweed militia on her way to fetching water from a well, participants rated their appreciation and enjoyment of the narrative (Oliver & Bartsch, 2010). Afterwards, they were offered a $1 bonus in addition to the $1 that they were already receiving as compensation for taking part in the study. Participants had then the option to choose which percentage of this $1 bonus they wanted to have paid to them and which percentage should be donated for them to the charity Save Darfur (10-percent increments between 0% and 100%).

Results. There was a significant main effect for interactivity ($F(1, 228) = 4.427$, $p = .036$, $\eta_p^2 = .019$), but not for presentation mode. Percentage donated was significantly higher

in the interactive conditions (M = 62.52, SD = 39.45) than in the noninteractive conditions (M = 50.76, SD = 40.98). Interactivity also significantly increased enjoyment ($F(1, 228) = 33.99$, $p < .001$, $\eta_p^2 = .13$) and appreciation ($F(1, 228) = 6.05$, $p = .015$, $\eta_p^2 = .026$) for the narrative. However, only the latter was significantly correlated with donating ($r = .25$, $p < .001$).

To further probe into the potential relation between interactivity, appreciation and donating behavior, a mediation analysis was calculated. To this end, two path-models were set up, as seen in Figure 5. The first path model examined the direct effect of interactivity on percentage donated, while the second path model included appreciation as a mediator variable. As had already been found in the ANOVA, the first path model revealed a significant direct effect of interactivity on percentage donated ($\beta = .14$, b = 11.76, SE = 5.298, t = 2.20, $p = .026$). The second path model revealed significant paths from interactivity to appreciation ($\beta = .17$, b = .49, SE = .182, t = 2.71, $p = .007$) and appreciation to percentage donated ($\beta = .23$, b = 6.69, SE = 1.86, t = 3.561, $p < .001$), while the path from interactivity to percentage donated was now no longer significant ($\beta = .104$, b = 8.46, SE = 5.24, t = 1.62, $p = .106$), indicating that the effect of interactivity on percentage donated is fully mediated by appreciation.

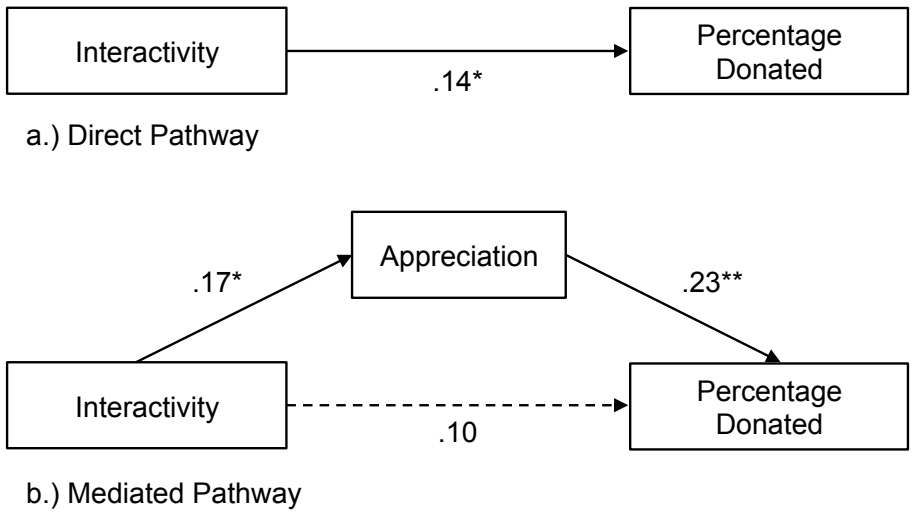

Figure 5. The relationship between interactivity and percentage donated, fully mediated by appreciation. * $p < .05$. ** $p < .01$.

Discussion and conclusion. For game designers and organizations aiming to create games for change, the main takeaway from this study is that while interactivity is crucial for

the effectiveness of games for change to encourage donating, presentation mode is seemingly less important. Notably, in this study this meant that using a simple interactive text was almost exactly as effective at motivating participants to donate, as the video game Darfur is Dying. However, an important limitation of this study is that the effectiveness of the game only refers to the behavior of players after being prompted to play the game. It is very possible that while presentation mode may not be important for increasing donating behavior, it may increase the likelihood of a player noticing or seeking out a game, as a video game may look more interesting than a text-based game.

Due to the systematic experimental manipulation of interactivity across the three presentation modes, conclusions may now be drawn as to their direct effects on the examined dependent variables. Only interactivity significantly increased enjoyment, appreciation, and donating behavior, while presentation mode had no significant impact. Specifically, enjoyment was increased by interactivity, but was not correlated with percentage donated. In other words, while participants clearly enjoyed the interactive conditions more than the non-interactive conditions, this did not necessarily make them donate a higher percentage. This recalls the findings of Cohen (2014), who found that while enjoyment increased the intention to share a game for change, it was not associated with a higher likelihood of later (self-reported) sharing. A caveat for this study however, is that while interactivity increased enjoyment, the primary goal of the game makers had most likely not been to make the narrative of Darfur is Dying enjoyable (Neys & Jansz, 2010).

Appreciation was particularly relevant in the context of the present study, as it was not only increased by interactivity, but also fully mediated the relationship between interactivity and percentage donated. While the nature of the present study does not allow for any causal inferences, this may suggest that participants found a narrative that they could actively participate in more meaningful than a narrative they were passively consuming and this then possibly encouraged them to donate a larger percentage of their bonus. This is in line with previous research on the ability of games to be thought-provoking (Bopp, Mekler, & Opwis, 2015; Iacovides & Cox, 2015; Marsh & Costello, 2012) and findings on the relationship between meaningfulness and prosocial behavior (Apter, Spirn, Sveback, & Apter, 1997;

Myrick & Oliver, 2014). However, this was the first study to find evidence for a potential connection between game properties, appreciation, and prosocial behavior. These findings indicate the importance of including appreciation in the examination of the effectiveness of games for change, as well as highlighting its potential for encouraging prosocial behavior, such as donating.

General Discussion

Manuscript 1 identified intrinsic motivation and need satisfaction as not only crucial to game enjoyment, but also some of the most thoroughly empirically validated operationalizations of the concept. Seeing how this framework was also deemed promising for gamification research and gameful design (e.g., Deterding, 2015), the studies described in manuscript 2 and 3 set out to explore and empirically validate these concepts' relevance for gamification. Even though both studies could denote significant performance gains, no significant effects on intrinsic motivation or need satisfaction, positive or negative, could be observed. Recent findings in communication studies and media psychology, based on an expanded version of the player experience need satisfaction framework (Ryan et al., 2006), – suggest that games may evoke both enjoyment and appreciation (Oliver et al., 2015), and the latter has also been associated with prosocial behavior (Myrick & Oliver, 2014). Hence, manuscript 4 shifted the focus towards the role of these types of experiences when engaging with a game for change. While interactivity did significantly promote enjoyment and appreciation, only the latter was associated with increased donating behavior.

All three empirical studies described in this thesis provide evidence that already very simple game(-like) properties, – points, levels, or a piece of interactive fiction, which asked players to choose their decisions via hyperlink, – have the potential to shape user behavior. However, the role of enjoyment and intrinsic motivation was less straightforward. In contrast to the assumption that enjoyment and intrinsic motivation facilitate behavior change, the empirical studies presented in this thesis suggest that under certain circumstances enjoyment is either not necessary (manuscripts 2 and 3) or less effective than the experience of appreciation (manuscript 4) for motivating people to showcase a certain behavior (generating tags resp. donating). Indeed, while manuscript 2 showcased that adding points, levels and leaderboards to a task need not invariably undermine intrinsic motivation, the results of manuscript 3 indicate that superficially implemented game elements act as extrinsic motivators.

It need be stated that adding features reminiscent of games, such as points, levels and leaderboards, does 'motivate' users to show more of a given behavior. However, it is important to note that both extrinsic and intrinsic motivation promote performance gains (see

Cerasoli et al., 2014, for an overview), but only the latter has been associated with improved psychological well-being, enhanced creativity and learning outcomes (Ryan & Deci, 2000), as well as increases in the extent and quality of effort that people put into a given task (Cerasoli et al., 2014). It also seems plausible that the addition of game-like patterns only works in the short-term. The findings of Koivisto and Hamari (2014) suggest that the effectiveness of many gamification interventions may be short-lived due to their novelty effect, ultimately resulting in performance decreases. In fact, while the results in manuscript 2 showed that points, and especially levels and leaderboards motivated participants to generate more tags over time, all three conditions show a decreasing trend over time. Following previous findings on intrinsic versus extrinsic motivation (Deci et al., 1999; Ryan & Deci, 2000), it is likely that an intrinsically motivating implementation of gamification may help maintain longer-term behavior change, a fact which seems especially relevant for learning or health-related applications.

Why did points, levels and leaderboards not increase intrinsic motivation?

The motivational appeal of many games lies in their ability to provide players with challenges to master, hence allowing them to experience feelings of competence (Deterding, 2015; Mekler, Bopp, et al., 2014; Przybylski et al., 2010). The image annotation task described in manuscripts 2 and 3, on the other hand, could hardly be considered challenging, as participants were free to create as many tags as they wanted. Even in the control condition participants generated on average more than 50 tags, which corresponds to reaching the first two performance goals set in the level and leaderboard conditions. According to Wang, Schneider, and Valacich (2015), moderate performance targets do not motivate people to put much effort into achieving that target goal. In this case, informational feedback does not further encourage people to achieve more challenging targets and is thus less likely to satisfy their need for competence. In short, it seems plausible that points, levels and leaderboards only afford competence need satisfaction for tasks that are actually experienced as challenging.

Moreover, the game elements might not have offered enough meaningful, informational feedback to help participants judge their performance (Deci et al., 1999). Points informed

participants about how many tags they generated, and the levels/leaderboard provided performance targets to aim for (Hamari & Koivisto, 2013; Jung, Schneider, & Valacich, 2010; Von Ahn & Dabbish, 2008). But the gamified image annotation task contained no explicit indication of how many tags actually constituted a "good" performance and participants could therefore not judge whether they were competent at the image annotation task. And even in the case of points, levels or the leaderboard providing sufficient informational feedback, their visual presentation was very understated and lacked "juiciness". In contrast, many digital games provide excessive positive – juicy – feedback in the form of sounds, visuals and animations (Przybylski et al., 2010). According to Juul (2012), juiciness does not simply communicate information *[...] but also gives the player an immediate, pleasurable experience [...] enhancing the experience of feeling competent, or clever, when playing a game* (pp. 45).

Next, while the study was engaged in voluntarily, participants were encouraged to solve human computation "tasks" rather than "games". Lieberoth (2015) however, could show that explicitly labeling an activity containing game design elements as a "game" may increase people's intrinsic motivation, compared to a control task featuring no such framing or game elements. Indeed, Huotari and Hamari (2012) state that the goal of well-thought out gamification is to provide gameful experiences, and Deterding (2015) stressed that rather than (re-)structuring objects to look more like games (i.e., superficially applying game design elements to an image annotation task), a non-game context should instead be framed in such a way that people experience it as "game-like". Participants in the present study likely did not experience the image annotation task as gameful / game-like. Perhaps, intrinsic motivation might have increased, if framing of the image annotation task as a game or game-like activity were facilitated, – even if the task itself remained unchanged. However, it is important to note that this may not hold true for all people under all circumstances, as previous research found that people may be suspicious of encountering games in unexpected settings (e.g., Littleton et al., 1999).

Extrinsically motivating game elements are not forcibly bad

The above suggests that it is challenging to develop enjoyable, intrinsically motivating gamified applications. Moreover, intrinsically motivating game elements ought to be well integrated within the gamified task (Deterding, 2015; Nicholson, 2012), and might not readily generalize to other non-game contexts. However, even if extrinsic motivation may not be associated with as many benefits as intrinsic motivation, it need be stated that even extrinsically motivating game elements may have their uses. First of all, even for simple tasks that are not perceived as particular interesting, the studies described in manuscript 2 and 3 found that game elements increased performance quantity, without compromising participants' intrinsic motivation or feelings of autonomy. Similarly, in a recent study by (Preist & Jones, 2015), learning activities were interwoven as optional sidequests into a mobile game. While the game may not have made the learning material more appealing to students, it nevertheless increased the amount of time students interacted with the learning material, which subsequently led to significantly better learning outcomes. Another study by Dergousoff and Mandryk (2015) further found evidence that incorporating game elements as an extrinsic incentive encouraged people to partake in micro-experiments. Taken together, the above suggests that while game elements as extrinsic incentives may not be as enjoyable as actual games, nor provide the same benefits as intrinsically motivating interventions (Habgood & Ainsworth, 2011; Ryan & Deci, 2000), they are still effective at promoting certain behaviors, at least in the short term. Note however that in contrast to the study of Hanus and Fox (2015), all of the above studies incorporated game features in voluntary, non-mandatory settings.

The potential of games to evoke appreciation and meaning

Manuscript 4 further emphasized the role of perceived meaningfulness on motivating behavior. Interactivity significantly increased participants' appreciation for the narrative, – note that the appreciation scale (Oliver & Raney, 2011) included items such as "I found the game to be very meaningful", – which in turn was associated with higher percentage donated. These findings highlight the potential of appreciation/meaning as a crucial component of the effectiveness of games for change. It remains to be seen what game properties other than

interactivity may potentially inspire player appreciation and how this subsequently relates to behavior change. Iacovides and Cox (2015) mention narrative, gameplay, and audio as factors that helped create a meaningful and thought-provoking experience in a game illustrating the dilemmas facing health professionals. Similarly, Bartsch, Kalch, and Oliver (2014) highlight the role that moving music can play in evoking appreciation for a film.

Several scholars have also noted that for gamification to be intrinsically motivating, users should perceive both the game elements and the purpose of the service as meaningful (Deterding, 2011; McGonigal, 2011; Nicholson, 2012). McGonigal (2011) states that 'meaning' forms a type of intrinsic reward in itself, because *[...] we want to belong to and contribute to something that has lasting significance beyond our own individual lives* (pp. 50). This notion was also supported by a study conducted by Mekler, Brühlmann, Opwis, and Tuch (2013a), who found that there mere addition of a rationale and acknowledging participants' efforts significantly increased intrinsic motivation, as well as participants' perception of the task as valuable and, to a lesser extent, as personally more important. Surprisingly, this is still an empirically relatively underexplored area of research (Seaborn & Fels, 2015).

Conclusion

The experiments described in this thesis found game properties, such as feedback in the form of points, levels and leaderboards, but also mere interactivity to be effective facilitators of behavior. But while games are typically experienced as enjoyable and intrinsically motivating, this was not particularly pronounced in the present studies. In other words, only because something is labelled as a game, looks like a game, or features characteristics reminiscent of games, will not necessarily make it invariably intrinsically motivating. Finally, while enjoyment/intrinsic motivation was not tied to the increases in behavior in manuscripts 2 and 3, the experiment described in manuscript 4 could show that other experiential constructs, such as appreciation, may indeed shape behavior.

However, even though the present studies did not find enjoyment/intrinsic motivation to be a crucial factor for promoting certain behaviors, it would be foolhardy to dismiss it without further empirical inquiry. Given a positive experience, recall that users will be more likely drawn towards and continue interacting with a product or service, because they perceive it as inherently enjoyable, interesting and/or valuable and meaningful (Ryan & Deci, 2000; Zhang, 2008).

The present findings also stress the importance of not merely testing the effectiveness of game features to promote certain behaviors, but to also take user's experience as a potential mediator into account in order to gain a better understanding of the underlying psychological mechanisms. Without a solid theoretical framework through which to design studies and interpret results, it is difficult to differentiate which and why specific game elements were essential in producing successful behavioral and experiential outcomes. Finally, not only is more research necessary on how other game features relate to enjoyment and different behaviors, but potential contextual, personal and situational moderators need also be taken into account. Further investigations into whether and when certain game features afford appreciation and meaning seems to be another promising avenue for future research.

References

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., ... Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. *Psychological bulletin*, *136*(2), 151.

Antin, J., & Churchill, E. (2011). Badges in social media: A social psychological perspective. In *Gamification: Using game design elements in non-gaming contexts, a workshop at chi.*

Apter, M. J., Spirn, N., Sveback, S., & Apter, M. (1997). Motives for donating blood. *Stress and health: A reversal theory perspective*, 145–156.

Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proc. chi '11* (pp. 2689–2698).

Bartsch, A., Kalch, A., & Oliver, M. B. (2014). Moved to think: The role of emotional media experiences in stimulating reflective thoughts. *Journal of Media Psychology: Theories, Methods, and Applications*, *26*(3), 125.

Bavelier, D., Green, C. S., Pouget, A., & Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annual review of neuroscience*, *35*, 391–416.

Bernhaupt, R. (2010). User experience evaluation in entertainment. In *Evaluating user experience in games* (pp. 3–7). Springer.

Birk, M., & Mandryk, R. L. (2013). Control your game-self: effects of controller type on enjoyment, motivation, and personality in game. In *Chi '13* (pp. 685–694).

Blythe, M., & Hassenzahl, M. (2005). The semantics of fun: Differentiating enjoyable experiences. In *Funology* (pp. 91–100). Springer.

Blythe, M., Monk, A., Overbeeke, C., & Wright, P. C. (2004). *Funology: From usability to user enjoyment*. Kluwer, Dordrecht.

Bogost, I. (2011). *How to do things with video games*. University of Minnesota Press.

Bopp, J. A., Mekler, E. D., & Opwis, K. (2015). It was sad but still good: Gratifications of

emotionally moving game experiences. In *Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems* (pp. 1193–1198).

Boyle, E. A., Connolly, T. M., Hainey, T., & Boyle, J. M. (2012). Engagement in digital entertainment games: A systematic review. *Computers in Human Behavior*, *28*(3), 771–780.

Brühlmann, F., & Schmid, G.-M. (2015). How to measure the game experience?: Analysis of the factor structure of two questionnaires. In *Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems* (pp. 1181–1186).

Caillois, R. (1958). *Les jeux et les hommes*. Gallimard.

Calvillo Gámez, E. H., Cairns, P., & Cox, A. L. (2009). From the gaming experience to the wider user experience. In *Proceedings of the 23rd british hci group annual conference on people and computers: Celebrating people and technology* (pp. 520–523).

Calvillo-Gámez, E. H., Cairns, P., & Romero, P. (2012). Catching the game: multi-method approaches to understanding gaming experiences. In *Game user research workshop chi '12*.

Caroux, L., Isbister, K., Le Bigot, L., & Vibert, N. (2015). Player–video game interaction: A systematic review of current concepts. *Computers in Human Behavior*, *48*, 366–381.

Carroll, J. M. (1982). The adventure of getting to know a computer. *Computer*(11), 49–58.

Carroll, J. M., & Thomas, J. C. (1988). Fun. *ACM SIGCHI Bulletin*, *19*(3), 21–24.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, *140*(4), 980–1008.

Cohen, E. L. (2014). What makes good games go viral? the role of technology use, efficacy, emotion and enjoyment in players' decision to share a prosocial digital game. *Computers in Human Behavior*, *33*, 321–329.

Deci, E. L., Koestner, R., & Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, *125*(6), 627–668.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale:

Self-determination in personality. *Journal of research in personality*, *19*(2), 109–134.

Dergousoff, K., & Mandryk, R. L. (2015). Mobile gamification for crowdsourcing data collection: Leveraging the freemium model. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 1065–1074).

Deterding, S. (2011). Meaningful play: getting "gamification" right. *Google Tech Talk*.

Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. *Human–Computer Interaction*, *30*(3-4), 294–335.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic mindtrek conference: Envisioning future media environments* (pp. 9–15).

Engl, S., & Nacke, L. E. (2013). Contextual influences on mobile player experience–a game user experience model. *Entertainment Computing*, *4*(1), 83–91.

ESA. (2015). *Essential facts about the computer and video game industry* (Tech. Rep.). Retrieved from `http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf` (Retrieved October 21, 2015 from `http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf`.)

Fitz-Walter, Z., Wyeth, P., Tjondronegoro, D., & Scott-Parker, B. (2013). Driven to drive: Designing gamification for a learner logbook smartphone application. In *Proceedings of the first international conference on gameful design, research, and applications* (pp. 42–49).

Forde, S. F., Mekler, E. D., & Opwis, K. (2015). Informational vs. controlling gamification: A study design. In *Proceedings of the 2015 annual symposium on computer-human interaction in play* (pp. 517–522).

Francisco-Aparicio, A., Gutiérrez-Vela, F. L., Isla-Montes, J. L., & Sanchez, J. L. G. (2013). Gamification: Analysis and application. In *New trends in interaction, virtual reality and modeling* (pp. 113–126). Springer.

Gajadhar, B. J., Nap, H. H., de Kort, Y. A., & IJsselsteijn, W. A. (2010). Out of sight, out of mind: co-player effects on seniors' player experience. In *Proc. fng '10* (pp. 74–83).

Gollwitzer, M., & Melzer, A. (2012). Macbeth and the joystick: Evidence for moral cleansing after playing a violent video game. *Journal of Experimental Social Psychology*, 1356–1360.

Habgood, M. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *The Journal of the Learning Sciences*, *20*(2), 169–206.

Hamari, J., & Koivisto, J. (2013). Social motivations to use gamification: An empirical study of gamifying exercise. In *proceedings of the 21 st european conference in information systems. utrecht, netherlands.*

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152–161.

Hartmann, T., & Vorderer, P. (2010). It's okay to shoot a character: Moral disengagement in violent video games. *Journal of Communication*, *60*(1), 94–119.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & information technology*, *25*(2), 91–97.

Hecker, C. (2010). *Achievements considered harmful.* (Retrieved June 1, 2013 from `http://chrishecker.com/Achievements_Considered_Harmful`)

Huotari, K., & Hamari, J. (2012). Defining gamification: a service marketing perspective. In *Proceeding of the 16th international academic mindtrek conference* (pp. 17–22).

Iacovides, I., & Cox, A. L. (2015). Moving beyond fun: Evaluating serious experience in digital games. In *Proc. chi '15* (pp. 2245–2254).

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, *66*(9), 641–661.

Jin, S.-A. A. (2011). "i feel present. therefore, i experience flow:" a structural equation modeling approach to flow and presence in video games. *Journal of Broadcasting & Electronic Media*, *55*(1), 114–136.

Jin, S.-A. A. (2012). "toward integrative models of flow": Effects of performance, skill,

challenge, playfulness, and presence on flow in video games. *Journal of Broadcasting & Electronic Media*, *56*(2), 169–186.

Johnson, D., Nacke, L., & Wyeth, P. (2015). All about that base: differing player experiences in video game genres and the unique case of moba games. In *Chi '15* (pp. 2265–2274).

Jordan, P. W. (2002). *Designing pleasurable products: An introduction to the new human factors*. CRC press.

Jung, J., Schneider, C., & Valacich, J. (2010). Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. *Management Science*, *56*(4), 724–742.

Juul, J. (2011). *Half-real: Video games between real rules and fictional worlds*. MIT press.

Juul, J. (2012). *A casual revolution: Reinventing video games and their players*. The MIT Press.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, *35*, 179–188.

Komulainen, J., Takatalo, J., Lehtonen, M., & Nyman, G. (2008). Psychologically structured approach to user experience in games. In *Proc. nordichi '08* (pp. 487–490).

Law, E. L.-C., & van Schaik, P. (2010). Modelling user experience–an agenda for research and practice. *Interacting with computers*, *22*(5), 313–322.

Lazzaro, N. (2009). Why we play: affect and the fun of games. *Human-Computer Interaction: Designing for Diverse Users and Domains*, 155.

Lieberoth, A. (2015). Shallow gamification testing psychological effects of framing an activity as a game. *Games and Culture*, *10*(3), 229–248.

Limperos, A. M., Schmierbach, M. G., Kegerise, A. D., & Dardis, F. E. (2011). Gaming across different consoles: exploring the influence of control scheme on game-player enjoyment. *Cyberpsychology, Behavior, and Social Networking*, *14*(6), 345–350.

Littleton, K., Ashman, H., Light, P., Artis, J., Roberts, T., & Oosterwegel, A. (1999). Gender, task contexts, and children's performance on a computer-based task. *European journal of psychology of education*, *14*(1), 129–139.

Malone, T. W. (1982). Heuristics for designing enjoyable user interfaces: Lessons from

computer games. In *Proceedings of the 1982 conference on human factors in computing systems* (pp. 63–68).

Marsh, T., & Costello, B. (2012). Experience in serious games: between positive and serious experience. In *Serious games development and applications* (pp. 255–267). Springer.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin Press HC.

Mekler, E. D., & Bopp, J. A. (2015). Exploring the false affective dichotomy in games – emotions and meta-emotions. In *Chi play '15 workshop "the false affective dichotomy between positive and negative affect in game play"*.

Mekler, E. D., Bopp, J. A., Tuch, A. N., & Opwis, K. (2014). A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proceedings of the 32nd annual acm conference on human factors in computing systems* (pp. 927–936).

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013a). Disassembling gamification: the effects of points and meaning on user motivation and performance. In *Chi'13 extended abstracts on human factors in computing systems* (pp. 1137–1142).

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013b). Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of common gamification elements. In *Proceedings of the first international conference on gameful design, research, and applications* (pp. 66–73).

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2015). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*.

Mekler, E. D., Tuch, A. N., Martig, A. L., & Opwis, K. (2014). A diary study exploring game completion and player experience. In *Proceedings of the first acm sigchi annual symposium on computer-human interaction in play* (pp. 433–434).

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: the quorom statement. *The Lancet*, *354*(9193), 1896–1900.

Müller, L. J., Mekler, E. D., & Opwis, K. (2015). Facets in hci: Towards understanding

eudaimonic ux–preliminary findings. In *Proceedings of the 33rd annual acm conference extended abstracts on human factors in computing systems* (pp. 2283–2288).

Myrick, J. G., & Oliver, M. B. (2014). Laughing and crying: Mixed emotions, compassion, and the effectiveness of a youtube psa about skin cancer. *Health communication*(ahead-of-print), 1–10.

Nabi, R. L., & Krcmar, M. (2004). Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory*, *14*(4), 288–310.

Nacke, L., & Drachen, A. (2011). Towards a framework of player experience research. In *Epex '11*.

Nacke, L., & Lindley, C. A. (2008). Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proc. future play '08* (pp. 81–88).

Nakamura, J., & Csikszentmihalyi, M. (2002). The concept of flow. *Handbook of positive psychology*, 89–105.

Neys, J., & Jansz, J. (2010). Political internet games: Engaging an audience. *European Journal of Communication*, *25*(3), 227–241.

Nicholson, S. (2012). A user-centered theoretical framework for meaningful gamification. In *Games+learning+society 8.0*.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic books.

Oliver, M. B., & Bartsch, A. (2010). Appreciation as audience response: Exploring entertainment gratifications beyond hedonism. *Human Communication Research*, *36*(1), 53–81.

Oliver, M. B., Bowman, N. D., Woolley, J. K., Rogers, R., Sherrick, B. I., & Chung, M.-Y. (2015). Video games as meaningful entertainment experiences. *Psychology of Popular Media Culture*.

Oliver, M. B., & Raney, A. A. (2011). Entertainment as pleasurable and meaningful: Identifying hedonic and eudaimonic motivations for entertainment consumption. *Journal of Communication*, *61*(5), 984–1004.

Paharia, R. (2013). *Loyalty 3.0: How to revolutionize customer and employee engagement*

with big data and gamification. McGraw Hill Professional.

Peng, W., Lee, M., & Heeter, C. (2010). The effects of a serious game on role-taking and willingness to help. *Journal of Communication*, *60*(4), 723–742.

Peng, W., Lin, J., Pfeiffer, K., & Winn, B. (2012). Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology*, *15*(2), 175–196.

Poels, K., Hoogen, W. v. d., Ijsselsteijn, W., & de Kort, Y. (2012). Pleasure to play, arousal to stay: The effect of player emotions on digital game preferences and playing time. *CyberPsychology, Behavior, and Social Networking*, *15*(1), 1–6.

Preist, C., & Jones, R. (2015). The use of games as extrinisic motivation in education. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 3735–3738).

Przybylski, A. K., Rigby, C., & Ryan, R. (2010). A motivational model of video game engagement. *Review of General Psychology*, *14*(2), 154–166.

Przybylski, A. K., Ryan, R. M., & Rigby, C. S. (2009). The motivating role of violence in video games. *Personality and Social Psychology Bulletin*, *35*(2), 243–259.

Reinecke, L., Tamborini, R., Grizzard, M., Lewis, R., Eden, A., & David Bowman, N. (2012). Characterizing mood management as need satisfaction: The effects of intrinsic needs on selective exposure and mood repair. *Journal of Communication*.

Robertson, M. (2010). *Can't play, won't play.* (Retrieved June 1, 2013 from `http://hideandseek.net/2010/10/06/cant-play-wont-play/`)

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54–67.

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

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31.

Seif El-Nasr, M., Durga, S., Shiyko, M., & Sceppa, C. (2015). Unpacking adherence and engagement in pervasive health games. In *Foundations of digital games '15.*

Sherry, J. L. (2004). Flow and media enjoyment. *Communication Theory*, *14*(4), 328–347.

Shim, K. J., Srivastava, J., & Hsu, K.-W. (2011). An exploratory study of player performance, motivation, and enjoyment in massively multiplayer online role-playing games. In *Proc. passat socialcom '11* (pp. 135–140).

Shneiderman, B. (2004). Designing for fun: how can we design user interfaces to be more fun? *interactions*, *11*(5), 48–50.

Steinemann, S. T., Mekler, E. D., & Opwis, K. (2015). Increasing donating behavior through a game for change: The role of interactivity and appreciation. In *Proceedings of the 2015 annual symposium on computer-human interaction in play* (pp. 319–329).

Sweetser, P., & Wyeth, P. (2005). Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, *3*(3), 3–3.

Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, involvement, and flow in digital games. In *Evaluating user experience in games* (pp. 23–46). Springer.

Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., & Organ, A. (2010). Defining media enjoyment as the satisfaction of intrinsic needs. *Journal of Communication*, *60*(4), 758–777.

Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, *51*(8), 58–67.

Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication theory*, *14*(4), 388–408.

Wang, X., Schneider, C., & Valacich, J. S. (2015). Enhancing creativity in group collaboration: How performance targets and feedback shape perceptions and idea generation performance. *Computers in Human Behavior*, *42*, 187–195.

Weber, R., Tamborini, R., Westcott-Baker, A., & Kantor, B. (2009). Theorizing flow and media enjoyment as cognitive synchronization of attentional and reward networks. *Communication Theory*, *19*(4), 397–422.

Weibel, D., Wissmath, B., & Mast, F. W. (2011). Influence of mental imagery on spatial presence and enjoyment assessed in different types of media. *Cyberpsychology, Behavior, and Social Networking*, *14*(10), 607–612.

Wyeth, P., Johnson, D. M., & Sweetser, P. (2012). Conceptualising, operationalising and measuring the player experience in videogames. In *Extended proc. fng '12* (pp. 90–93).

Zhang, P. (2008). Motivational affordances: reasons for ict design and use. *Communications of the ACM*, *51*(11), 145–147.

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly Media.

Acknowledgements

- First and foremost, I am extremely grateful to Klaus Opwis, my thesis supervisor, for offering me the opportunity to embark on this academic journey, as well as for his encouragement, trust and support in enabling me to pursue my interest in game research.
- My co-author and master thesis supervisor Alexandre N. Tuch for teaching me the 'tools of the trade' of good research.
- I am especially grateful to my 'student' co-authors, Julia Ayumi Bopp, Florian Brühlmann, and Sharon T. Steinemann, for their amazing work, unfailing commitment and invaluable contributions to the manuscripts outlined in the present thesis. You rock!
- All my master students and research assistants – Julia Ayumi Bopp, Florian Brühlmann, Sarah Endress, Seamus F. Forde, Thomas Keller, Anja Lea Martig, Fabian Martinis, Livia J. Müler, Gian-Marco Schmid, and Sharon T. Steinemann, – for sharing my enthusiasm for research, as well as the trials and tribulations that sometimes accompany it, for countless inspiring and challenging discussions, and for their support throughout my PhD journey.
- A huge thank you to Lennart E. Nacke, for volunteering to be Second Reviewer, but more importantly, for being an incredible inspiration!
- Special thanks to Sandra P. Roth – if it weren't for you, I probably would not be working in HCI today.
- The doctoral committee for evaluating this work: Rui Mata (Chairman), Klaus Opwis (First Reviewer) and Lennart E. Nacke (Second Reviewer).
- Last but not least I want to thank my family, in particular my parents, Yolanda and Philipp Mekler-Lupolover, for their support, guidance and for always believing in me, even in times of self-doubt. I love you.

Appendix

1. Mekler, E. D., Bopp, J. A., Tuch, A. N. & Opwis, K. (2014). A Systematic Review of Quantitative Studies on the Enjoyment of Digital Entertainment Games. *CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 927-936. http://dx.doi.org/10.1145/2556288.2557078

2. Mekler, E. D., Brühlmann, F., Opwis, K. & Tuch, A. N. (2013). Do Points, Levels and Leaderboards Harm Intrinsic Motivation? An Empirical Analysis of Common Gamification Elements. *Gamification 2013: Proceedings of the First International Conference on Gameful Design, Research, and Applications*, 66-73. http://dx.doi.org/10.1145/2583008.2583017

3. Mekler, E. D., Brühlmann, F., Tuch, A. N. & Opwis, K. (in press). Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance. *Computers in Human Behavior*. http://dx.doi.org/10.1016/j.chb.2015.08.048

4. Steinemann, S. T., Mekler, E. D. & Opwis, K. (2015). Increasing Donating Behavior Through a Game for Change: The Role of Interactivity and Appreciation. *CHI PLAY '15 Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. http://dx.doi.org/10.1145/2793107.2793125

A Systematic Review of Quantitative Studies on the Enjoyment of Digital Entertainment Games

Elisa D. Mekler, Julia Ayumi Bopp, Alexandre N. Tuch, Klaus Opwis
University of Basel, Department of Psychology, Center for Cognitive Psychology & Methodology
Missionsstrasse 62a, 4055 Basel, Switzerland
elisa.mekler@unibas.ch, j.bopp@stud.unibas.ch, a.tuch@unibas.ch, klaus.opwis@unibas.ch

ABSTRACT
Enjoyment has been identified as a central component of the player experience (PX), but various, overlapping concepts within PX make it difficult to develop valid measures and a common understanding of game enjoyment. We conducted a systematic review of 87 quantitative studies, analyzing different operationalizations and measures of game enjoyment, its determinants, and how these were related to other components of PX, such as flow, presence and immersion. Results suggest that game enjoyment describes the positive cognitive and affective appraisal of the game experience, and may in part be associated with the support of player needs and values. Further, we outline that enjoyment is distinct from flow in that it may occur independently of challenge and cognitive involvement, and argue that enjoyment may be understood as the valence of the player experience. We conclude with a discussion of methodological challenges and point out opportunities for future research on game enjoyment.

Author Keywords
Digital Games; Enjoyment; Player Experience; Flow

ACM Classification Keywords
K.8.0 Personal Computing: Games; J.4 Social and Behavioral Sciences: Sociology, Psychology

INTRODUCTION
Digital games offer some of the most intense, rich and engaging experiences of all interactive products. Not only are they steadily increasing in popularity [1], but their growing importance as a medium for education, persuasion, and self-expression [12], as well as the different cultural practices that have formed around them (e.g., let's play videos, modding, game journalism, etc.) make games an immensely fascinating topic for the HCI community to study.

People play games for the experience they create [47]. However, the experience of being entertained through games is not well understood yet from a psychological perspective [83, 89]. Various concepts have been discussed, such as immersion [39], flow [59] or presence [78], making it difficult to decide upon which psychological concepts are best suited to study the player experience (PX) [8]. Meanwhile, enjoyment has been characterized as the core experience of all entertainment media [83] and games in particular [58, 77], and has previously been identified as a key motive for why people play games (for an overview refer to [13]).

Unfortunately, the definition of game enjoyment is fuzzy [77] and not well differentiated from other potentially related psychological components of PX, such as flow [58]. This raises the issue of how the experience of playing digital games can be effectively operationalized and measured [16, 89]. Moreover, many models of PX often neglect previous findings and/or are based on little empirical grounding [58, 89]. For these reasons, a common terminology for discussing and measuring PX and game enjoyment is still lacking, making it difficult to compare study findings and assess which factors contribute most to enjoyable game experiences.

Synthesizing and building upon existing knowledge on game enjoyment may improve the validity of measures, which in turn strengthens our understanding of what makes games so enjoyable. Hence, the present paper aims to provide a comprehensive review of previous quantitative research on game enjoyment. We take stock of the different conceptualizations and operationalizations of game enjoyment currently employed, in order to gain a clearer understanding of how the experience of enjoyment when playing digital games may be adequately measured. We also investigate how other psychological components of PX relate to game enjoyment, in order to further clarify the concept. Finally, we identify and categorize relevant determinants of game enjoyment and uncover opportunities for further research. We believe that by elaborating what constitutes game enjoyment, we may not only get more insights on how to properly measure the construct, but also gain a clearer understanding of the overall player experience.

RELATED WORK
Enjoyment describes an individual's positive response towards media technology and its content [83] and has become a central concept within HCI research [11], as well as one of the most frequently assessed dimensions of user experience [5]. The enjoyment of websites, for instance, has already been well conceptualized and operationalized [49], but may not be readily applicable to assess players' enjoyment of digital games, as websites generally serve some utilitarian goal, whereas the *single most important goal of games* is enjoyment [77].

CHI 2014, April 26 - May 01, 2014, Toronto, ON, Canada.

http://dx.doi.org/10.1145/2556288.2557078

Similarly, Boyle et al. [13] distinguish between different stages of the engagement process in digital games and consider the subjective experience of enjoyment key to engagement. They define the subjective experience as moment-to-moment feelings of enjoyment that players experience while gaming. In contrast, enjoyment as a motive for playing games denotes the more enduring reason to play. In other words, people are motivated to play games, because they wish to experience enjoyment.

Many models and concepts have been proposed to explain this subjective experience of enjoyment in games (for an overview see [78]). For instance, game enjoyment has often been equated with the experience of flow [77, 85]. Flow describes the subjective experience of engaging in challenging, yet manageable activities (e.g., digital games), further characterized by complete cognitive absorption, time distortion and enjoyment [61]. But some argue that flow is too restrictive, as it is only concerned with extreme experiences [39, 78] and may not adequately cover more casual experiences of enjoyment, such as gaming on the phone while waiting for the bus. And indeed, Nakamura and Csikszentmihalyi have shown that enjoyment may occur independently of the flow experience [61].

Others still, consider flow and enjoyment distinct concepts. The Game Experience Questionnaire (GEQ), for instance, defines enjoyment as a multi-dimensional construct, made up of challenge, competence, (minimal) frustration and positive affect, whereas flow constitutes part of the involvement construct, which also includes immersion and (lack of) boredom [29]. In contrast, Jennett et al. [39] define immersion as the key component of good game experiences. However, in one experiment involving a non-game task they found that high immersion could also be accompanied by anxiety and negative affect, which makes it difficult to infer whether the task was not enjoyable or whether participants simply felt tense due to the task's increasing challenge. In short, these inconsistent, often overlapping conceptualizations make it difficult to form a common definition of game enjoyment, which poses serious challenges to effectively measuring the construct, thereby hampering our understanding of PX [16, 58, 89].

METHOD

The aim of this paper is to clarify the concept of *game enjoyment*, in order to better understand how the construct may be operationalized and measured. To do so, we analyzed 87 quantitative studies measuring game enjoyment. The selection of the publications was done according to an adapted QUOROM procedure, which has previously been employed by Bargas-Avila and Hornbæk for their review on empirical studies in user experience research [5].

Data collection

Source selection. Relevant publications on PX are spread across multiple scientific journals and conferences, so we chose not to limit our search to a pre-defined set of venues, lest relevant sources be missed. Instead, the following three publication databases were searched: ACM Digital Library (DL), ScienceDirect (ScD) and ISI Web of Knowledge (WoK). A time frame of five years (2008 to 2012) was set to restrict the search and because research on PX grew considerably in the last few years [89].

Search procedure. In all three databases, the terms *game* and *enjoyment* were combined to search all publications. Although many different terms, such as *liking*, *fun* or *preference*, are used synonymously for *enjoyment*, it has been argued that the aforementioned terms do not cover the exact same meaning [11, 57]. *Liking*, for instance, reflects reactions to the media message only, whereas enjoyment reflects reactions to both the media message and the media experience (i.e., the situation and context in which the media message was received) [57]. Similarly, Blythe and Hassenzahl [11] consider *fun* and *pleasure* two distinct forms of enjoyable experiences (i.e., enjoyment in terms of distraction or absorption respectively). We argue that games may provide both absorbing experiences and be played for distraction, and thus settled for the general term *enjoyment*. The search resulted in a total of 3'036 publications (DL=977, ScD=1'716, WoK=343).

Screening criteria. Four screening criteria were defined to narrow the entries: The papers had to be about (1) digital games and (2) more specifically, about entertainment games, that is, games, whose "only" intended purpose is to generate enjoyment [58, 77]. Enjoyment is without doubt crucial for games with specific purposes, such as serious games or exergames (e.g., [53]), but – for the sake of manageability, – they are beyond the scope of this review. Similarly, papers studying pervasive and augmented reality games were not taken into account, because it has been suggested that they differ from more "traditional" digital entertainment games in several dimensions and thereby require their own criteria for enjoyment [38]. Lastly, the publications included for analysis had to be (3) original full papers and (4) written in English.

Note that these categories are not mutually exclusive – as exclusion criteria the first obvious category was chosen. The screening was done in entirety by the first author. To control for interrater effects, an independent rater performed the same categorization using 20% of the entries. Interrater reliability was found to be Kappa = 0.928 ($p < .001$), 95% CI (0.788, 1.067). A Kappa value of .8 and higher is considered *almost perfect*.

Selection criteria for inclusion in the final analysis

As we were concerned with the conceptualization and subsequent operationalization of game enjoyment, it had to be measured in some form. For this reason, papers that did not contain quantitative empirical user data (e.g., theoretical papers) were excluded. Furthermore, only publications in which the authors explicitly mention that they measured enjoyment were included. Concepts such as *liking* or *fun* were only included, if the authors explicitly equated them with enjoyment (e.g., [76]). Exceptions were made if a study employed the GEQ [29] or based its questionnaire on the GameFlow model [77], as the original authors explicitly link these to game enjoyment. Lastly, only studies about the subjective experience

of enjoyment were considered for analysis, resulting in a final sample of 68 publications.

Again, the screening was done by the first author and an independent rater. The interrater reliability was found to be Kappa = 0.873 ($p < .001$), 95% CI (0.701, 1.045). Then, all papers were coded in terms of aim of the study, measures, results, measuring times, study designs, participants, game genre, duration of gameplay, and if provided, definition of enjoyment. Note that some papers contained two or more studies, each containing measures of enjoyment. In this case, we treated those experiments as individual studies. Thus, a total of 87 studies were included in the final analysis. A table containing the collected data for each study is available online at ACM Digital Library `http://dx.doi.org/10.1145/2556288.2557078`.

RESULTS

In the following, the results of the analysis of the 87 studies are reported. Due to the current lack of a common terminology for game enjoyment research, we structured the results section into general methodological observations, followed by a more in-depth look at the measures employed. Next, we summarize the current state of research on determinants of game enjoyment and compare how enjoyment has been operationalized in relation to other PX components.

General methodological observations

Purpose of the studies. The majority of studies analyzed specific determinants of game enjoyment (n=49). Seventeen studies were about the evaluation of a particular game and 15 studies developed or compared methods for assessing game enjoyment and/or the overall player experience. The remaining studies were concerned with other aspects of gaming, such as the impact of violence on aggressiveness (e.g., [4]).

Participants. In more than half of the studies (n=44), participants' mean age ranged between 20 and 29 years, whereas in 14 studies participants were 19 years or younger. There were only 3 studies, where the mean age lay above 40 years. Moreover, 7 studies provided only the range of participants' age and 19 studies did not state any information on participants' age. Most studies (n=56) featured students as participants. Regarding gender distribution, the majority of studies had more female than male participants, apart from 10 studies where only male participants were present.

Most studies (n=52) did not assess game expertise in any form. Twenty-three studies asked players about the frequency of game-play, 2 studies let participants rate their expertise and 10 studies used other methods (e.g., Fang et al. [27] assessed the number of years participants have been playing digital games).

Games and Genre. First-person shooter games were by far the most frequently studied genre (n=22), followed by racing (n=13) and sports games (n=12, including bowling, boxing, football and tennis games). These three genres represent more than a half of the entries (n=47). Twenty-four studies did not provide any explanation on why they chose a certain game title or genre for their study.

Study setting, gameplay duration and game metrics. The time participants spent playing a game ranged between 2 to 60 minutes, with a median of 10 minutes. Apart from the online surveys, only 3 studies investigated game enjoyment outside of the lab setting [56, 63, 66]. Out of the 87 studies, 26 looked at game metrics, that is, the quantified measures of in-game data (e.g., time spent playing, actions taken by the player).

Measuring point of enjoyment. More than half the studies (n=46) assessed enjoyment after participants interacted with the game. Twelve studies measured enjoyment in-between the different gaming sessions or experimental conditions and 8 studies measured enjoyment during game-play. Thirteen studies did not provide any information on the time of measurement.

Measurement of enjoyment

Techniques used to measure enjoyment. Overall, three groups of measures were identified: Subjective self-reports in the form of questionnaires (n=82), physiological measures (n=11) and other (n=2), which includes one study which let participants rank the games they played according to their preference [9] and another study that observed the amount of player behavior that expressed enjoyment [74]. We categorized the latter along with physiological measures as a more objective means to assess game enjoyment, seeing how participants have less direct control over their responses compared to self-reports [55, 60]. Most studies employed only subjective measures (n=75), 4 studies used only objective techniques, and 8 studies employed both subjective and objective measures to assess game enjoyment.

The most frequently used standardized questionnaires were the Intrinsic Motivation Inventory (IMI; n=15), the GEQ (n=8, including iGEQ) and the self-assessment manikin scale (SAM; n=2) [66, 82]. Several studies referenced and adapted other questionnaires (e.g., [20, 75]. Except the GEQ and the scale developed by Fang et al. [27], all questionnaires were uni-dimensional. Out of 82 studies that used questionnaires to assess game enjoyment, only 31 provided psychometric properties. As illustrated in Table 1, most studies (n=29) that investigated determinants of game enjoyment utilized self-developed questionnaires.

We also took account of the individual items used to measure game enjoyment. As seen in Table 2, *enjoy* was by far the most frequently used term. Strictly speaking, items measuring the subjective experience of enjoyment were the most common (n=54; e.g., "I enjoyed playing Madden very much" [73]), the remaining items measured the "enjoyability" of the game itself (e.g., "The game was enjoyable" [69]) or some individual aspect of it (e.g., "This interface made the game play more enjoyable" [19]).

There were some differences when comparing subjective and objective measures. Overall, facial electromyography (EMG) corresponded the most to subjective reports [18, 59, 66, 82], as game enjoyment was associated with an increase in activation of the zygomaticus major, orbicularis oculis and decrease in activation of the corrugator supercilii region. Other physiological measures, such as electrodermal activity, electrocar-

Table 1. Enjoyment measures employed by study purpose

				self-developed				
Purpose	IMI	GEQ	referenced other	single item	items unknown	all listed	partly listed	objective measures
Determinants of enjoyment	15	2	6	12	6	6	5	3
Game Evaluation	–	1	2	6	–	3	1	3
Method Validation	–	3	–	–	–	7	1	4
Other	1	2	6	6	1	3	–	2
Total	16	8	14	24	7	19	7	11

Note. Data do not sum up to 87 because some studies use more than one measure or serve more than one purpose.

Table 2. The 11 most frequent terms used to measure game enjoyment.

Term	Examples and sources	N
Enjoyable, Enjoying	"How much did you enjoy the game?" [30]	77
Fun	"Playing the game was fun" [81]	30
Interesting	"I thought playing Madden was interesting" [73]	17
Good	"I felt good" [18]	17
Future play intent	"Would you like to play the game again?" [39]	16
Boring	"I thought the game was boring" [67]	13
Happy	"I feel happy when playing this game" [27]	11
Frustrating	"I felt frustrated" [60]	10
Challenge	"I felt challenged" [29]	10
Entertaining	"This game was entertaining" [80]	9
Irritating	"I felt irritable" [71]	9

diography or electroencephalography were not as clearly related to game enjoyment, even though they correlated significantly with different dimensions of the GEQ [18, 24, 45, 60].

Determinants of game enjoyment

We grouped the factors that potentially predict game enjoyment broadly into game system, player and context variables, as suggested by Nacke and Drachen [58].

Game System. Challenge was the most frequently examined factor (n=21) and was found to be an important determinant of game enjoyment, although this effect was further determined by player skills [3, 41] and motives [76]. Game outcome yielded mixed results, as some studies showed that winning a game increased enjoyment [41, 65], whereas others did not [3, 48, 75]. However, winning only by a bit was experienced as particularly enjoyable [3, 65]. Intuitive control schemes facilitated feelings of being in-control and self-efficacy [48, 72, 81], which in turn also contributed to game enjoyment. This was further reflected in some of the evaluation studies, as players enjoyed interfaces that were easy to control and allowed for best performances the most [15, 42, 51].

Quick et al. [68] found that fantasy was an important determinant of game enjoyment, and several studies confirm this notion, as narrative [64], avatar resemblance [23], as well as identification with the avatar [44, 50] and other playable characters [22] all significantly contributed to enjoyable game experiences. Sound and music also enhanced enjoyment to a certain degree [17, 60].

Of the 12 studies that examined violence in games, most found that violent game content did not or only marginally enhance enjoyment (e.g., [6]). In fact, there was some evidence that players enjoy games less, when they contain violence against humans, as they cause moral distress [30]. However, players that were already familiar with a violent game experienced less guilt and negative affect, as well as more enjoyment [34].

Player. Player types and motives were important determinants of game enjoyment [2, 21, 68, 76]. Similarly, personality traits, such as sensation seeking and self-forgetfulness were found to correlate positively with enjoyment of games [28]. Mental imagery capability had no influence on enjoyment [88]. Two studies found that game enjoyment of certain genres was dependent on the player's gender [48, 50].

Various psychological outcomes were also associated with enjoyable game experiences, such as feelings of being in-control, self-efficacy and need satisfaction [67, 79, 80]. Further, Downs and Sundar [23] found a strong positive correlation between winning, ego-enhancement and game enjoyment. Similarly, Reinecke et al. [69, 70] found that enjoyment was significantly related to mood repair and recovery experience, that is, allowing for the satisfaction of psychological needs that were previously thwarted. On the other hand, there were some indications that feelings of guilt are negatively correlated with enjoyment [34, 50].

Context. Context factors were rarely examined (n=5) and yielded mixed results. For instance, there was no clear indication that the co-presence of other players increases enjoyment (e.g., [35]). Lack of communication between players had little impact on enjoyment [9], as did location (i.e., playing at home vs. in a laboratory) [18].

Enjoyment in relation to other PX components

Twenty-four studies measured flow in some form (Abuhamdeh and Csikszentmihalyi never mention "flow" and refer to "optimal experience" instead [2, 3]), with varying definitions of how enjoyment and flow relate to each other. Some use the terms "flow" and "enjoyment" interchangeably (e.g., [40]), a notion which was further reflected by the activation of reward-related midbrain structures during flow experiences [43].

Others [21, 51] refer specifically to the GameFlow model of player enjoyment [77], whereas studies that rely on the GEQ argue that flow is a dimension of player involvement rather than enjoyment (e.g., [29]). Lastly, some state that the two constructs share similarities and that enjoyment results from the flow experience (e.g., [86]).

These inconsistent definitions stem largely from the different aspects of flow that the studies chose to focus on. Several studies associate flow with focused attention [40, 85] and the balance of skill and challenge [41, 43, 59]. However, Shim

et al. [76] found that the balance of skill and challenge – arguably a defining factor of flow – did only partially account for game enjoyment on its own. Rather, this was further dependent on players' motivations for game-play. In contrast, Limperos et al. [48] found that the experience of control – an aspect also associated with flow – was related to players' enjoyment, but not to other characteristics of flow.

Notably, several studies found a strong positive relation between presence and flow [40, 59, 87]. While presence seemed not to affect enjoyment directly [72, 86], Weibel et al. found that it increased enjoyment indirectly through flow. They concluded that being immersed in a virtual environment may facilitate cognitive involvement and subsequently, enable the experience of flow.

Similarly, Jennett et al. [39] consider immersion key to a good gaming experience, characterized by real world dissociation, as well as cognitive and emotional involvement. Unfortunately, although their measure of immersion also includes enjoyment as a subconstruct, they do not report how it relates to the aforementioned aspects. In contrast, Nacke and Lindley [59] offer an alternative definition of immersion in terms of the sensory experience that a game provides and found that immersive level design was associated with positive affect and feelings of competence, most likely due to providing a sense of spatial presence.

DISCUSSION

On the basis of the 87 studies reviewed, we compare the various operationalizations of game enjoyment and differentiate it from related constructs, in order to establish a working definition of the concept and consider its relevance for PX. Moreover, we discuss what implications our results have for future PX research.

Conceptualizing game enjoyment

As illustrated in Table 2, game enjoyment was frequently associated with fun and interest, as well as being the opposite of boredom, especially in studies that operationalized enjoyment as intrinsic motivation and hence employed the IMI (e.g., [67, 73]). This is in line with previous conceptualizations of enjoyment as a positive cognitive appraisal of media [57, 83]. Game enjoyment was also operationalized as positive affect, using both self-reports and biometrics (e.g., [66]). However, only Fang et al. [27] attempted to distinguish between affective and cognitive aspects of enjoyment. Notably, all studies that employed measures of frustration either assessed it as a separate dimension of game enjoyment, or acknowledged that it should not be considered the antipode of enjoyment [23], because even though many games provide moments of frustration, the overall experience may still be enjoyable.

Interestingly, game enjoyment was often associated with feelings of being in-control (e.g., [72]), competence (e.g., [41]) and improved mood after game-play, due to to the satisfaction of psychological needs [70]. It has been stated that need satisfaction may in part explain how game enjoyment comes into being [67, 79]. In fact, several of the reviewed studies indicated that this aspect of enjoyment is more pronounced

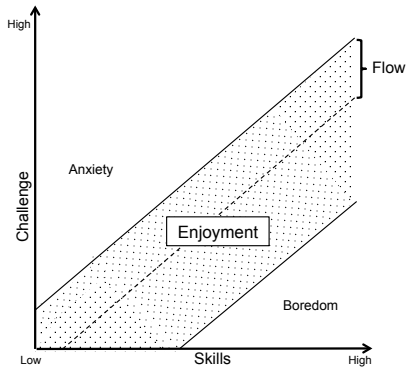

Figure 1. Flow results from the balance of game challenge and player skills. Less challenging gameplay may still be experienced as enjoyable.

for games than for non-interactive media [69, 70, 80]. Concurrently, game enjoyment was to certain degree negatively correlated with feelings of guilt [34, 50], and Gollwitzer and Melzer even operationalized it as the opposite of moral distress [30]. This was further reflected in some of the items used to measure enjoyment, such as in the instrument developed by Fang et al. [27] ("The activities in this game [...] are respectable"). It seems that – beyond 'mere' fun, – game enjoyment not only denotes a positive cognitive and affective appraisal of the gaming experience, but is in part also characterized by certain psychological outcomes, namely need satisfaction and the absence of guilt. This recalls the concept of serious fun proposed by Lazzaro [47], which suggests that players experience enjoyment when a game reflects their values (e.g., the absence of guilt) and positively affects their thoughts and feelings (e.g., through need satisfaction).

This raises some interesting research questions. For instance, whether it is more enjoyable if a game, albeit less challenging, provides ample opportunities for competence need satisfaction, or whether players experience more enjoyment when they achieve a rare moment of triumph in an extremely challenging game, and how this differs with regards to player motives. On the other hand, researchers may look into how games that were intentionally designed as non-enjoyable – such as *Torture Game 2*, which seeks to elicit disinterest by "allowing" players to commit atrocities [12], – induce guilt, and compare them to popular (i.e., enjoyable), albeit controversial games, such as *Grand Theft Auto V*, in order to clarify the relationship between need satisfaction, identification, guilt and enjoyment.

Relationship between enjoyment and other PX components

Presence did not impact game enjoyment directly, but facilitated flow experience, which itself gave rise to enjoyment [40, 86]. Arguably, game enjoyment shares many similarities with flow. But even though enjoyment is a crucial aspect of the flow experience [61], they still differ in specific ways: It seems that the experience of challenge, enjoyment and deep concentration are the main characteristics of flow (e.g., [43]), while other aspects, such as a sense of control, facilitate flow and make games more enjoyable, but do not trigger flow by

themselves [48]. Put differently, flow encompasses both enjoyment and involvement, triggered by the optimal balance of challenging gameplay and player skills [41], but players may experience enjoyment independently of flow [48, 76], if their skills exceed the challenges posed by the game [59, 61] (see Figure 1). Therefore, we conclude that limiting game enjoyment to the experience of flow would fail to account for the variety of enjoyable experiences that games may provide. Nevertheless, more empirical research is required to further probe under what circumstances game enjoyment and flow occur and whether the flow experience is characterized by particularly deep enjoyment.

This conceptualization of game enjoyment has some potential theoretical implications for PX research, namely that game enjoyment may be understood as the valence of the player experience. In contrast, the intensity of the player experience may perhaps be best represented by players' immersion, involvement or engagement, that is, the extent to which players's attention is held by gameplay challenges and the game environment (i.e., presence) [14, 39]. Put differently, the more immersive a game, the more intense the player experience, eventually culminating in cognitive absorption and time distortion. While it is beyond the scope of the present paper to establish a working definition of this concept, flow may thus be explained in terms of a very intense, yet positive experience (see Figure 2).

This framework supports previous conceptualizations of flow as effortless attention [16] and immersion as a graded experience [39]. However, more research is needed to empirically assess its validity. For instance, whether a non-intense gaming experience may be just as enjoyable as an intense one, and how the valence and the intensity of the player experience relate to each other. Moreover, in order to deepen our understanding of PX, it is necessary to consider how the intensity of the player experience can be operationalized and measured, how it is affected by challenge and presence and how this relates to need satisfaction, as well as to cognitive and affective aspects of game enjoyment. Unraveling this issue is especially relevant in light of the increasing realism of digital games, the upcoming resurgence of VR technology, such as the Oculus Rift, and for research on the negative effects of gaming [14].

Methodological challenges

Our review uncovered some methodological issues currently present in research on game enjoyment. Firstly, few studies employed standardized questionnaires, although validated and standardized questionnaires are easier to compare and may be taken with more confidence [36]. As illustrated in Table 1, many studies on the determinants of game enjoyment relied on self-developed scales, which makes it difficult to compare results and also adds to the problem of distinguishing the different components of PX. Additionally, many studies omitted item descriptions, as well as information on the reliability and validity of the scales. To facilitate interpretation of study results and advance PX research, future studies on game enjoyment would be well advised to provide as much information as possible on the measures used.

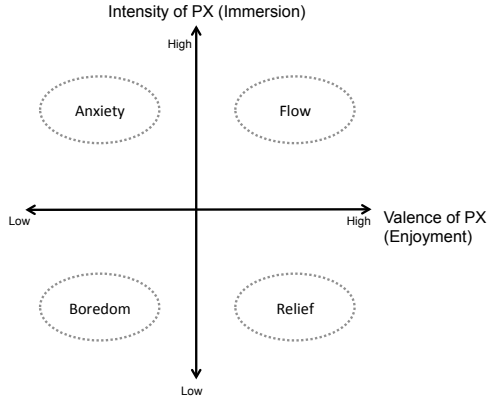

Figure 2. Valence and intensity of PX framework. Enjoyment describes the valence of the player experience, whereas immersion may denote its intensity.

Interestingly, physiological measures seem to be more common in the study of PX than in user experience research [5]. However, apart from EMG, physiological measures yielded mixed results with regards to game enjoyment. But due to the low number of studies, it would be foolhardy to dismiss physiological measures as not suitable, as they show a lot of promise, given that they allow the objective measure of player experience without interrupting the player-game interaction.

Next, the majority of studies recruited students between 20 and 29 years, although this represents only around a third of the gaming population [1]. And even though game expertise and familiarity has been found to affect enjoyment (e.g., [34]), most studies (n=52) did not take this into account. Notably, FPS games were the most frequently studied genre, even though they are mostly popular with the 21 - 29 year old male demographic [26], while other popular genres, such as casual or computer role-playing games [1] were underrepresented. To establish a more comprehensive understanding of game enjoyment, it would be beneficial to study a wide spectrum of game genres and player demographics.

Lastly, as of now, there is scant quantitative evidence on what characterizes long-term game enjoyment, as well as how players experience games before actually engaging with them [16, 58]. Seeing how some players are willing to invest hundreds of hours into a game, it would be interesting to study what aspects of the player, game system or context determine game enjoyment, how it changes over time and how this affects other components of PX.

Further research

Strikingly, although the importance of the context surrounding the player experience has been emphasized time and again [35, 58], only little attention has been paid to how contextual factors affect game enjoyment. Yet many players enjoy gaming with others [1], indicating that social aspects are indeed an important determinant of game enjoyment [47].

Moreover, Lazarro's 4 keys to fun model [47] associates many of the aforementioned determinants with particular

emotions, such as challenge and *fiero* (i.e., pride). Further research on how aspects of the game system, player and context elicit these emotions, and how this impacts overall game enjoyment, would benefit both game designers and PX researchers alike. Also, many recent games strive to provide more complex experiences than mere "fun". Seeing how previous research on non-interactive media showed that sad or frightening media content is often experienced as enjoyable [57], future studies should look into games that, for instance, inspire negative affective experiences, such as when the plot of a game demands for the death of a character, in order to study how these affect players' overall experience of enjoyment, whether they impact the cognitive and affective aspects of enjoyment differently, and how they relate to need satisfaction and player values.

Finally, it has been argued that enjoyment may facilitate media effects such as learning, aggressiveness or behavior change [57]. Not only is more research necessary to examine how game enjoyment and the intensity of the player experience are linked to negative effects, such as aggressiveness or addiction [14], but in light of the steadily increasing interest in games as a medium for expressing ideas, teaching and behavior change [12, 53], it would be especially valuable to study in what ways game enjoyment is related to these beneficial outcomes of game-play.

Limitations of the present paper

The present paper features several shortcomings. First, while our review procedure attempted to cover the gamut of quantitative studies on game enjoyment, we might have missed out on relevant studies that studied enjoyment under another term (e.g., fun). Secondly, we solely focused on studies that employed quantitative measures. For a more comprehensive understanding of enjoyable game experiences, it is necessary to also consider qualitative studies – which may in turn inform the development of quantitative measures [16], – and multimethod approaches [16], such as the biometric user studies conducted by Mirza-Babaei et al. [55]. Also, we only included studies examining "traditional" entertainment games. More research is required to find out whether the present findings may be extended to serious games and pervasive games. Thirdly, due to the nature of the review, we discussed a wide array of studies, but as a consequence, discussed individual studies only briefly. Finally, although our conclusions were founded on substantial amounts of empirical data, they in turn await empirical scrutiny.

CONCLUSION

This paper sought to investigate the concept of game enjoyment and its various operationalizations and measures. Based on the review of 87 recent quantitative studies, we found that game enjoyment is commonly understood as the positive cognitive and affective appraisal of the game experience, and is in part associated with need satisfaction and the absence of guilt. Moreover, we provide a clearer outline of how enjoyment, flow, immersion and presence differ and interrelate, and suggest that the player experience may be studied in terms of its valence, that is, how enjoyable it is. Correspondingly, we discuss the strengths and shortcomings of both objective and subjective means by which the extent of game enjoyment can be measured. More research is needed to tackle the challenge of further disentangling the different psychological components of PX, especially immersion, and elaborating upon how other relevant components, such as need satisfaction and affect relate to game enjoyment and the overall player experience.

ACKNOWLEDGEMENTS

We would like to thank the reviewers for their very helpful comments. Special thanks to Florian Brühlmann.

REFERENCES

1. Essential facts about the computer and video game industry. Tech. rep., Entertainment Software Association, 2013. Retrieved June 1, 2013 from `http://www.theesa.com/facts/pdfs/ESA_EF_2013.pdf`.
2. Abuhamdeh, S., and Csikszentmihalyi, M. Intrinsic and extrinsic motivational orientations in the competitive context: An examination of person–situation interactions. *Journal of personality 77*, 5 (2009), 1615–1635 *.
3. Abuhamdeh, S., and Csikszentmihalyi, M. The importance of challenge for the enjoyment of intrinsically motivated, goal-directed activities. *Personality and Social Psychology Bulletin 38*, 3 (2012), 317–330 *.
4. Anderson, C. A., and Carnagey, N. L. Causal effects of violent sports video games on aggression: Is it competitiveness or violent content? *Journal of Experimental Social Psychology 45*, 4 (2009), 731–739 *.
5. Bargas-Avila, J. A., and Hornbæk, K. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proc. CHI '11*, ACM (2011), 2689–2698.
6. Bastian, B., Jetten, J., and Radke, H. R. Cyber-dehumanization: Violent video game play diminishes our humanity. *Journal of Experimental Social Psychology 48*, 2 (2012), 486–491 *.
7. Bateman, S., Doucette, A., Xiao, R., Gutwin, C., Mandryk, R. L., and Cockburn, A. Effects of view, input device, and track width on video game driving. In *Proc. GI '11*, Canadian Human-Computer Communications Society (2011), 207–214 *.
8. Bernhaupt, R. User experience evaluation in entertainment. In *Evaluating User Experience in Games*. Springer, 2010, 3–7.
9. Beznosyk, A., Quax, P., Coninx, K., and Lamotte, W. The influence of cooperative game design patterns for remote play on player experience. In *Proc. APCHI '12*, ACM (2012), 11–20 *.
10. Bharambe, A., Douceur, J. R., Lorch, J. R., Moscibroda, T., Pang, J., Seshan, S., and Zhuang, X. Donnybrook: enabling large-scale, high-speed, peer-to-peer games. In *Proc. SIGCOMM '08*, vol. 38, ACM (2008), 389–400 *.

11. Blythe, M., and Hassenzahl, M. The semantics of fun: Differentiating enjoyable experiences. In *Funology*. Springer, 2005, 91–100.
12. Bogost, I. *How to Do Things with Video Games*. University of Minnesota Press, 2011.
13. Boyle, E. A., Connolly, T. M., Hainey, T., and Boyle, J. M. Engagement in digital entertainment games: A systematic review. *Computers in Human Behavior* (2011).
14. Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology 45*, 4 (2009), 624–634.
15. Browne, K., and Anand, C. An empirical evaluation of user interfaces for a mobile video game. *Entertainment Computing 3*, 1 (2012), 1–10 *.
16. Calvillo-Gámez, E. H., Cairns, P., and Romero, P. Catching the game: multi-method approaches to understanding gaming experiences. In *Game User Research Workshop CHI '12* (2012).
17. Cassidy, G., and MacDonald, R. A. The effects of music on time perception and performance of a driving game. *Scandinavian journal of psychology 51*, 6 (2010), 455–464 *.
18. Chanel, G., Kivikangas, J. M., and Ravaja, N. Physiological compliance for social gaming analysis: Cooperative versus competitive play. *Interacting with Computers 24*, 4 (2012), 306–316 *.
19. Charbonneau, E., Miller, A., Wingrave, C., and LaViola Jr, J. J. Understanding visual interfaces for the next generation of dance-based rhythm video games. In *Proc. Sandbox '09*, ACM (2009), 119–126 *.
20. De Simone, J., Verbruggen, T., Kuo, L.-H., and Mutlu, B. Is cheating a human function? the roles of presence, state hostility, and enjoyment in an unfair video game. *Computers in Human Behavior* (2012), 2351–2358 *.
21. Dias, R., and Martinho, C. Adapting content presentation and control to player personality in videogames. In *Proc. ACE '11*, ACM (2011), 18 *.
22. Dimas, J., Pereira, G., Santos, P. A., Prada, R., and Paiva, A. I'm happy if you are happy.: a model for emotional contagion in game characters. In *Proc. ACE '11*, ACM (2011), 2 *.
23. Downs, E., and Sundar, S. S. "we won" vs."they lost": Exploring ego-enhancement and self-preservation tendencies in the context of video game play. *Entertainment Computing 2*, 1 (2011), 23–28 *.
24. Drachen, A., Nacke, L. E., Yannakakis, G., and Pedersen, A. L. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proc. Sandbox '10*, ACM (2010), 49–54 *.
25. Duh, H. B.-L., Yew Yee, S. L. C., Gu, Y. X., and Chen, V. H.-H. A narrative-driven design approach for casual games with children. In *Proc. Sandbox '10*, ACM (2010), 19–24 *.
26. Elliott, L., Golub, A., Ream, G., and Dunlap, E. Video game genre as a predictor of problem use. *Cyberpsychology, Behavior, and Social Networking 15*, 3 (2012), 155–161.
27. Fang, X., Chan, S., Brzezinski, J., and Nair, C. Development of an instrument to measure enjoyment of computer game play. *Intl. Journal of Human–Computer Interaction 26*, 9 (2010), 868–886 *.
28. Fang, X., and Zhao, F. Personality and enjoyment of computer game play. *Computers in Industry 61*, 4 (2010), 342–349 *.
29. Gajadhar, B. J., Nap, H. H., de Kort, Y. A., and IJsselsteijn, W. A. Out of sight, out of mind: co-player effects on seniors' player experience. In *Proc. FnG '10*, ACM (2010), 74–83 *.
30. Gollwitzer, M., and Melzer, A. Macbeth and the joystick: Evidence for moral cleansing after playing a violent video game. *Journal of Experimental Social Psychology* (2012), 1356–1360 *.
31. Gualeni, S., Janssen, D., and Calvi, L. How psychophysiology can aid the design process of casual games: A tale of stress, facial muscles, and paper beasts. In *Proc. FDG '12*, ACM (2012), 149–155 *.
32. Gürkök, H., Hakvoort, G., Poel, M., and Nijholt, A. User expectations and experiences of a speech and thought controlled computer game. In *Proc. ACE '11*, ACM (2011), 53 *.
33. Hagelbäck, J., and Johansson, S. J. Measuring player experience on runtime dynamic difficulty scaling in an rts game. In *Proc. CIG '09*, IEEE (2009), 46–52 *.
34. Hartmann, T., and Vorderer, P. It's okay to shoot a character: Moral disengagement in violent video games. *Journal of Communication 60*, 1 (2010), 94–119 *.
35. Heeter, C., Sarkar, C. D., Palmer-Scott, B., and Zhang, S. Engineering sociability: Friendship drive, visibility, and social connection in anonymous co-located local wi-fi multiplayer online gaming. *Intl. J. of Gaming and Computer-Mediated Simulations 4*, 2 (2012), 1–18 *.
36. Hornbæk, K. Current practice in measuring usability: Challenges to usability studies and research. *Intl. journal of human-computer studies 64*, 2 (2006), 79–102.
37. Jefferson, C., Moncur, W., and Petrie, K. E. Combination: Automated generation of puzzles with constraints. In *Proc. SAC '11*, ACM (2011), 907–912 *.
38. Jegers, K. Pervasive game flow: understanding player enjoyment in pervasive gaming. *Computers in Entertainment 5*, 1 (2007), 9.

39. Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. Measuring and defining the experience of immersion in games. *Intl. journal of human-computer studies 66*, 9 (2008), 641–661 *.

40. Jin, S.-A. A. "i feel present. therefore, i experience flow:" a structural equation modeling approach to flow and presence in video games. *Journal of Broadcasting & Electronic Media 55*, 1 (2011), 114–136 *.

41. Jin, S.-A. A. "toward integrative models of flow": Effects of performance, skill, challenge, playfulness, and presence on flow in video games. *Journal of Broadcasting & Electronic Media 56*, 2 (2012), 169–186 *.

42. Jörg, S., Normoyle, A., and Safonova, A. How responsiveness affects players' perception in digital games. In *Proc. SAP '12*, ACM (2012), 33–38 *.

43. Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A., and Mathiak, K. Neural contributions to flow experience during video game playing. *Social cognitive and affective neuroscience 7*, 4 (2012), 485–495 *.

44. Klimmt, C., Hefner, D., Vorderer, P., Roth, C., and Blake, C. Identification with video game characters as automatic shift of self-perceptions. *Media Psychology 13*, 4 (2010), 323–338 *.

45. Kuikkaniemi, K., Laitinen, T., Turpeinen, M., Saari, T., Kosunen, I., and Ravaja, N. The influence of implicit and explicit biofeedback in first-person shooter games. In *Proc. CHI '10*, ACM (2010), 859–868 *.

46. Lai, J.-H., Chen, C.-L., Wu, P.-C., Kao, C.-C., and Chien, S.-Y. Tennis real play: an interactive tennis game with models from real videos. In *Proc. MM '11*, ACM (2011), 483–492 *.

47. Lazzaro, N. The four fun keys. *Game Usability: Advancing the Player Experience (K. Isbister and N. Schaffer, Eds.). Burlington: Elsevier* (2008), 315–344.

48. Limperos, A. M., Schmierbach, M. G., Kegerise, A. D., and Dardis, F. E. Gaming across different consoles: exploring the influence of control scheme on game-player enjoyment. *Cyberpsychology, Behavior, and Social Networking 14*, 6 (2011), 345–350 *.

49. Lin, A., Gregor, S., and Ewing, M. Understanding the nature of online emotional experiences: a study of enjoyment as a web experience. In *Proc. ICEC '09*, ACM (2009), 259–268.

50. Lin, S.-F. Gender differences and the effect of contextual features on game enjoyment and responses. *Cyberpsychology, Behavior, and Social Networking 13*, 5 (2010), 533–537 *.

51. Macret, M., Antle, A. N., and Pasquier, P. Can a paper-based sketching interface improve the gamer experience in strategy computer games? In *Proc. IHCI '12*, IEEE (2012), 1–6 *.

52. Maruyama, Y., Masoodian, M., and Rogers, B. A survey of japanese gamers' ratings of experience elements for different game genres. In *Proc. ACE '11*, ACM (2011), 47 *.

53. Mellecker, R., Lyons, E. J., and Baranowski, T. Disentangling fun and enjoyment in exergames using an expanded design, play, experience framework: A narrative review. *GAMES FOR HEALTH: Research, Development, and Clinical Applications* (2013).

54. Merritt, T., McGee, K., Chuah, T. L., and Ong, C. Choosing human team-mates: perceived identity as a moderator of player preference and enjoyment. In *Proc. FDG '11*, ACM (2011), 196–203 *.

55. Mirza-Babaei, P., Nacke, L. E., Gregory, J., Collins, N., and Fitzpatrick, G. How does it play better?: exploring user testing and biometric storyboards in games user research. In *Proc. CHI '13*, ACM (2013), 1499–1508.

56. Moser, C., Fuchsberger, V., and Tscheligi, M. Rapid assessment of game experiences in public settings. In *Proc. FnG '12*, ACM (2012), 73–82 *.

57. Nabi, R. L., and Krcmar, M. Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory 14*, 4 (2004), 288–310.

58. Nacke, L., and Drachen, A. Towards a framework of player experience research. In *EPEX '11* (2011).

59. Nacke, L., and Lindley, C. A. Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proc. Future Play '08*, ACM (2008), 81–88 *.

60. Nacke, L. E., Grimshaw, M. N., and Lindley, C. A. More than a feeling: Measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interacting with Computers 22*, 5 (2010), 336–343 *.

61. Nakamura, J., and Csikszentmihalyi, M. The concept of flow. *Handbook of positive psychology* (2002), 89–105.

62. Obaid, M., Han, C., and Billinghurst, M. Feed the fish: an affect-aware game. In *Proc. IE '08*, ACM (2008), 6 *.

63. Obrist, M., Igelsböck, J., Beck, E., Moser, C., Riegler, S., and Tscheligi, M. Now you need to laugh!: investigating fun in games with children. In *Proc. ACE '09*, ACM (2009), 81–88 *.

64. Park, N., Min Lee, K., Jin, S.-A. A., and Kang, S. Effects of pre-game stories on feelings of presence and evaluation of computer games. *Intl. journal of human-computer studies 68*, 11 (2010), 822–833 *.

65. Piselli, P., Claypool, M., and Doyle, J. Relating cognitive models of computer games to user evaluations of entertainment. In *Proc. FDG '09*, ACM (2009), 153–160 *.

66. Poels, K., Hoogen, W. v. d., Ijsselsteijn, W., and de Kort, Y. Pleasure to play, arousal to stay: The effect of player

emotions on digital game preferences and playing time. *CyberPsychology, Behavior, and Social Networking 15*, 1 (2012), 1–6 *.

67. Przybylski, A. K., Ryan, R. M., and Rigby, C. S. The motivating role of violence in video games. *Personality and Social Psychology Bulletin 35*, 2 (2009), 243–259 *.
68. Quick, J. M., Atkinson, R. K., and Lin, L. Empirical taxonomies of gameplay enjoyment: Personality and video game preference. *Intl. Journal of Game-Based Learning 2*, 3 (2012), 11–31 *.
69. Reinecke, L., Klatt, J., and Krämer, N. C. Entertaining media use and the satisfaction of recovery needs: Recovery outcomes associated with the use of interactive and noninteractive entertaining media. *Media Psychology 14*, 2 (2011), 192–215 *.
70. Reinecke, L., Tamborini, R., Grizzard, M., Lewis, R., Eden, A., and Bowman, N. D. Characterizing mood management as need satisfaction: The effects of intrinsic needs on selective exposure and mood repair. *Journal of Communication* (2012), 437–453 *.
71. Schild, J., LaViola, J., and Masuch, M. Understanding user experience in stereoscopic 3d games. In *Proc. CHI '12*, ACM (2012), 89–98 *.
72. Schmierbach, M., Limperos, A. M., and Woolley, J. K. Feeling the need for (personalized) speed: How natural controls and customization contribute to enjoyment of a racing game through enhanced immersion. *Cyberpsychology, Behavior, and Social Networking* (2012), 364–369 *.
73. Schmierbach, M., Xu, Q., Oeldorf-Hirsch, A., and Dardis, F. E. Electronic friend or virtual foe: Exploring the role of competitive and cooperative multiplayer video game modes in fostering enjoyment. *Media Psychology 15*, 3 (2012), 356–371 *.
74. Seif El-Nasr, M., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H., and Mah, S. Understanding and evaluating cooperative games. In *Proc. CHI '10*, ACM (2010), 253–262 *.
75. Shafer, D. M. Causes of state hostility and enjoyment in player versus player and player versus environment video games. *Journal of Communication* (2012), 719–737 *.
76. Shim, K. J., Srivastava, J., and Hsu, K.-W. An exploratory study of player performance, motivation, and enjoyment in massively multiplayer online role-playing games. In *Proc. PASSAT SocialCom '11*, IEEE (2011), 135–140 *.
77. Sweetser, P., and Wyeth, P. Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE) 3*, 3 (2005), 3–3.
78. Takatalo, J., Häkkinen, J., Kaistinen, J., and Nyman, G. Presence, involvement, and flow in digital games. In *Evaluating user experience in games*. Springer, 2010, 23–46.
79. Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., and Organ, A. Defining media enjoyment as the satisfaction of intrinsic needs. *Journal of Communication 60*, 4 (2010), 758–777 *.
80. Tamborini, R., Grizzard, M., Bowman, N. D., Reinecke, L., Lewis, R. J., and Eden, A. Media enjoyment as need satisfaction: The contribution of hedonic and nonhedonic needs. *Journal of Communication 61*, 6 (2011), 1025–1042 *.
81. Trepte, S., and Reinecke, L. The pleasures of success: game-related efficacy experiences as a mediator between player performance and game enjoyment. *Cyberpsychology, Behavior, and Social Networking 14*, 9 (2011), 555–557 *.
82. van den Hoogen, W., Poels, K., IJsselsteijn, W., and de Kort, Y. Between challenge and defeat: Repeated player-death and game enjoyment. *Media Psychology 15*, 4 (2012), 443–459 *.
83. Vorderer, P., Klimmt, C., and Ritterfeld, U. Enjoyment: At the heart of media entertainment. *Communication theory 14*, 4 (2004), 388–408.
84. Weaver, A. J., and Lewis, N. Mirrored morality: An exploration of moral choice in video games. *Cyberpsychology, Behavior, and Social Networking 15*, 11 (2012), 610–614 *.
85. Weber, R., Tamborini, R., Westcott-Baker, A., and Kantor, B. Theorizing flow and media enjoyment as cognitive synchronization of attentional and reward networks. *Communication Theory 19*, 4 (2009), 397–422 *.
86. Weibel, D., and Wissmath, B. Immersion in computer games: The role of spatial presence and flow. *Intl. Journal of Computer Games Technology* (2011), 6 *.
87. Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., and Groner, R. Playing online games against computer-vs. human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in Human Behavior 24*, 5 (2008), 2274–2291 *.
88. Weibel, D., Wissmath, B., and Mast, F. W. Influence of mental imagery on spatial presence and enjoyment assessed in different types of media. *Cyberpsychology, Behavior, and Social Networking 14*, 10 (2011), 607–612 *.
89. Wyeth, P., Johnson, D. M., and Sweetser, P. Conceptualising, operationalising and measuring the player experience in videogames. In *Extended Proc. FnG '12*, IRIT Press (2012), 90–93.
90. Yun, C., Shastri, D., Pavlidis, I., and Deng, Z. O'game, can you feel my frustration?: improving user's gaming experience via stresscam. In *Proc. CHI '09*, ACM (2009), 2195–2204 *.

* Denotes a reference among the reviewed studies.

Do Points, Levels and Leaderboards Harm Intrinsic Motivation? An Empirical Analysis of Common Gamification Elements

Elisa D. Mekler[1], **Florian Brühlmann**[1], **Klaus Opwis**[1], **Alexandre N. Tuch**[1 2]

[1]Center for Cognitive Psychology and Methodology
University of Basel
Switzerland
{elisa.mekler,klaus.opwis}@unibas.ch
florian.bruehlmann@stud.unibas.ch

[2]Dpt. of Computer Science
University of Copenhagen
Denmark
a.tuch@unibas.ch

ABSTRACT

It is heavily debated within the gamification community whether specific game elements may actually undermine users' intrinsic motivation. This online experiment examined the effects of three commonly employed game design elements – points, leaderboard, levels – on users' performance, intrinsic motivation, perceived autonomy and competence in an image annotation task. Implementation of these game elements significantly increased performance, but did not affect perceived autonomy, competence or intrinsic motivation. Our findings suggest that points, levels and leaderboards by themselves neither make nor break users' intrinsic motivation in non-game contexts. Instead, it is assumed that they act as progress indicators, guiding and enhancing user performance. While more research on the contextual factors that may potentially mediate the effects of game elements on intrinsic motivation is required, it seems that the implementation of points, levels, and leaderboards is a viable means to promote specific user behavior in non-game contexts.

Author Keywords

Gamification; gameful design; motivation; game design elements

ACM Classification Keywords

H.5.2 Informations interfaces and presentation: User interfaces

General Terms

Human Factors

INTRODUCTION

Digital games have become increasingly popular over the last few years [1], with many players investing countless hours in gaming [17, 19]. Industry professionals have taken notice of this trend and have attempted to apply games' motivational appeal to various non-gaming contexts to foster user engagement. This practice is nowadays best known under the moniker "gamification", commonly defined as *the use of game design elements in non-game contexts* [8], and has become a heavily debated subject in its own right [7, 20].

Most prominently, gamification has been commonly associated with points, levels and leaderboards [8, 27] – "the things that are least essential to games" [20], – which has irked game designers and psychologists alike. Some have cautioned against the over-reliance on such elements, as they may diminish intrinsic interest in both game- and non-game contexts, ultimately leading users to stop interacting with the application or service altogether [6, 11]. In fact, previous research in psychology provides ample evidence that different forms of rewards, feedback, and other external events can have detrimental effects on intrinsic motivation (for an overview see [4]). However, there is still a lack of empirical evidence on whether and under what circumstances these game elements may actually undermine users' intrinsic motivation [11].

While several studies have already examined how game elements, such as points, levels [9, 23] or badges [5, 10], affect user behavior, to our knowledge, none have looked yet into their impact on intrinsic motivation. This makes it difficult to infer how game elements affect users' intrinsic motivation and behavior in non-game contexts. Yet, this issue is highly relevant to gamification designers and researchers alike, as these elements have been and continue to be applied to a broad spectrum of non-game contexts with varying degrees of success and are still widely considered "[...] the heart of any gaming system" ([27], pp. 36).

The present paper aims to address the aforementioned research gaps by investigating how three of the most commonly employed game elements – points, leaderboards, levels, – affect users' behavior and intrinsic motivation. By no means do we claim that the implementation of these game elements form *good* or *bad* examples of gamification. Rather, we believe that the prevalence of these game elements in many gamified applications warrants a closer examination of their effects, to form a clearer understanding of when their imple-

Gamification'13, October 02 - 04, 2013, Stratford, ON, Canada.

http://dx.doi.org/10.1145/2583008.2583017

mentation may prove beneficial or harmful to user engagement.

RELATED WORK

Points, levels and leaderboards

Points, levels and leaderboards form three of the most basic game patterns [26, 27]. Zagal et al. categorize them as *goal metrics*, as all three are used to keep track of and provide feedback on player performance in games [26]. Due to their apparent connection to digital games and due to them being readily applicable to various non-game contexts, points, levels and leaderboards have become the poster children of gamification (e.g., [27]). Von Ahn and Dabbish, for example, consider them essential to increase enjoyment in human computation tasks [24]. However, research on the effectiveness and potential side effects of goal metrics in non-game contexts is still few and far between.

Farzan et al. notably studied the potential of a point-based incentive system (i.e., points, "status" levels and a leaderboard) to promote user activity in an enterprise social networking site [9]. Indeed, user activity initially increased, compared to the control group that was not presented with such game elements. However shortly after launch, user activity reverted back to baseline, and after the game elements were removed, user activity even dropped below to what it had been before implementation of the incentive system [23]. In fact, some users hinted at feeling driven by the leaderboard to keep up with other users [9]. This further suggests that the incentive system extrinsically motivated users to engage with the social network, as after the removal of the game elements, they were less inclined – in other words, intrinsically motivated, – to do so.

Intrinsic Motivation

Ryan and Deci differentiate two forms of motivation [21]: *Intrinsic motivation* denotes the pursuit of an activity, because it is inherently interesting or enjoyable, whereas *extrinsic motivation* is defined as doing something due to a separable outcome, such as money or deadlines. However, such extrinsic incentives have been found to reduce intrinsic motivation in various contexts [4]. In other words, a person is no longer intrinsically drawn towards engaging in an activity, because s/he is pushed to do so through external means. This is unfortunate, as intrinsic motivation is not only associated with improved psychological well-being, but also benefits the extent and quality of effort that people put into a given task, which results in enhanced performance, creativity and learning outcomes in a variety of domains [21].

External events, such as feedback and rewards may impact intrinsic motivation in different ways, depending on whether they are perceived as informational or controlling [4]. Put differently, the effects of external events on intrinsic motivation are mediated by a person's perception of how these events influence the need for competence and autonomy. Competence signifies the perceived extent of one's own actions as the cause of desired consequences in one's environment [21] and thrives when met with direct and positive feedback. But if perceived as controlling, even positive feedback may thwart people's inherent need for autonomy and hence, decrease intrinsic motivation [4], whereas feedback that is perceived as both noncontrolling and informational, supports people's need for competence and subsequently boosts their intrinsic motivation (see Figure 1).

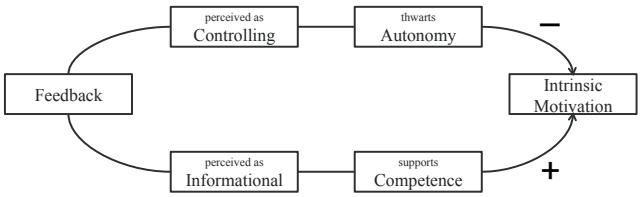

Figure 1. Feedback may be perceived as controlling or informational, thereby affecting need satisfaction and intrinsic motivation in different ways. (Figure adapted from [4, 7])

Aim of the study

Points, levels and leaderboards are commonly considered a form of extrinsic incentive in non-game contexts (e.g., [14]) and may therefore threaten users' intrinsic motivation to engage with a gamified system. However, as they often transmit a form of positive feedback, it is imaginable that they could enhance feelings of competence, and therefore increase intrinsic motivation [7, 12]. While the aforementioned studies point towards a possible detrimental effect of game elements on intrinsic motivation [9, 23], this might have been due to users feeling pressured to engage in the social networking service of their *employer* [3]. The incentive system may have simply exacerbated this, as users' activity or lack thereof was made apparent to their co-workers.

As such situational factors may mediate how points, levels and leaderboards affect intrinsic motivation [6], it is important to examine their effects in different non-game contexts, in order to gain a better understanding of when the implementation of these game elements may or may not harm users' intrinsic motivation. By investigating the effects of goal metrics on both user performance and intrinsic motivation, as well as on autonomy and competence need satisfaction, we wish to learn more about whether they are invariably perceived as controlling.

The present study aims to expand upon existing research by investigating how goal metrics affect user behavior and intrinsic motivation in an image annotation task. This non-game context was deemed suitable for several reasons. First, human computation tasks (e.g., image annotation) are often engaged in voluntarily, for fun or for pastime [2]. Thus, the threat of contextual factors acting as confounding influence (e.g., the workplace being a potentially controlling setting) is minimized. Secondly, von Ahn and Dabbish explicitly state that the aforementioned game elements increase user motivation in human computation contexts [24]. However, to our knowledge, no actual empirical evidence to back this claim exists, which provides an additional motivation for the present study.

In line with previous findings on the effects of game design elements on user behavior [5, 9, 10], we formulate the following hypothesis:

H1: *Points, levels and leaderboards significantly boost performance in the image annotation task, compared to the control condition.*

As there exists no definite previous scientific evidence on the effects of points, levels and leaderboards on intrinsic motivation in non-game contexts, the following hypotheses are based on existing research on the effects of rewards and feedback in educational settings [4, 21]. Thus, if these game elements are to be considered a form of extrinsic reward, we posit that:

H2: *Points, levels and leaderboards significantly decrease autonomy need satisfaction and intrinsic motivation, compared to the control condition.*

METHOD

To test our hypotheses, we conducted a between-subject online experiment. The independent variable were three of the most common game elements: Points vs. leaderboard vs. levels vs. control condition. The dependent variables were user performance (amount of tags, "cheating behavior", time spent on task), intrinsic motivation and satisfaction of autonomy and competence needs.

Materials

Image Tagging Platform

The image annotation task consisted of 15 abstract paintings that were taken from Machajdik and Hanbury's study on affective image classification [15]. In order to control for social factors, a single player image tagging platform was designed, loosely modeled after the one created by Wang and Yu [25]. An image was presented for 5 seconds, before flipping over and revealing the input area, where participants could enter their tags.

In the *control condition*, no game design elements were present and the right-hand side of the screen was left blank.

In the *points* condition, participants earned 100 points for each tag they entered. The current score was displayed in the upper right corner of the screen (see Figure 2 and Figure 3). Points had no further meaning, other than depicting how many tags a participant had generated. After completing all 15 images, participants were presented with their final score.

In the *leaderboard* condition, participants could compare their current score to four fictitious participants in a leaderboard on the right-hand side of the screen (see Figure 2). Participants were deliberately left unaware of the fact that fictional participants occupied the leaderboard. This static leaderboard was implemented so that all participants had the same chance to rise in ranks, as the leaderboard posititioning may have had a confounding effect on motivation otherwise [24]. To reach the lowest position on the leaderboard, participants had to generate at least ten tags. For each subsequent position, participants had to generate even more tags.

Put differently, the four competing players had a score of 1000, 3000, 6000, and 10000 respectively. These step sizes were chosen to allow participants to reach a reasonably high position on the leaderboard, but it was expected to be still reasonably challenging for participants to come up with more than 100 tags.

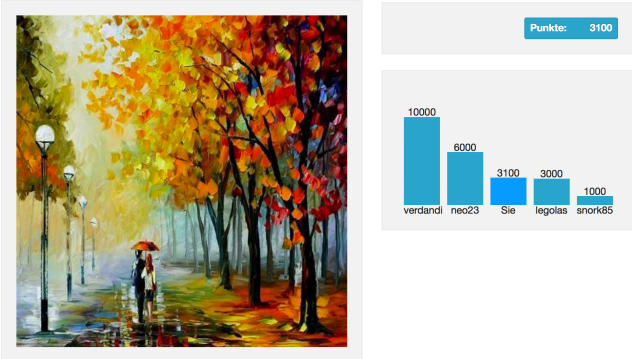

Figure 2. Screenshot of the tagging platform with points and leaderboard

In the *levels* condition, participants were presented with a vertical progress bar labeled with "next level" and the corresponding points necessary to reach the indicated level (see Figure 3). Progression to the next level mirrored the leaderboard condition, albeit without the option for (seemingly) social comparison. Whenever participants reached a score of 1000, 3000, 6000, 10000 and finally 15000, they would gain another level symbolized by an asterisk.

Measurements

Performance was measured by tracking the amount of tags generated per participant. Additionally, the amount of time participants spent on the image annotation task was tracked. According to von Ahn and Dabbish, the *throughput* (i.e., the number of tags generated per human-hour) and the overall time spent on the task determine whether the gamification of a human computation task was successful [24]. *Intrinsic motivation* (α = .95, range .93 - .96) and satisfaction of the *autonomy* (α = .68, range .65 - .74) and *competence* needs (α = .86, range .84 - .88) were assessed with the Intrinsic Motivation Inventory (IMI) [22] (7-point Likert scale, 1 = not at all true, 7 = very true).

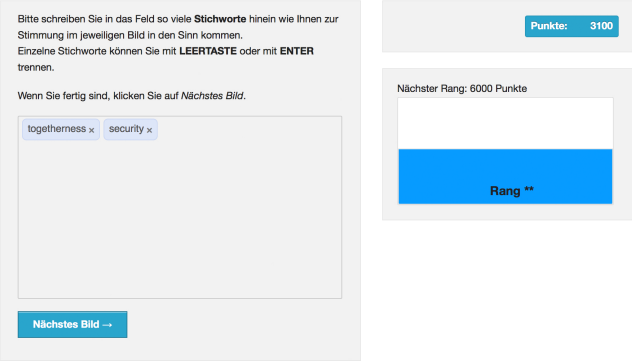

Figure 3. Screenshot of the input area with points and levels

Participants

Participants were recruited by e-mail from the university's own database, where people may sign up, if they wish to participate in studies. A total of 295 participants (93 male, 191 female, 11 not specified; mean age 32.85 years (SD = 12.33), range 17-68 years) completed the online study. Five $50 gift coupons for an online consumer electronics retailer were raffled among all participants. The raffle was deliberately chosen as incentive, because it was assumed that it would not distort the experimental effects of game elements on intrinsic motivation, due to being a form of unexpected, task-noncontingent reward. In their meta-analysis, Deci, Koestner and Ryan found that task-noncontingent rewards do not affect intrinsic motivation, as these rewards do not require doing or completing the task and hence are not perceived as controlling [4].

Procedure

Upon clicking the invitation link to the study, participants were randomly assigned to one of the four experimental conditions. Following a brief demographic questionnaire, they were then introduced to the image annotation task and informed that their tags would help improve affective image categorization. In order to isolate the effect of game elements on intrinsic motivation, special care was taken to ensure that the study description did not contain any wording that might be perceived as controlling (e.g., "you must", "you should") [22]. A test trial consisting of three images, which was the same for every condition with no game elements displayed, preceded the actual experiment.

Before starting the actual experiment, participants' attention was drawn towards the game elements, except for the control condition. Again, because the focus of this experiment was to examine the effects of points, levels and leaderboards and not the task context per se, we made sure that task instructions were worded as noncontrolling as possible, in order to avoid detrimental effects on intrinsic motivation [4]. In the points and level conditions, participants were informed that their score and level would help them estimate their contribution to the study. In the leaderboard condition, participants were told that they had the option to compare themselves to other participants.

Images were presented in random order. After completing the image annotation task, participants in the game element conditions were presented their final score, level or position on the leaderboard. Additionally, participants in the leaderboard condition had the option to enter a nickname on the leaderboard. Afterwards, all participants filled in the IMI [22] and had the option to comment on the study. Overall, participants took on average around 22 minutes to complete the study.

RESULTS

In order to investigate the effects of points, levels and leaderboard on user performance, intrinsic motivation and need satisfaction, analyses of variance (ANOVA) were calculated, unless otherwise noted. To assure homogeneity of variance, data were square-root transformed. For all statistical tests an alpha level of .05 was used.

Performance

Tag quality was determined by matching all generated tags with a German dictionary consisting of over 1.3 million entries (http://germandict.sourceforge.net/). All nonsensical tags and articles (e.g., the) were discarded from subsequent analyses. As participants could receive points even for nonsensical tags, we first checked whether conditions differed in cheating behavior (the use of nonsensical tags). A chi-square test showed that the amount of nonsensical tags differed significantly among conditions ($\chi^2 = 37.71$, $p < .001$). Descriptive analysis indicated that "cheating behavior" was less common in the level condition (4.9% nonsensical tags) than in the leaderboard (6.2%), points (7.4%) and control conditions (8.2%).

As illustrated in Figure 4 and supporting H1, participants in the game element conditions generated significantly more tags than participants in the control condition ($F(3, 291) = 11.109$, $p < .001$, $\eta_p^2 = .102$). Planned contrasts showed that participants in the points condition significantly outperformed participants in the control condition ($F(1, 153) = 10.523$, $p = .001$, $\eta_p^2 = .064$), and were in turn significantly outperformed by participants in the leaderboard ($F(1, 154) = 5.23$, $p = .024$ $\eta_p^2 = .033$) and level conditions ($F(1, 151) = 3.91$, $p = .050$ $\eta_p^2 = .026$). Performance did not differ between the leaderboard and levels conditions.

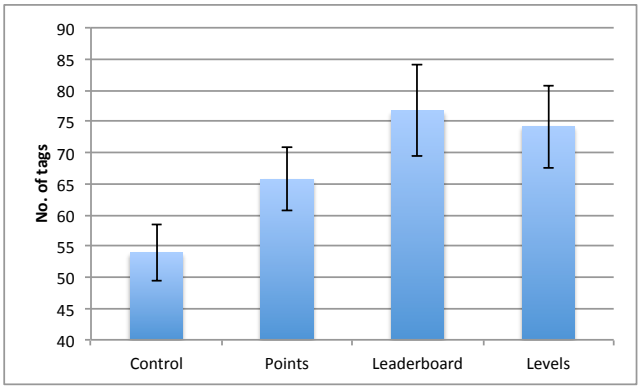

Figure 4. Average number of user-generated tags per condition. Error bars indicate 95% confidence intervals.

To check whether tagging performance changed over time for the different conditions, we calculated a 3x4 repeated measures ANOVA with time (3 blocks with 5 images each) as within-subject factor and condition as between-subject factor. A significant, significant time x condition interaction ($F(6, 582) = 2.462$, $p = .023$, $\eta_p^2 = .025$), as well as a significant main effect of time on performance ($F(6, 582) = 17.447$, $p < .001$, $\eta_p^2 = .057$) was found. Although performance over all experimental conditions decreased over time, participants' performance in the leaderboard and levels conditions declined more slowly than in the other two conditions (see Figure 5).

Conditions did not differ in the overall time participants spent on the image annotation task ($F(3, 291) = 2.015$, $p < .112$, $\eta_p^2 = .02$). But game elements significantly impacted time spent per tag ($F(3, 291) = 2.956$, $p = .033$, $\eta_p^2 = .03$).). Participants in the game element conditions spent less time per tag than

	Control ($N = 71$)		*Points* ($N = 83$)		*Leaderboard* ($N = 73$)		*Levels* ($N = 68$)	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Tags	54.24	19.120	66.01	23.527	76.86	31.234	74.31	27.175
Competence	3.81	1.161	3.85	1.124	4.13	1.179	3.91	1.017
Autonomy	5.42	1.004	5.20	1.147	5.28	1.156	5.39	.818
IM^{a}	4.87	1.454	4.54	1.554	4.79	1.609	4.90	1.445

Table 1. Means and standard deviations of dependent variables for all conditions. [a] Intrinsic motivation.

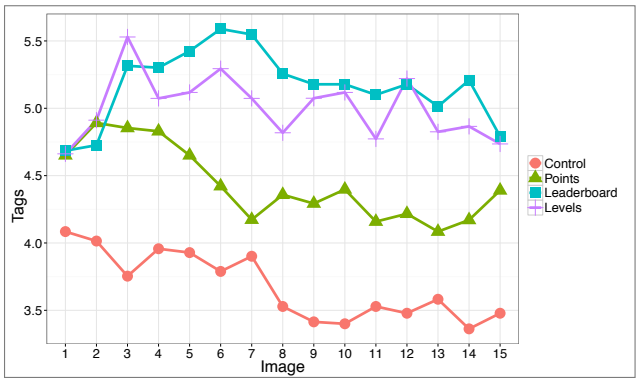

Figure 5. Average number of user-generated tags per condition over the course of 15 images.

participants in the control condition. In other words, participants in the points, leaderboard, and levels conditions generated more tags in the same amount of time as participants in the control condition.

The distribution of participants' number of tags showed some interesting differences (see Figure 6). In the control and points conditions, two (2.8%) resp. five (6.9%) participants came up with more than 100 tags, yet in the leaderboard and level conditions 17 (23.3 %) resp. 14 (20.6%) participants generated more than 100 tags each.

Intrinsic motivation & need satisfaction

Against our expectations, no significant effect of game elements on intrinsic motivation was found (p = .499). Participants reported similar levels of task enjoyment and interest, regardless of whether they received feedback in form of points, leaderboard, levels, or none at all (see Table 1). Also, no significant effects on either autonomy (p = .570) or competence (p = .340) need satisfaction were found. H2 could thus not be confirmed. Overall, participants were rather engaged in the image annotation task (see Table 1) and several commented that they enjoyed coming up with suitable tags for the paintings.

DISCUSSION

Our motivation for the present study was to experimentally assess whether and how points, leaderboards and levels affect participants' performance and intrinsic motivation in an image annotation task. In line with existing research on the potential of game elements to promote user behavior [5, 9, 10, 23], points, and especially levels and the leaderboard prompted participants to generate significantly more tags in less amount of time. Moreover, implementation of the leaderboard and leveling system inspired participants to maintain their performance for longer, compared to the points and control conditions. Arguably, the addition of game elements enhanced the quantity of tags, which confirms von Ahn and Dabbish's recommendations for increasing the throughput [24]. Against expectations, none of the game elements affected intrinsic motivation or need satisfaction in any way. All participants were motivated to similar degrees and reported intrinsic motivation did not mirror their performance in the image annotation task.

The observed performance gains may be due to the employed game elements being all forms of goal metrics [26]. By communicating how many tags have been generated, points likely formed a clear connection between participants' effort and their performance in the image annotation task [12, 24]. The levels and leaderboard perhaps further reinforced this behavior by setting explicit goals for participants to aspire to [9, 10, 12, 24]. In fact, previous research on goal setting in information systems found that users' performance increased when given a clear goal, as opposed to users who were simply asked "to do their best", even if the latter were aware of their performance [12]. Thus, it seems plausible that in the present study, participants in the points condition "did their best", but had no point of reference to judge their performance. In contrast, participants in the level and leaderboard conditions were always shown how many points they needed to reach the next level/next rank on the leaderboard, which may have prompted them to generate more tags than the other conditions. Still, even though the level and leaderboard conditions generated the most tags, their throughput eventually declined with time, which suggests that these elements did not really add to the task's "interestingness".

Yet against our expectations, participants' reported intrinsic motivation did not reflect their performance, nor was it negatively affected by either points, levels or the leaderboard. Apparently, participants did not perceive the game elements as particularly controlling, compared to the control condition. Perhaps, points, levels and leaderboards by themselves did not affect intrinsic motivation negatively, because they were not linked to other potentially pressuring external events, such as cash prizes for the best performance [4]. In contrast to the results on game elements in enterprise social network sites [9, 23], participants in the present study found themselves in a

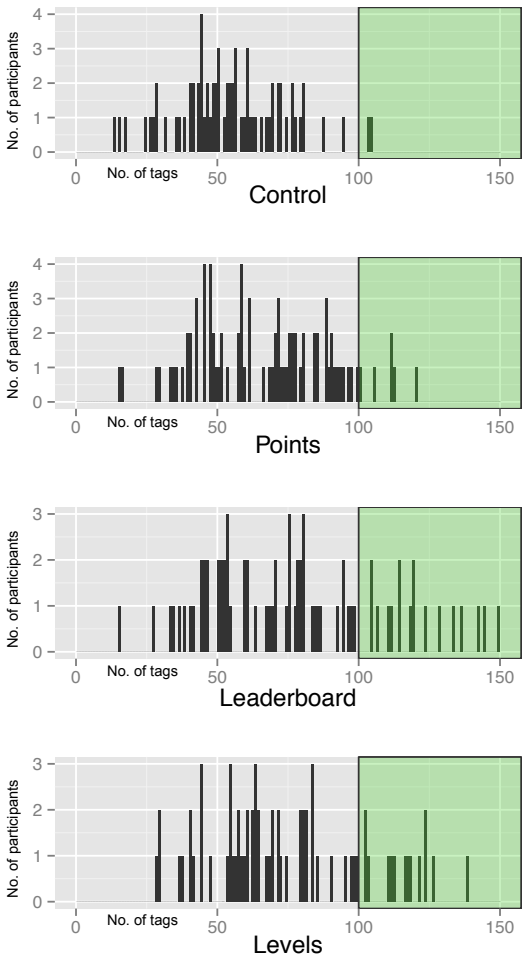

Figure 6. Distribution of the number of tags per participant for each condition.

relatively laid-back setting, as they chose to participate in the study voluntarily and, in case of the leaderboard condition, competing against people whom they did not know. Thus, we assume that due to the relative absence of controlling factors, participants' need for autonomy was not threatened.

On the other hand, game elements also did not increase feelings of competence. Hence, it may be assumed that points, levels and the leaderboard were not considered informational feedback [4]. This might be due to the goal metrics not offering enough meaningful, informational feedback to help participants judge their performance [19], such as whether a tag was fitting an image or not. On the other hand, this may not be as much due to the game elements themselves, but rather due to the nature of the task. While the image annotation task was deemed pleasant, it could hardly be considered a challenge, as participants were free to create any and as many tags as they wanted. For instance, it has been argued that the motivational appeal of games lies in their ability to provide players with new challenges to master [16], hence allowing them to experience feelings of competence [17, 19]. Seeing how many popular games rely on goal metrics [26] (e.g., *Tetris*), these game elements may perhaps only facilitate intrinsic motivation for tasks that demand learning and skill mastery.

Similarly, achievement goal theory differentiates between two types of goals, namely, mastery and performance goals [18]. While mastery goals refer to skill development and task mastery, performance goals focus on the demonstration of competence relative to normative standards. A meta-analysis on the effects of performance and mastery goals on intrinsic motivation found that informational feedback only increased intrinsic motivation for mastery goals, whereas performance goals were left unaffected [18]. Indeed, the image annotation task used in the present study bears more resemblance to a performance than a mastery goal, as participants simply had to "demonstrate" their competence in tagging paintings, relative to the norm set by levels or the leaderboard.

While the findings of the present study raise many questions on intrinsic motivation and gamification, they yield a few practical implications for the design of gamified applications. Designers may consider implementing points, levels and leaderboards in their projects as a quick and easy way to boost user performance for simple tasks, as these game elements may set clear goals for users to strive towards. While intrinsic motivation should remain unaffected by the addition of goal metrics, situational factors should still be kept in mind, as they may determine whether game elements are perceived as controlling and hence, damage intrinsic motivation. However, if the aim of gamification is to facilitate intrinsic motivation, then the mere addition of points, levels and leaderboards is not sufficient to make non-game contexts more engaging.

Limitations and further research

We provide evidence that points, levels and leaderboards are an effective means to increase short-term performance in an image annotation task. Against expectations, goal metrics did not significantly affect participants' intrinsic motivation. However, the present study featured several shortcomings that have to be addressed.

First, we only examined the short-term effects of points, levels and leaderboards. While research in psychology has found that rewards affect intrinsic motivation even for simple, short tasks [4], previous findings on the effects of game elements on user engagement have shown that it is important to also study the long-term effects of game elements, in order to better assess whether and under what circumstances they shape user behavior in the long run (e.g., [5, 10]). Also, while previous research [10, 12] and the findings of the present study suggest that gamification improves performance by means of goal-setting, it still has to be seen whether the implementation of goals without game elements (e.g., by asking participants to generate a set number of tags) affects user motivation and performance differently than when goal metrics, such as points, levels or leaderboards are employed.

Secondly, we only measured participants' self-reported intrinsic motivation. While self-reported and "free choice" measures of intrinsic motivation yielded comparable results

in previous studies [4], it would have been interesting to employ a behavioral "free choice" measure of intrinsic motivation by letting participants choose whether they want to continue engaging with the image annotation task, even after the conclusion of the experiment. Future studies should consider combining self-reported and behavioral measures of intrinsic motivation for additional methodological robustness. Also, because participants from the university's database usually engage voluntarily in studies, it is possible that they already had a minimum level of intrinsic motivation from the get-go, which might have affected the results of the present study. More research is required to investigate how users' initial eagerness to engage in a gamified application affects their subsequent motivation.

Thirdly, while von Ahn and Dabbish consider throughput and overall time spent on the main indicators of success of gamified human computation tasks [24], it would be interesting to examine more closely how goal metrics affect the *quality* of tags (i.e., how well tags fit the image). Future studies should look into how game elements may not only be implemented to increase the quantity of a certain behavior, but also how to enhance the quality of a certain behavior. This would not only yield better tags, but turn the image annotation task into a mastery goal, thereby making it potentially more challenging and interesting for users.

Finally, the results of the present study are specific to the image annotation context and the task itself scored only somewhat above average on intrinsic motivation. As rewards only threaten intrinsic motivation for activities that people find *interesting* [4], goal metrics may not have affected intrinsic motivation, due to the task not being interesting enough. Hence, our findings should only cautiously be applied to other gamified applications. It still has to be seen whether these results can be replicated for other gamified tasks. However, we believe that only by studying the effects of game elements in different non-game contexts, can we gain a more comprehensive understanding of how and when points, levels and leaderboards should be implemented to promote user behavior and intrinsic motivation.

More research is required to further investigate the role of contextual, social and situational aspects, as they at least partially determine how game elements affect intrinsic motivation and behavior [6]. While it has already been shown that goal metrics may harm intrinsic motivation in situations that may inherently be perceived as controlling [9, 23], it still has to be seen whether they also drive user behavior for other, potentially more voluntary contexts, other than human computation. For instance, it would be interesting to study the effects of goal metrics for applications focusing on sustainable living, as people usually freely choose to pursue such goals.

Also, Hamari found that badges had only a limited effect on user engagement in a peer-to-peer sharing service [10]. Thus, he argued that game elements might affect user engagement differently, depending on whether they are implemented in utilitarian or hedonic services, and stresses that use scenarios should always be kept in mind when designing gamified applications. Similarly, previous research in psychology suggests that rewards may either undermine or enhance intrinsic motivation depending on whether they are endogenous or exogenous to a given task [13]. Lastly, as previous research suggests that games are motivating due to them providing players with the possibility of expressing their choices and skills [17, 19], future studies should also compare whether "mastery" and "performance" tasks [18] are impacted differently by game elements, in order to examine whether certain non-game contexts may benefit more from goal metrics than others.

CONCLUSION

Points, levels and leaderboards are not only some of the most basic, but also three of the most commonly employed game elements in game and non-game contexts. While it has been argued that they may negatively impact users' intrinsic motivation, no actual empirical evidence exists to back this claim. The findings of the present study suggest that gamification by means of implementing points, levels and leaderboards may be an easy, viable and effective way to drive user behavior – at least in the short term. Perhaps by establishing a clear connection between user effort and performance, and by providing explicit performance goals, these game elements significantly enhanced participants' performance in an image annotation task. While significant performance gains were achieved, intrinsic motivation remained unaffected by the mere presence of points, levels and leaderboards. However, designers of gamified services should still be wary of potential social or contextual factors that may determine whether these game elements diminish intrinsic motivation. Also, as these game elements did not increase intrinsic motivation, they should not solely be relied upon to sustain long-term user engagement.

ACKNOWLEDGEMENTS

We thank Roland Hübscher and the reviewers for their very helpful suggestions on improving this paper. Alexandre N. Tuch was supported by the Swiss National Science Foundation under fellowship number PBBSP1 144196.

REFERENCES

1. Essential facts about the computer and video game industry. Tech. rep., Entertainment Software Association, 2013. Retrieved June 1, 2013 from `http://www.theesa.com/facts/pdfs/ESA_EF_2013.pdf`.
2. Antin, J., and Shaw, A. Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the us and india. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, ACM (2012), 2925–2934.
3. Deci, E., Eghrari, H., Patrick, B., and Leone, D. Facilitating internalization: The self-determination theory perspective. *Journal of personality 62*, 1 (1994), 119–142.
4. Deci, E., Koestner, R., and Ryan, R. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin 125*, 6 (1999), 627–668.

5. Denny, P. The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 763–772.
6. Deterding, S. Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using Game Design Elements in Non-Gaming Contexts, a workshop at CHI* (2011).
7. Deterding, S. Coding conduct: Games, play, and human conduct between technical artifacts and social framing, 2012. Retrieved June 1, 2013 from **http://www.slideshare.net/dings/coding-conduct-games-play-and-human-conduct-between-technical-code-and-social-framing**.
8. Deterding, S., Dixon, D., Khaled, R., and Nacke, L. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ACM (2011), 9–15.
9. Farzan, R., DiMicco, J. M., Millen, D. R., Dugan, C., Geyer, W., and Brownholtz, E. A. Results from deploying a participation incentive mechanism within the enterprise. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2008), 563–572.
10. Hamari, J. Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications* (2013). **http://dx.doi.org/10.1016/j.elerap.2013.01.004**.
11. Hecker, C. Achievements considered harmful, 2010. Retrieved June 1, 2013 from **http://chrishecker.com/Achievements_Considered_Harmful**.
12. Jung, J., Schneider, C., and Valacich, J. Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. *Management Science 56*, 4 (2010), 724–742.
13. Kruglanski, A. W. The endogenous-exogenous partition in attribution theory. *Psychological Review 82*, 6 (1975), 387.
14. Kumar, J. M., and Herger, M. *Gamification at Work: Designing Engaging Business Software*. The Interaction Design Foundation, Aarhus, Denmark, 2013.
15. Machajdik, J., and Hanbury, A. Affective image classification using features inspired by psychology and art theory. In *Proceedings of the international conference on Multimedia*, ACM (2010), 83–92.
16. Malone, T. Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 conference on Human factors in computing systems*, ACM (1982), 63–68.
17. Przybylski, A., Rigby, C., and Ryan, R. A motivational model of video game engagement. *Review of General Psychology 14*, 2 (2010), 154–166.
18. Rawsthorne, L. J., and Elliot, A. J. Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review 3*, 4 (1999), 326–344.
19. Rigby, S., and Ryan, R. *Glued to games: How video games draw us in and hold us spellbound*. ABC-CLIO, 2011.
20. Robertson, M. Can't play, won't play, 2010. Retrieved June 1, 2013 from **http://hideandseek.net/2010/10/06/cant-play-wont-play/**.
21. Ryan, R., and Deci, E. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology 25*, 1 (2000), 54–67.
22. Ryan, R., Mims, V., and Koestner, R. 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 (1983), 736–750.
23. Thom, J., Millen, D., and DiMicco, J. Removing gamification from an enterprise sns. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM (2012), 1067–1070.
24. Von Ahn, L., and Dabbish, L. Designing games with a purpose. *Communications of the ACM 51*, 8 (2008), 58–67.
25. Wang, J., and Yu, B. Labeling images with queries: A recall-based image retrieval game approach. In *Proceedings of the ACM SIGIR Workshop on Crowdsourcing for Information Retrieval* (2011).
26. Zagal, J. P., Mateas, M., Fernández-Vara, C., Hochhalter, B., and Lichti, N. Towards an ontological language for game analysis. In *Proceedings of International DiGRA Conference: Changing Views – Worlds in Play.* (2005), 3–14.
27. Zichermann, G., and Cunningham, C. *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly Media, 2011.

Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance

Elisa D. Mekler[a,*], Florian Brühlmann[a], Alexandre N. Tuch[a], Klaus Opwis[a]

[a] *Center for Cognitive Psychology and Methodology, University of Basel, Switzerland*

Abstract

Research on the effectiveness of gamification has proliferated over the last few years, but the underlying motivational mechanisms have only recently become object of empirical research. It has been suggested that when perceived as informational, gamification elements, such as points, levels and leaderboards, may afford feelings of competence and hence enhance intrinsic motivation and promote performance gains. We conducted a 2x4 online experiment that systematically examined how points, leaderboards and levels, as well as participants' goal causality orientation influence intrinsic motivation, competence and performance (tag quantity and quality) in an image annotation task. Compared to a control condition, game elements did not significantly affect competence or intrinsic motivation, irrespective of participants' causality orientation. However, participants' performance did not mirror their intrinsic motivation, as points, and especially levels and leaderboard led to a significantly higher amount of tags generated compared to the control group. These findings suggest that in this particular study context, points, levels and leaderboards functioned as extrinsic incentives, effective only for promoting performance quantity.

Keywords: Gamification, motivation, self-determination theory

2010 MSC: 00-01, 99-00

*Corresponding author. Tel.: +41 (0)612673568.

Email addresses: elisa.mekler@unibas.ch (Elisa D. Mekler), florian.bruehlmann@unibas.ch (Florian Brühlmann), a.tuch@unibas.ch (Alexandre N. Tuch), klaus.opwis@unibas.ch (Klaus Opwis)

1. Introduction

Digital games have become increasingly popular over the last few years (ESA, 2015) and empirical research in psychology has further lent evidence for their motivational appeal (e.g., Peng et al., 2012; Przybylski et al., 2010). Industry professionals have taken notice of this trend and have attempted to apply games' motivational potential to various non-gaming contexts to foster user engagement. This practice is nowadays best known under the moniker "gamification", commonly defined as *the use of game design elements in non-game contexts* (Deterding et al., 2011), and has become a heavily debated subject in its own right (Deterding, 2012; Hamari et al., 2014; Seaborn & Fels, 2015).

Most prominently, gamification has been commonly associated with points, levels and leaderboards (Hamari et al., 2014; Seaborn & Fels, 2015). While several studies have shown that the implementation of game elements may promote user behavior in various contexts (refer to Hamari et al., 2014; Seaborn & Fels, 2015, for an overview), some have cautioned against the over-reliance on such elements, as they may diminish users' intrinsic interest and hence lead them to stop engaging with the application or service altogether (Deterding, 2011; Koivisto & Hamari, 2014; Seaborn & Fels, 2015). In fact, previous research in psychology provides ample evidence that certain forms of rewards, feedback, and other external events can have detrimental effects on intrinsic motivation (for an overview see Deci et al., 1999), and a recent study suggests that the same may hold true for gamification under certain circumstances (Hanus & Fox, 2015). On the other hand, it has been argued that – provided a non-controlling setting, – the well-thought out implementation of game elements may indeed improve intrinsic motivation by satisfying users' innate psychological needs for autonomy, competence and relatedness (Deterding, 2014; Francisco-Aparicio et al., 2013; Pe-Than et al., 2014; Peng et al., 2012).

Deterding (2011, 2012) suggested that in order to gain a better understanding of the psychological mechanisms underlying gamification, the effects of *individual* game design elements on user motivation should be studied, referring to the concept of *motivational affordance*, that is, *the properties of an object that determine whether and how it [...] supports one's motivational needs* (Zhang (2008) , pp. 145). While efforts have since been undertaken to link game design elements to the satisfaction of motivational needs (Francisco-Aparicio et al., 2013; Pe-Than et al., 2014; Peng et al., 2012; Wang et al., 2015), to date only few studies attempted to experimentally investigate the effects of *individual* game elements on motivation and performance (Deterding, 2011; Hamari et al., 2014; Seaborn & Fels, 2015).

Yet, this issue is highly relevant to gamification research. Firstly, the majority of currently available gamification literature focuses predominantly on studying the effectiveness of game design elements in promoting certain behavioral outcomes (Hamari et al., 2014; Seaborn & Fels, 2015), largely ignoring the underlying psychological mechanisms that may actually account for these effects (Antin & Churchill, 2011; Deterding, 2014), (but refer to Hanus & Fox, 2015; Lieberoth, 2015; Mekler et al., 2013b, for notable exceptions). Secondly, game elements, such as points, levels and leaderboards have been and continue to be applied to a broad spectrum of non-game contexts with varying degrees of success (Hamari et al., 2014; Seaborn & Fels, 2015). But most empirical gamification studies investigate the impact of multiple game elements, making it difficult to pinpoint how and to what extent these game elements contribute to user motivation and behavior (Hamari et al., 2014; Seaborn & Fels, 2015). Moreover, most pattern-based approaches to gamification, such as the one described by Francisco-Aparicio et al. (2013), offer little guidance in deciding whether points, levels or leaderboards are suitable for a given context, or how they should be implemented (Deterding, 2015). Studying the effects of individual game elements on both behavioral outcomes and users' intrinsic motivation thus contributes to gamification research by providing a more nuanced

understanding of how particular game elements function in a given context, and may potentially benefit designers, as it allows for more informed decisions on how and under what circumstances game elements, such as points, levels or leaderboards, should or should not be implemented (Seaborn & Fels, 2015).

Based on self-determination theory (SDT), one of the most established theoretical frameworks within gamification and game motivation research (Deterding, 2015; Seaborn & Fels, 2015), the present paper aims to address the aforementioned research gaps by systematically assessing the impact of individual game design elements on both user motivation and behavior. Specifically, this study examines how points, leaderboards, and levels, – three of the most commonly employed game elements (Hamari et al., 2014; Seaborn & Fels, 2015), – affect need satisfaction, intrinsic motivation and performance in an image annotation task. Moreover, because apart from situational factors, individual differences may also account for the differing effects of gamification (Hamari et al., 2014), we additionally examine whether users' causality orientation further determines the effects of gamification.

2. Theoretical background

2.1. Intrinsic motivation, cognitive evaluation and causality orientation

Self-determination theory (SDT) differentiates two forms of motivation (Ryan & Deci, 2000) – (but refer to Vansteenkiste et al., 2010, for a more nuanced differentiation of varying types of extrinsic motivation): *Extrinsic motivation* is defined as doing something due to a separable outcome, such as pressure or "extrinsic rewards" in the form of money or verbal feedback (e.g., praise) (Deci et al., 1999), whereas *intrinsic motivation* denotes the pursuit of an activity, because it is inherently interesting or enjoyable. A recent literature review by Seaborn & Fels (2015) identified intrinsic and extrinsic motivation as some the most frequently discussed, yet rarely empirically studied constructs in gamification research. It is important to note that both extrinsic and intrinsic motivation

promote performance gains (see Cerasoli et al., 2014, for an overview), but only the latter has been associated with improved psychological well-being, enhanced creativity and learning outcomes (Ryan & Deci, 2000), as well as increases in the extent and quality of effort that people put into a given task (Cerasoli et al., 2014).

While certain extrinsic rewards have been found to reduce intrinsic motivation in various domains (Deci et al., 1999; Ryan & Deci, 2000), external rewards must not invariably undermine people's intrinsic motivation (Cerasoli et al., 2014; Deci et al., 1999). According to cognitive evaluation theory – a subtheory of SDT (Ryan & Deci, 2000; Vansteenkiste et al., 2010), – the effects of extrinsic rewards on intrinsic motivation are mediated by a person's perception of these events as informational or controlling (Deci et al., 1999; Ryan & Deci, 2000), which in turn determines how these events influence the innate psychological needs for competence and autonomy (see Figure 1). *Competence* signifies the perceived extent of one's own actions as the cause of desired consequences in one's environment (Ryan & Deci, 2000) and thrives when met with direct and positive (i.e., *informational*) feedback. However, feelings of competence will not increase intrinsic motivation unless they are accompanied by a sense of *autonomy*, that is, people must experience their behavior as self-determined rather than *controlled* by some outside source. If perceived as controlling, even positive feedback may thwart people's inherent need for autonomy and hence, decrease intrinsic motivation (Deci et al., 1999), whereas feedback that is perceived as both non-controlling and informational, supports people's need for competence and subsequently boosts their intrinsic motivation.

Finally, according to causality orientation theory (Deci & Ryan, 1985), another subtheory of SDT (Vansteenkiste et al., 2010), people differ in the extent to which they experience their actions as self-determined, which further influences whether they perceive feedback as informational or controlling (see Figure 1). Hence, a person's causality orientation acts as a moderator of the effects of feed-

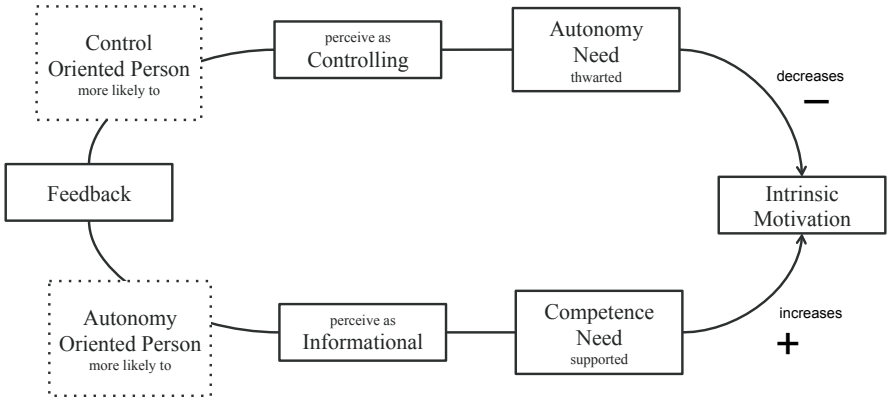

Figure 1: Feedback may be perceived as controlling or informational, thereby affecting need satisfaction and intrinsic motivation in different ways. A person's causality orientation may further moderate how feedback affects need satisfaction and intrinsic motivation. (Figure adapted from Deterding, 2012)

back on need satisfaction. Autonomy oriented individuals are more likely to act according to their own interests and values and interpret external events as informational rather than controlling (Deci & Ryan, 1985; Vansteenkiste et al., 2010), therefore experiencing more competence need satisfaction. Control oriented people, in contrast, are more likely to act due to external demands and perceive external events as pressuring and therefore experience less feelings of autonomy.

2.2. Need satisfaction and game design elements

The intrinsically motivating nature of digital games has been attributed to their potential to satisfy the psychological needs for autonomy, competence and relatedness (Przybylski et al., 2010). Satisfaction of those needs has also been found to be positively associated with the enjoyment of human computation games (Pe-Than et al., 2014). Additionally, Peng et al. (2012) compared different versions of an exergame, designed with a variety of autonomy-supportive (i.e., avatar customization) and competence-supportive game features (i.e., dy-

namic difficulty adjustment, various performance indicators). As posited by cognitive evaluation theory (Ryan & Deci, 2000), they found that need satisfaction mediated the effects of the game elements on participants' enjoyment (as measured by the Intrinsic Motivation Inventory; Ryan et al. (1983)), motivation for future play and game recommendation. However, since their study combined several game elements in each experimental condition, Peng et al. (2012) acknowledge that it is not possible to assess which and to what degree each game element accounted for the increases in need satisfaction and intrinsic motivation. As this allows only for limited carry-over to gamification and since several psychological mechanisms may underly the same game element (Antin & Shaw, 2012), Deterding (2011) stressed the need for studying how *individual* game elements influence need satisfaction in various non-game contexts.

Points, levels and leaderboards, in particular, have become the poster children of gamification (Hamari et al., 2014; Seaborn & Fels, 2015), due to their apparent connection to digital games (Zagal et al., 2005) and due to them being readily applicable to various non-game contexts. Zagal et al. (2005) categorize them as *goal metrics*, as all three are used to keep track of and provide feedback on player performance in games. According to Przybylski et al. (2010), they function as *positive*, *informational* performance feedback and thus form an important part of digital games' motivational appeal, since they afford opportunities for players to satisfy their need for competence.

Correspondingly, Francisco-Aparicio et al. (2013) list points, levels and leaderboards as a means to promote competence need satisfaction in gamified services – provided they are presented in a non-controlling manner and a voluntary setting, – but their framework still awaits empirical validation. Similarly, Jung et al. (2010) found that providing feedback (i.e., points) and clear goals (i.e., levels and leaderboards) in an idea generation task yielded significant performance gains, compared to the control group, and theorized that these results were due to the game elements satisfying people's intrinsic need for competence. Draw-

ing upon causality orientation theory, Wang et al. (2012) further found that users performed best when provided with a challenging, but attainable performance target (i.e., levels) instead of a moderate one. But while these studies claim that this may be due to performance targets promoting competence need satisfaction, it is not clear whether this was actually the case, as they did not measure feelings of competence. In fact, their conclusions actually contradict those posited by causality orientation theory, as control oriented participants outperformed autonomy oriented participants. Finally, Wang et al. (2015) also examined the effects of informational versus controlling performance feedback on competence need satisfaction. Whilst participants reported feeling less competent in the challenging condition when provided with controlling feedback, no significant differences between informational and controlling feedback for moderately challenging performance targets were found.

However, it should be noted that all of the aforementioned studies took place in form of a group collaboration setting with up to 50 students present on site per session, and it could be argued that even the informational feedback employed by Wang et al. (2015) was worded in a manner that could have been perceived as controlling by some ("You have contributed X ideas. On average, people with the same goal have contributed Y ideas by this time."). Unfortunately, since neither Jung et al. (2010) nor Wang et al. (2012, 2015) actually measured intrinsic motivation, the effects on intrinsic motivation and how these in turn relate to performance remain unclear.

2.3. Aim of study

As showcased by the literature review above, and further supported by previous systematic reviews of current gamification research (Hamari et al., 2014; Seaborn & Fels, 2015), few studies attempted to investigate the effects of *individual* game design elements on the interplay between users' need satisfaction, intrinsic motivation and behavior. While it has been suggested that points, levels and leaderboards could enhance feelings of competence, and therefore

boost intrinsic motivation and performance (Francisco-Aparicio et al., 2013; Jung et al., 2010; Przybylski et al., 2010) – provided they are encountered in a non-controlling, voluntary situation, – to date, no empirical research has been conducted on the subject.

The present study aims to expand upon existing research by investigating the effects of points, levels and leaderboards on participants' performance and motivation in an image annotation task. This non-game context was deemed suitable for three reasons. First, Von Ahn & Dabbish (2008) consider points, levels and leaderboards essential to increase enjoyment in human computation tasks (e.g., image annotation). Secondly, human computation tasks are often engaged in voluntarily, for fun and for pastime (Antin & Shaw, 2012). Thus, the threat of contextual factors acting as confounding influence (e.g., school being a potentially controlling setting (Hanus & Fox, 2015)) is minimized. Thirdly, user performance can readily be measured through the quantity and quality of generated tags.

Points, levels and leaderboards are commonly implemented in digital games to provide players with performance feedback (Przybylski et al., 2010; Zagal et al., 2005) and previous gamification research has demonstrated their effectiveness for promoting certain behaviors, since they form a clear connection between user actions and their performance (Cechanowicz et al., 2013; Denny, 2013; Hamari, 2013; Wang et al., 2012). Moreover, in contrast to points, levels/leaderboards set clear performance targets for users to aspire to, which have been associated with further performance gains (Jung et al., 2010). Hence, we formulate the following hypotheses:

H1a: *Points, levels and leaderboards significantly increase the number of tags generated in the image annotation task, compared to the plain condition.*

H1b: *Levels and leaderboards significantly increase the number of tags gen-*

erated in the image annotation task, compared to the points condition.

As previous studies on the effects of feedback on performance quality yielded mixed results (Cerasoli et al., 2014; Jung et al., 2010), no hypotheses concerning tag quality were formed.

Given the assumption that points, levels and leaderboards may afford competence need satisfaction in non-game contexts (Francisco-Aparicio et al., 2013). The effect should be more pronounced for levels and leaderboards, since they provide more performance feedback than points only (Jung et al., 2010). Lastly, need satisfaction has been found to mediate the effects of game elements on intrinsic motivation (Peng et al., 2012), hence we posit that:

H2a: *Points, levels and leaderboards significantly increase competence need satisfaction, compared to the plain condition.*

H2b: *Levels and leaderboards significantly increase competence need satisfaction, compared to the points condition.*

H3a: *Points, levels and leaderboards significantly increase intrinsic motivation, compared to the plain condition.*

H3b: *Levels and leaderboards significantly increase intrinsic motivation, compared to the points condition.*

H4: *The effect of points, levels and leaderboards on intrinsic motivation predicted in H3a and H3b is mediated by competence need satisfaction.*

Finally, it has been suggested that causality orientation may moderate the effects of feedback on user performance (Wang et al., 2012) and that autonomy oriented individuals tend to perceive feedback as more informational than

control oriented individuals (Deci & Ryan, 1985; Vansteenkiste et al., 2010), therefore experiencing more competence need satisfaction. Hence, we expect that:

H5: *Autonomy oriented study participants generate significantly more tags than control oriented participants, when in the points, level and leaderboard conditions.*

H6: *Autonomy oriented study participants report significantly more competence need satisfaction in the points, level and leaderboard conditions, compared to control oriented participants.*

H7: *Autonomy oriented study participants report significantly more intrinsic motivation in the points, level and leaderboard conditions, compared to control oriented participants.*

3. Method

To test our hypotheses, we conducted a 4 x 2 between-subject online experiment. The independent variable were three of the most common game elements: points vs. leaderboard vs. levels vs. plain condition without any game elements, as well as participants' causality orientation (autonomy vs. control oriented). The dependent variables were user performance (amount of tags and tag quality), intrinsic motivation and satisfaction of autonomy and competence needs.

3.1. Materials

3.1.1. Image Tagging Platform

The image annotation task consisted of 15 abstract paintings that were taken from Machajdik & Hanbury (2010)'s study on affective image classification. In order to control for social factors, a single player image tagging platform was designed. An image was presented for 5 seconds, before flipping over and revealing the input area, where participants could enter their tags. It has to be noted

that we did not aim to create the most efficient or effective image annotation platform, but rather focused on developing a system that would allow us to systematically study various game design elements in a controlled manner. In a previous study employing the same platform (Mekler et al., 2013a), participants rated the image annotation task as rather intrinsically motivating (mean intrinsic motivation = 4.72 on a 7-point Likert scale).

In the *plain* condition, no game design elements were present and the right-hand side of the screen was left blank.

In the *points* condition, participants earned 100 points for each tag they entered. The current score was displayed in the upper right corner of the screen (see Figure 2 and Figure 3). Points had no further meaning, other than depicting how many tags a participant had generated. After completing all 15 images, participants were presented with their final score.

In the *leaderboard* condition, participants could compare their current score to four fictitious participants in a leaderboard on the right-hand side of the screen (see Figure 2). Participants were deliberately left unaware of the fact that fictional participants occupied the leaderboard. This static leaderboard was implemented so that all participants had the same chance to rise in ranks, as the leaderboard posititioning may have had a confounding effect on motivation otherwise (Von Ahn & Dabbish, 2008). To reach the lowest position on the leaderboard, participants had to generate at least ten tags. For each subsequent position, participants had to generate a specific number of tags: 30 tags for 3rd position, 60 tags for 2nd position, and 100 tags for top rank. Put differently, the four competing players had a score of 1000, 3000, 6000, and 10000 respectively. These step sizes were chosen to allow participants to reach a reasonably high position on the leaderboard, but it was expected to be still reasonably challenging for participants to come up with more than 100 tags.

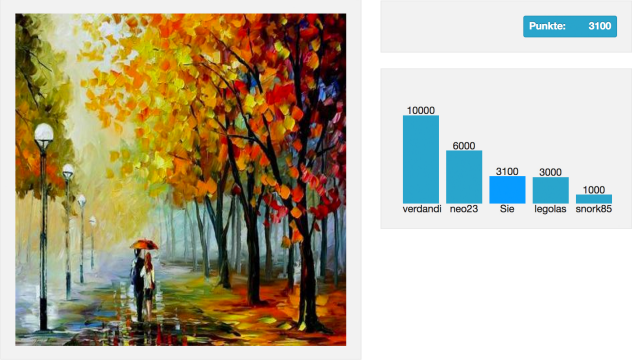

Figure 2: Screenshot of the tagging platform with points and leaderboard

In the *levels* condition, participants were presented with a vertical progress bar labeled with "next level" and the corresponding points necessary to reach the indicated level (see Figure 3). Progression to the next level mirrored the leaderboard condition, albeit without the option for (seemingly) social comparison. Whenever participants reached a score of 1000, 3000, 6000 and finally 10000, they would gain another level symbolized by an asterisk.

3.1.2. Measurements

Performance was measured by tracking the amount of tags generated per participant. In order to assess *tag quality*, the dataset was cleaned up using Open Refine (`www.openrefine.org`) to remove typos and spelling mistakes. All nonsensical tags and articles (e.g., the) were discarded from subsequent analyses. Two independent evaluators rated how well the remaining tags fit the images, following the instructions used in the study. They were asked to rate the images from 1 to 3 (1 = tag does not represent the emotional content of the image, 3 = tag reflects the emotional content of the image). The inter-rater reliability was found to be Kappa =.825. A Kappa value of .8 and higher is considered *almost perfect*. Based on the ratings of the two evaluators mean tag quality scores were calculated for each study participant.

Intrinsic motivation ($\alpha = .95$) and satisfaction of the need for *competence* ($\alpha = .86$) were assessed with the Intrinsic Motivation Inventory (IMI; Ryan et al. (1983)) All IMI items were rated on a 7-point Likert scale, ranging from 1 (not at all true) to 7 (very true). Since previous research found that game elements may impair intrinsic motivation (Hanus & Fox, 2015), we also included measures of *autonomy* need satisfaction ($\alpha = .68$), in order to find out whether these game elements are invariably perceived as controlling.

Lastly, *general causality orientation* was measured with the 12-vignette General Causality Orientations Scale (GCOS; Deci & Ryan (1985)), which has previously also been used by Wang et al. (2012). The complete GCOS questionnaire can be downloaded from `http://www.selfdeterminationtheory.org/general-causality-orientations-scale/`. Each vignette describes an incident and lists two ways of responding to it, whereupon participants state how likely it is that they would respond in such a way (7-point Likert scale, 1 = very unlikely, 7 = very likely). For instance:

You have been offered a new position in a company where you have worked for some time. The first question that is likely to come to mind is:

a) Will I make more at this position?

b) I wonder if the new work will be interesting.

Answer *a)* depicts the control-oriented response, whereas *b)* illustrates the autonomy-oriented response to the event. As previously described by Wang et al. (2012), each participant's causality orientation was identified by standardizing their total score on the two scales (autonomy vs. control orientation). Participants were then classified as autonomy-oriented when the z value of the autonomy scale was higher than the z value of the control scale and vice versa.

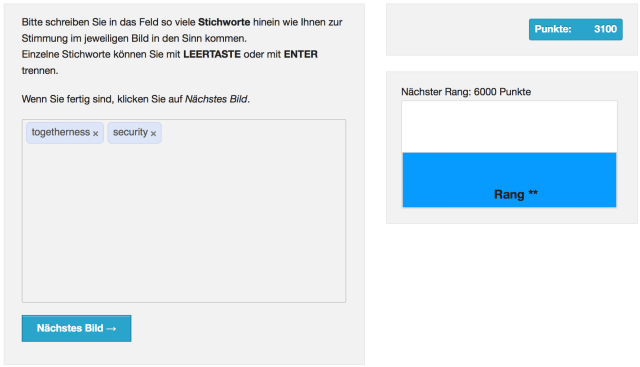

Figure 3: Screenshot of the input area with points and levels

3.2. Participants

Participants were recruited by e-mail from the university's own database, where people may sign up, if they wish to participate in studies. A total of 273 participants (84 male, 178 female, 11 not specified; mean age 32.80 years (SD = 12.21), range 17-68 years) completed the online study. Five $50 gift coupons for an online consumer electronics retailer were raffled among all participants. The raffle was deliberately chosen as incentive, because it was assumed that it would not distort the experimental effects of game elements on intrinsic motivation, due to being a form of unexpected, task-noncontingent reward. In their meta-analysis, Deci et al. (1999) found that task-noncontingent rewards do not affect intrinsic motivation, as these rewards do not require doing or completing the task and hence are not perceived as controlling. Concerning causality orientation, 130 participants were identified as autonomy oriented, whereas 143 participants were control oriented. Participants of both causality orientations were equally distributed among experimental conditions (see Table 1). Similarly, men and women were evenly distributed among the different experimental conditions and causality orientations.

			TagQuantity		*TagQuality*		*Autonomy*		*Competence*		*IM*[a]	
Condition	*CO*[b]	*N*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Plain	Autonomy	37	56.76	17.103	1.45	.252	5.42	1.08	3.81	1.16	5.07	1.41
	Control	31	51.29	21.704	1.43	.206	5.46	.886	3.97	1.14	4.55	1.53
Points	Autonomy	34	66.50	31.111	1.41	.207	5.39	1.27	3.87	1.34	4.83	1.82
	Control	41	62.61	24.461	1.37	.248	4.99	.906	3.77	.954	4.28	1.37
Leaderboard	Autonomy	29	81.97	32.109	1.45	.194	5.21	.891	4.07	1.19	4.94	1.63
	Control	36	70.53	28.360	1.40	.244	5.38	1.26	4.04	1.18	4.62	1.58
Levels	Autonomy	30	75.93	22.770	1.45	.225	5.40	.838	3.94	1.03	5.01	1.59
	Control	35	70.71	29.577	1.43	.206	5.34	.801	3.79	.890	4.70	1.30

Table 1: Descriptive statistics for all experimental conditions. [a] IM = Intrinsic Motivation.
[b] CO = Causality orientation.

3.3. Procedure

Upon clicking the invitation link to the study, participants were randomly assigned to one of the four experimental conditions. Following a brief demographic questionnaire, they were then introduced to the image annotation task and informed that their tags would help improve affective image categorization. In order to isolate the effect of game elements on intrinsic motivation, special care was taken to ensure that the study description did not contain any wording that might be perceived as controlling, such as "you must" and "you should" (Ryan et al., 1983). A test trial consisting of three images, which was the same for every condition with no game elements displayed, preceded the actual experiment.

Before starting the actual experiment, participants' attention was drawn towards the game elements, except for the control condition. Again, because the focus of this experiment was to examine the effects of points, levels and leaderboards and not the task context per se, we made sure that task instructions were worded as non-controlling as possible, in order to avoid detrimental effects on intrinsic motivation (Deci et al., 1999). In the points and level conditions, participants were informed that their score and level would help them estimate their contribution to the study. In the leaderboard condition, participants were told that they had the option to compare themselves to other participants.

Images were presented in random order. After completing the image annotation task, participants in the game element conditions were presented their final score, level or position on the leaderboard. Additionally, participants in the leaderboard condition had the option to enter a nickname on the leaderboard. Afterwards, all participants filled in the IMI (Ryan et al., 1983) and the GCOS (Deci & Ryan, 1985) and had the option to comment on the study. Overall, participants took on average around 22 minutes to complete the study.

4. Results

In order to investigate the effects of points, levels and leaderboard and goal causality orientation on user performance, intrinsic motivation and need satisfaction, analyses of variance (ANOVA) were calculated, unless otherwise noted. To assure homogeneity of variance, data were square-root transformed. For all statistical tests an alpha level of .05 was used. Before any further analysis was conducted, we controlled the data for age and gender effects, but did not find any significant differences. Overall, tag quantity was negatively correlated with tag quality (see Table 2), as well as slightly positively correlated with intrinsic motivation and competence need satisfaction. Intrinsic motivation was positively correlated with both autonomy and competence need satisfaction. An overview of our hypotheses can be found in Table 4.

	Tag Quality	Autonomy	Competence	Intrinsic Motivation
Tag Quantity	-.38**	.01	.18**	.14*
Tag Quality		-.02	-.10	-.05
Autonomy			.31**	.44**
Competence				.41**

Table 2: Pearson's Correlation for dependent variables over all conditions. * Significant at $p < .05$. ** Significant at $p < .01$.

4.1. Tag Quantity

Firstly, while no significant game elements x causality orientation interaction was found, results yielded a significant main effect of game elements on the amount of tags generated. As illustrated in Figure 4 and supporting H1a, participants in the game element conditions generated significantly more tags than participants in the control condition ($F(3, 265) = 10.09$, $p < .001$, $\eta_p^2 = .103$). Confirming H1b, planned contrasts showed that participants in the points condition significantly outperformed participants in the control condition ($t(141) = 2.613$, $p = .01$, $d = .44$), and were in turn significantly outperformed by participants in the leaderboard ($t(138) = 2.299$, $p = .023$, $d = .39$) and level conditions ($t(138) = 2.032$, $p = .044$, $d = .35$). Performance did not differ between the leaderboard and levels conditions. Additionally, a significant main effect of causality orientation on tag quantity was found ($F(1, 265) = 5.31$, $p = .022$, $\eta_p^2 = .02$). Autonomy oriented participants generated on average more tags than control oriented participants (see Table 1 for descriptive statistics), partially supporting H5.

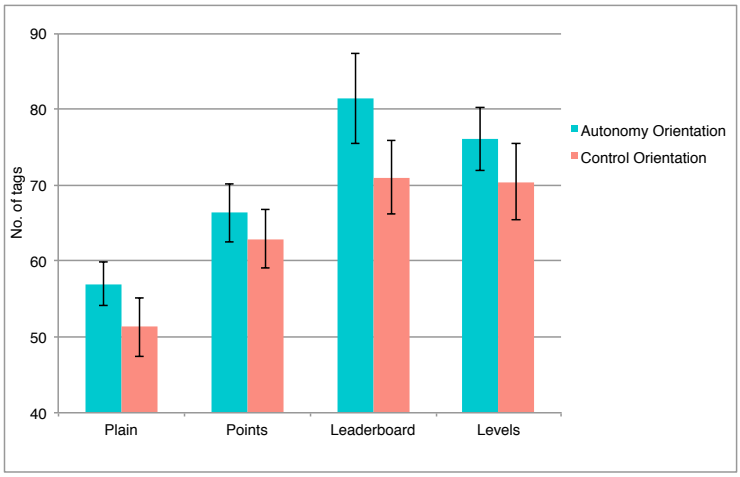

Figure 4: Average number of user-generated tags per condition. Error bars are indicate standard error of the mean.

4.2. Tag Quality

No significant interaction or main effects for game elements and causality orientation on tag quality were found. As depicted in Figure 5 participants in the the game element conditions did not generate tags of significantly higher quality ($F(3, 265) = 0.727$, $p = .537$, $\eta_p^2 = .008$). Neither a significant main effect for causality orientation on tag quality ($F(1, 265) = 1.188$, $p = .277$, $\eta_p^2 = .004$) nor a significant interaction between conditions and causality orientation were observed ($F(3, 265) = 0.049$, $p = .986$, $\eta_p^2 = .001$). Mean tag quality did not differ significantly between the different game element conditions, nor did it depend on participants' motivational orientation. Overall, mean tag quality was rather low (see Table 1).

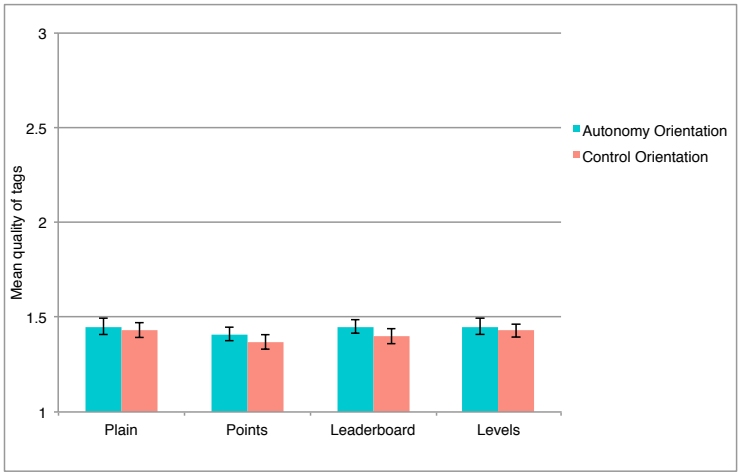

Figure 5: Average quality rating of user-generated tags per condition. Error bars indicate standard error of the mean.

As reported earlier, tag quality was significantly negatively correlated with tag quantity. To see if game elements predicted the quality of tags when controlling for tag quantity, a multiple linear regression was conducted. Therefore, the four conditions were dummy-coded into three variables and included in the model. Tag quantity predicted tag quality significantly, but game elements and the interaction of game elements x tag quantity did not (see Table 3).

Variable	B	SE(B)	β	t	Sig. (p)
(Constant)	1.978	0.135		14.637	$< .001$
Tag quantity	-0.075	0.02	-0.540	-4.044	$< .001$
Condition points	-0.164	0.18	-0.327	-0.887	.376
Condition levels	-0.045	0.19	-0.085	-0.240	.810
Condition leaderboard	-0.161	0.20	-0.307	-0.824	.411
Tag quantity x condition points	0.020	0.02	0.327	0.841	.401
Tag quantity x condition levels	0.014	0.02	0.240	0.614	.540
Tag quantity x condition leaderboard	0.029	0.02	0.475	1.176	.241

Note: Tag quantity was square-root transformed for analysis.

Table 3: Summary of multiple linear regression analysis for tag quality (N = 273) with $R^2 = 0.17$, $F(7, 265) = 7.866$, $p < .001$. The four experimental conditions were dummy-coded into three dichotomous variables.

4.3. Intrinsic motivation & need satisfaction

Against our expectations, neither a significant main effect of game elements ($p = .702$) nor a significant game elements x causality orientation interaction ($p = .927$) on intrinsic motivation emerged. In contrast to H3a and H3b, participants were motivated to similar degrees, regardless of whether they received feedback in form of points, leaderboard, levels, or none at all. Because of the lack of significant differences, it was not possible to assess competence need satisfaction as mediating the effects of the game elements on intrinsic motivation. Still, there was a significant main effect of causality orientation on intrinsic motivation ($F(7, 265) = 6.903$, $p = .009$, $\eta_p^2 = .03$). Autonomy oriented participants were more intrinsically motivated to engage in the image annotation task than control oriented participants, irrespective of the experimental condition. Lastly, no significant interaction or main effects on either autonomy ($p = .273$ - $.585$) or competence ($p = .656$ - $.861$) need satisfaction were found. H2a, H2b and H6 could thus not be confirmed.

Hypothesis	Confirmed?
H1a: *Points, levels and leaderboards significantly increase the number of tags generated in the image annotation task, compared to the plain condition.*	Yes
H1b: *Levels and leaderboards significantly increase the number of tags generated in the image annotation task, compared to the points condition.*	Yes
H2: *Points, levels and leaderboards significantly increase competence need satisfaction, compared to the plain condition.*	No
H3: *Points, levels and leaderboards significantly increase intrinsic motivation, compared to the plain condition.*	No
H4: *The effect of points, levels and leaderboards on intrinsic motivation is mediated by competence need satisfaction.*	No
H5: *Autonomy oriented study participants generate significantly more tags than control oriented participants, when in the points, level and leaderboard conditions.*	Partially
H6: *Autonomy oriented study participants report significantly more competence need satisfaction in the points, level and leaderboard conditions, compared to control oriented participants.*	No
H7: *Autonomy oriented study participants report significantly more intrinsic motivation in the points, level and leaderboard conditions, compared to control oriented participants.*	Partially

Table 4: Overview of hypotheses and results.

5. Discussion

Our motivation for the present study was to systematically assess whether points, leaderboards and levels increase performance, competence need satisfaction and intrinsic motivation in an image annotation task, while taking participants' general causality orientation into account. In line with previous research on the potential of game elements to promote user behavior (e.g., Cechanowicz et al., 2013; Denny, 2013; Hamari, 2013; Von Ahn & Dabbish, 2008), points, and especially levels and the leaderboard prompted participants to generate significantly more tags, although tag quality remained unaffected. Against our expectations, the different conditions did not differ concerning intrinsic motivation, or competence need satisfaction, nor did participants' causality orientation influence the effects of game elements on performance, need satisfaction or intrinsic motivation. As postulated by Deci & Ryan (1985), autonomy oriented participants reported more intrinsic motivation than control oriented individuals, and produced also significantly more tags. In line with SDT, intrinsic motivation was positively correlated with autonomy and competence need satisfaction, and slightly positively correlated with tag quantity.

5.1. Game elements and performance gains

Zagal et al. (2005) categorize points, levels and leaderboards as different types of goal metrics, as they usually represent and sometimes even define player success. By communicating how many tags have been generated, points likely also formed a clearer connection between participants' effort and their performance in the image annotation task (Jung et al., 2010; Von Ahn & Dabbish, 2008), which may have led to increased tag quantity compared to the plain condition. Seeing how Jung et al. (2010) found that users' performance increased when given a clear goal as opposed to users who were simply asked "to do their best", even if the latter were aware of their performance (i.e., points), it seems plausible that levels and leaderboard further reinforced tagging performance by setting explicit goals for participants to aspire to (Hamari, 2013; Jung et al.,

2010; Von Ahn & Dabbish, 2008).

Next, while game elements did not affect tag quality, tag quantity was a significant negative predictor of tag quality in all experimental conditions. While, the present study does not allow for establishing a causal relationship between increased tag quantity and decreased quality, this may indicate that participants motivated to generate many tags, might have disregarded tag quality in favor of tag quantity. Notably, this also means that participants in the gamified conditions generated more tags than participants in the plain condition, but at a comparable quality. Overall, it could be argued that they performed *better* than participants who were not presented with any game elements, further hinting at the potential effectiveness of gamification to promote certain behaviors. However, these findings should only be applied with caution since game elements might only work short-term due to novelty effects (Hamari et al., 2014; Koivisto & Hamari, 2014).

Further, while tag quantity was slightly positively correlated with intrinsic motivation, participants' reported intrinsic motivation did not reflect their performance. Interestingly, a recent meta-analysis by Cerasoli et al. (2014) found that intrinsic motivation only moderately predicted performance quantity in a variety of domains, whereas extrinsic incentives were found to be strong positive predictors. While in the present study game elements did not *decrease* intrinsic motivation, the fact that they improved tag quantity without increasing intrinsic motivation or competence need satisfaction, suggests that in this particular study context, points, levels and leaderboards may have functioned as (effective) extrinsic incentives. Note that extrinsic incentives need not invariably be perceived as controlling and subsequently undermine intrinsic motivation (Cerasoli et al., 2014; Deci et al., 1999). And although game elements led to overall performance gains, recall that only intrinsic motivation is associated with increases in the extent and quality of effort that people put into a given task (Cerasoli et al., 2014), a fact that is possibly reflected in the lack of improvements in

terms of tag quality.

5.2. Lack of effects on competence need satisfaction and intrinsic motivation

None of the game elements significantly affected intrinsic motivation or need satisfaction, nor was this further moderated by participants' causality orientation. Against our assumptions, game elements were apparently not perceived as particularly informational (Deci et al., 1999) and did not lead to more feelings of competence or intrinsic motivation compared to the plain condition. Contrary to previous claims on the need-supportive potential of game elements in non-game contexts (e.g., Francisco-Aparicio et al., 2013; Jung et al., 2010), this suggests that points, levels and leaderboards do not readily afford competence need satisfaction, even when encountered in a non-controlling setting. In the following, we discuss several reasons that might account for this finding:

Firstly, the game elements might not have offered enough meaningful, informational feedback to help participants judge their performance (Deci et al., 1999). Points informed participants about how many tags they generated, and the levels/leaderboard provided performance targets to aim for (Hamari, 2013; Jung et al., 2010; Von Ahn & Dabbish, 2008). Yet the present study featured no explicit indication of how many tags actually constituted a "good" performance and participants could therefore not judge whether they were competent at the image annotation task. Also, Wang et al. (2015) found that low performers experienced significantly less competence. Hence, it is possible that performance – or the mere impression that one is performing poorly, – further moderates the effects of game elements on competence need satisfaction.

Secondly, the motivational appeal of many games lies in their ability to provide players with challenges to master, hence allowing them to experience feelings of competence (Deterding, 2015; Przybylski et al., 2010). The image annotation task, on the other hand, could hardly be considered challenging, as participants were free to create as many tags as they wanted. Even in the

plain condition participants generated on average more than 50 tags, which corresponds to reaching the first two performance goals set in the level and leaderboard conditions. According to Wang et al. (2015), moderate performance targets do not motivate people to put much effort into achieving that target goal. In this case, informational feedback does not further encourage people to achieve more challenging targets and is thus less likely to satisfy their need for competence. In short, it seems plausible that points, levels and leaderboards only afford competence need satisfaction for tasks that are actually experienced as challenging.

Thirdly, even in the case of points, levels or the leaderboard providing sufficiently informational feedback, their visual presentation was very understated and lacked "juiciness". In contrast, many digital games provide excessive positive – juicy – feedback in the form of sounds, visuals and animations (Przybylski et al., 2010). According to Juul (2012), *juiciness does not simply communicate information [...] but also gives the player an immediate, pleasurable experience [...] enhancing the experience of feeling competent, or clever, when playing a game* (pp. 45). While the topic still requires empirical work, it seems plausible that a visually and aurally more impressive presentation of points, levels and leaderboards might have amplified their potential to afford competence need satisfaction and subsequently increase intrinsic motivation.

Fourthly, the present study's findings may not as much be due to the game elements themselves, but rather due to the nature of the task. Participants were only scored for tag quantity, and did not receive any feedback on whether a tag was fitting an image or not. If the image annotation platform would have also rewarded tag quality, participants would not only have received more feedback, but maximizing tag quality could have posed another challenge for people to master. The human computation games by Von Ahn & Dabbish (2008), for example, turn maximizing tag quality into a game mechanic by pairing two players together and having them guess how the *other* player would describe a given im-

age, while avoiding a list of "taboo" tags. While points, levels and leaderboards may further add to the task's enjoyment (Von Ahn & Dabbish, 2008), players may actually mostly enjoy engaging in the guesswork involved, which provides both challenge and team-play thereby affording opportunities for satisfying the needs for competence and relatedness. Points, levels and leaderboards on their own might not be as meaningful or informational without this additional "game mechanic".

Similarly, achievement goal theory differentiates between two types of goals, namely, mastery and performance goals (Rawsthorne & Elliot, 1999). While mastery goals refer to skill development and task mastery, performance goals focus on the demonstration of competence relative to normative standards. A meta-analysis on the effects of performance and mastery goals on intrinsic motivation found that informational feedback only increased intrinsic motivation for mastery goals, whereas performance goals were left unaffected (Rawsthorne & Elliot, 1999). Indeed, the image annotation task used in the present study bears more resemblance to a performance than a mastery goal, as participants simply had to "demonstrate" their competence in tagging paintings, relative to the norm set by levels or the leaderboard.

Finally, while the study was engaged in voluntarily, participants were encouraged to solve human computation "tasks" rather than "games". Lieberoth (2015) however, could show that explicitly labeling an activity containing game design elements as a "game" may increase people's intrinsic motivation, compared to a control task featuring no such framing or game elements. Indeed, Huotari & Hamari (2012) state that the goal of well-thought out gamification is to provide gameful experiences, and Deterding (2014, 2015) stressed that rather than (re-)structuring objects to look more like games (e.g., superficially applying points, levels and leaderboards to an image annotation task), a non-game context should instead be framed in such a way that people experience it as "game-like". Participants in the present study likely did not experience the

image annotation task as gameful / game-like. Perhaps, intrinsic motivation might have increased, if framing of the image annotation task as a game or game-like activity were facilitated, – even if the task itself remained unchanged. However, it is important to note that this may not hold true for all people under all circumstances, as previous research found that people may be suspicious of encountering games in unexpected settings (e.g., Littleton et al., 1999).

5.3. Limitations and further research

Firstly, it has to be noted that the nature of the image annotation task employed may have influenced the outcome of the studies. In contrast to other image annotation tasks (e.g., Von Ahn & Dabbish, 2008), participants were asked to describe the emotional content, and not the actual content of the paintings, which arguably may have also been shaped by participants' subjective experience of the paintings. Seeing how some of the paintings were more abstract, some participants may have experienced difficulties in coming up with suitable tags. The rather free-form nature of the image annotation task may have made it difficult for participants to know when they performed well, because there was no apparent correct way on how to describe the emotional content of the pictures. In order to further understand the psychological mechanisms underlying gamification and its effects on performance outcomes, it would be necessary to also study tasks with relatively clearly defined quality metrics, which make it easier to assess and provide feedback on performance quality.

Secondly, participants interacted only for a short time with the image annotation platform. However, Hamari et al. (2014) and Koivisto & Hamari (2014) found indications that gamification may only promote user engagement for a short time. For instance, the findings of Hanus & Fox (2015) have shown the importance of studying the long-term effects of gamification, in order to better assess whether and under what circumstances game design elements shape user behavior in the long run. However, our findings are still relevant for non-game contexts in which long-term user engagement and retention may not necessarily

be the primary goal, such as increasing participation and performance in crowdsourcing tasks.

Thirdly, only participants self-reported intrinsic motivation was measured. While self-reported and free choice measures of intrinsic motivation yielded comparable results in previous studies (Deci et al., 1999), employing a behavioral free choice measure of intrinsic motivation by letting participants choose whether they want to continue engaging with a given task, may yield additional insights. For instance, Cechanowicz et al. (2013) found that , compared to the plain condition, gamification of a market research survey motivated participants to continue engaging with the survey, even after conclusion of the "mandatory" part of the experiment. Future studies should thus consider combining self-reported and behavioral measures of intrinsic motivation for additional methodological robustness. Also, because participants from the universitys database usually engage voluntarily in studies, it is possible that they already had a minimum level of intrinsic motivation from the get-go, which might have affected the results of the present study. More research is required to investigate how users initial motivation to engage in a gamified application affects their subsequent motivation.

Fourthly, while our study covered a wide age spectrum, our sample contained relatively few male participants (84 men vs 178 women). Although we did not find any age or gender differences, demographic factors have been found to affect people's reaction to gamification, especially in the long run (Koivisto & Hamari, 2014).

Finally, the results of the present study are specific to the image annotation context. Hence, our findings should only cautiously be applied to other gamified applications. It still has to be seen whether these results can be replicated for other non-game contexts. Moreover, further research into individual factors, such as people's general causality orientation or varying degrees of competitive-

ness among users (e.g., Song et al., 2013), is required, since they may potentially moderate the effects of gamification on user behavior and motivation (Hamari, 2013; Hamari et al., 2014). Similarly, more research is required to further investigate the role of contextual, social and situational aspects, as they at least partially determine the motivational affordance of game design elements (Deterding, 2011, 2014). Lastly, game elements may also support other needs besides competence, such as autonomy and relatedness (Francisco-Aparicio et al., 2013; Pe-Than et al., 2014; Peng et al., 2012). It would thus prove insightful to empirically test whether specific game elements afford the satisfaction of individual needs in a variety of non-game contexts.

6. Conclusion

The present study is one of the first to cover several aspects still underexplored in current gamification research (Hamari et al., 2014; Seaborn & Fels, 2015). Firstly, we attempted to empirically evaluate the impact of gamification on intrinsic motivation and need satisfacition, two of the most frequently appealed to, yet seldom empirically studied constructs in gamification literature (Seaborn & Fels, 2015). We did so by employing a validated, theory-based instrument, the Intrinsic Motivation Inventory, in addition to measures of two different behavioral outcomes, tag quantity and quality. As of now, our study is also one of the first to isolate the effects of individual game elements in a comparative experiment. In the context of the present study this meant that while points, levels and leaderboards increased tag quantity, the lack of effects on intrinsic motivation, need satisfaction or tag quality suggest that they may have actually functioned as extrinsic incentives (Cerasoli et al., 2014). However, seeing how they did not impair intrinsic motivation in contrast to previous findings (Hanus & Fox, 2015), points, levels and leaderbords seem to be an effective means for promoting performance quantity. More empirical research is necessary on why particular game elements act as extrinsic or intrinsic motivators in a given context, and how this in turn shapes user enjoyment and behavior, but

we believe our study is a valuable first step in this direction and may serve as a blueprint for future studies.

7. References

Antin, J., & Churchill, E. (2011). Badges in social media: A social psychological perspective. In *Gamification: Using Game Design Elements in Non-Gaming Contexts, a workshop at CHI*.

Antin, J., & Shaw, A. (2012). Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the us and india. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (pp. 2925–2934). ACM.

Cechanowicz, J., Gutwin, C., Brownell, B., & Goodfellow, L. (2013). Effects of gamification on participation and data quality in a real-world market research domain. In *Gamification '13* (pp. 58–65). ACM.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin*, *140*, 980–1008.

Deci, E., Koestner, R., & Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin*, *125*, 627–668.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, *19*, 109–134.

Denny, P. (2013). The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 763–772). ACM.

Deterding, S. (2011). Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using Game Design Elements in Non-Gaming Contexts, a workshop at CHI*.

Deterding, S. (2012). Coding conduct: Games, play, and human conduct between technical artifacts and social framing. Retrieved June 1, 2013 from http://www.slideshare.net/dings/coding-conduct-games-play-and-human-conduct-between-technical-code-and-social-framing.

Deterding, S. (2014). Eudaimonic design, or: Six invitations to rethink gamification. In M. Fuchs, S. Fizek, P. Ruffino, & N. Schrape (Eds.), *Rethinking Gamification* (pp. 305 – 331). meson press.

Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. *Human–Computer Interaction*, *30*, 294–335.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). ACM.

ESA (2015). *Essential Facts about the Computer and Video Game Industry*. Technical Report Entertainment Software Association. Retrieved July 28, 2015 from http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf.

Francisco-Aparicio, A., Gutiérrez-Vela, F. L., Isla-Montes, J. L., & Sanchez, J. L. G. (2013). Gamification: Analysis and application. In *New Trends in Interaction, Virtual Reality and Modeling* (pp. 113–126). Springer.

Hamari, J. (2013). Transforming homo economicus into homo ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications*, *12*, 236–245.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work?–literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences*.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison,

satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152–161.

Huotari, K., & Hamari, J. (2012). Defining gamification: a service marketing perspective. In *Proceeding of the 16th International Academic MindTrek Conference* (pp. 17–22). ACM.

Jung, J., Schneider, C., & Valacich, J. (2010). Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. *Management Science*, *56*, 724–742.

Juul, J. (2012). *A casual revolution: Reinventing video games and their players*. The MIT Press.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, *35*, 179–188.

Lieberoth, A. (2015). Shallow gamification testing psychological effects of framing an activity as a game. *Games and Culture*, *10*, 229–248.

Littleton, K., Ashman, H., Light, P., Artis, J., Roberts, T., & Oosterwegel, A. (1999). Gender, task contexts, and children's performance on a computer-based task. *European journal of psychology of education*, *14*, 129–139.

Machajdik, J., & Hanbury, A. (2010). Affective image classification using features inspired by psychology and art theory. In *Proceedings of the international conference on Multimedia* (pp. 83–92). ACM.

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013a). Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1137–1142). ACM.

Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013b). Do points, levels and leaderboards harm intrinsic motivation?: an empirical analysis of

common gamification elements. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (pp. 66–73). ACM.

Pe-Than, E. P. P., Goh, D. H.-L., & Lee, C. S. (2014). Making work fun: Investigating antecedents of perceived enjoyment in human computation games for information sharing. *Computers in Human Behavior*, *39*, 88–99.

Peng, W., Lin, J., Pfeiffer, K., & Winn, B. (2012). Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology*, *15*, 175–196.

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

Rawsthorne, L. J., & Elliot, A. J. (1999). Achievement goals and intrinsic motivation: A meta-analytic review. *Personality and Social Psychology Review*, *3*, 326–344.

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*, 54–67.

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

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31.

Song, H., Kim, J., Tenzek, K. E., & Lee, K. M. (2013). The effects of competition and competitiveness upon intrinsic motivation in exergames. *Computers in Human Behavior*, *29*, 1702–1708.

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview,

emerging trends, and future directions. *Advances in motivation and achievement*, *16*, 105–165.

Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, *51*, 58–67.

Wang, X., Schneider, C., & Valacich, J. S. (2012). Fine-tuning group collaboration environments: How differences in general causality orientation and performance targets shape interaction and performance. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (pp. 553–561). IEEE.

Wang, X., Schneider, C., & Valacich, J. S. (2015). Enhancing creativity in group collaboration: How performance targets and feedback shape perceptions and idea generation performance. *Computers in Human Behavior*, *42*, 187–195.

Zagal, J. P., Mateas, M., Fernández-Vara, C., Hochhalter, B., & Lichti, N. (2005). Towards an ontological language for game analysis. In *Proceedings of International DiGRA Conference: Changing Views – Worlds in Play.* (pp. 3–14).

Zhang, P. (2008). Motivational affordances: reasons for ict design and use. *Communications of the ACM*, *51*, 145–147.

Increasing Donating Behavior Through a Game for Change: The Role of Interactivity and Appreciation

Sharon T. Steinemann, Elisa D. Mekler, Klaus Opwis
University of Basel, Department of Psychology, Center for Cognitive Psychology & Methodology
Missionsstrasse 62a, 4055 Basel, Switzerland
sharon.steinemann@unibas.ch, elisa.mekler@unibas.ch, klaus.opwis@unibas.ch

ABSTRACT
Games for change have attracted the interest of humanitarian aid organizations and researchers alike. However, their effectiveness to promote behavior such as donating remains unclear. Furthermore, little is known about how key game properties interactivity and presentation mode impact the effectiveness of these games, or how player attitudes and experiences relate to the interplay between game properties and donating behavior. In this study, experimental conditions were systematically varied in their interactivity and presentation mode. Thereby, 234 participants played, watched, or read through one of six variations of the narrative of the game *Darfur is Dying*. Following this, they were asked to choose the percentage of an unexpected bonus to donate to a charity. While interactivity increased donating by an average of 12%, presentation mode had no significant impact on the percentage donated. Thus, between presentation mode and interactivity, interactivity was found to be the more impactful game property. Moreover, appreciation fully mediated the relationship between interactivity and donating, hinting at its relevance for the evaluation of the effectiveness of games for change.

Author Keywords
Games for change, persuasive games, appreciation, donating

ACM Classification Keywords
J.4 Social and Behavioral Sciences:: Sociology, Psychology; K.8.0 Personal Computing Games

INTRODUCTION
For organizations focused on humanitarian aid, the rise of new media and technology brings with it the potential for creating new ways to make the world a better place. Most such organizations depend on reaching the public and persuading individuals to help [39]. Working on a limited budget, finding ways to do this in a way that is both efficient and effective is pivotal [42]. An interesting option, which in recent years has caught the attention of both humanitarian aid organizations and researchers alike, are games for change [7, 35].

CHI PLAY 2015, October 03 - 07, 2015, London, United Kingdom.

ACM 978-1-4503-3466-2/15/10...$15.00.
DOI: http://dx.doi.org/10.1145/2793107.2793125

Games for change, also known as social impact games or serious games for social change, are digital games with the purpose of not only entertaining, but reaching players and ideally animating them to support the social change the game is advocating [31, 36]. *Darfur is Dying* for example confronts players with the fear and constant lack of security facing Darfurian refugees, by forcing the player to attempt to bring water back to their camp, while avoiding the patrolling Janjaweed militia. Meanwhile, *Spent* illustrates how quickly poverty can spiral into homelessness by having players try to survive on a minimal income, while being faced with choices such as deciding whether to get an expensive treatment for a dental infection for half their monthly income or to buy numbing cream and try to ignore the pain. The appearance of games for change can vary greatly, from high-quality video games, to simple cartoon-like animation, to text-based gameplay [22]. The vast majority of these games however have in common that no matter how simple the design, they force players to face challenges and make difficult choices they would not have to in their regular life. This aspect of interactivity also sets these games apart from other forms of media conveying similar messages.

While interest in various types of serious games has been on the rise in recent years, little is known about the effectiveness of these games in achieving the goals for which they were designed [26, 44]. Especially in the context of games for change, effectiveness can be hard to discern, as their purpose may not be clearly defined or it may be difficult to distill their success down to measurable values [21, 28, 36]. However, as argued by Iacovides and Cox [21], despite these added difficulties it is especially vital to evaluate the experience invoked by games that go beyond fun, to understand whether they are effective facilitators of social change.

Related to the question of the overall effectiveness of games for change is the examination of how individual game properties, such as interactivity or presentation mode, contribute to this effect, as understanding the specific impact of individual game properties can help organizations create impactful games efficiently, by focusing on the most effective properties. Although still a very sparse field of research, the few studies that have aimed to evaluate the effectiveness of games for change have found support for their impact on a number of factors such as player attitudes or knowledge gain [34] and some have even found effects on attitudes weeks later [20, 38].

However, so far very few studies have examined the impact of games for change on behavior-related variables. The only

studies that have to our knowledge touched on behavior did this by asking participants about their willingness to show a certain behavior or via self-reports, for example by asking how willing participants would be to donate money [34] or asking them a week after the study, whether they had shared the game with a friend [12]. So far to our knowledge no studies have explored the impact of games for change on directly measurable behavior.

The goal of the current study is first and foremost to examine the effectiveness of a game for change on impacting behavior, specifically donating. In particular, we wish to understand how two specific game properties contribute to the effectiveness of these games by analyzing the impact of systematically changing presentation mode or adding and removing interactivity, while keeping the information content itself constant. Beyond that, we also wish to gain insight into how role-taking, willingness to help, enjoyment, and appreciation are affected by playing a game for change and how they relate to donating. Examined together, these questions should help us improve our understanding of how games for change can impact player attitudes, experience, and behavior and thereby further understand these games' value as a tool for meaningful social change.

RELATED WORK

A key game property, both for games as a whole and for games for change in particular, is interactivity [36]. Klimmt and Vorderer [23] defined interactivity in the context of digital games as the game property that gives the player the ability to interact with and influence the story told in the game. Previous research in the context of games for change, other serious games, and in the broader field of game research, has identified the interactivity of a game as a factor, which contributes to different attitudinal and motivational changes, as well as to knowledge gain.

Peng, Lee, and Heeter [34] for example had participants either play Darfur is Dying, watch a recording of the gameplay (but not play the game themselves), or read a text describing the same events as were played out in the game. They found that, compared to the other two conditions, playing the game led to significantly higher role-taking with the character and willingness to help people, who, like their character, were affected by the crisis in Darfur. Similarly, Ruggiero [38] found that playing Spent led to participants improving their attitude towards homelessness more than a control group and more than a group that read an article about homelessness. The effects were weaker three weeks later, but had decreased less for the game group than for the control or reading group. Ritterfeld et al. [37] also compared an interactive serious game, in this case for education, with the noninteractive recorded gameplay and found a significant effect for interactivity on knowledge gain in the game's subject of the human digestive system. Further, while not exclusively focusing on interactivity, a meta-analysis of games for education by Wouters et al. [44] found games to be more effective at encouraging learning and retention than conventional, mainly non-interactive instructional methods. An explanation for this finding could be the ability of interactive games to allow players to experience and manipulate the game's material and outcome in ways other instructional methods cannot [19]. While this effect may be especially relevant in the context of games for learning, this aspect, of seeing the consequences of certain actions and the therewith connected learning experience, may well also hold importance for the creation of new or adapted attitudes and thereby also be of significant importance for games for change. In the wider game context, interactivity has also been examined, as Lin [24] compared an interactive violent video game with the noninteractive recorded gameplay and a noninteractive corresponding scene in a movie, on which the game was based. Lin found interactivity to have a significant short-term effect on both aggressive affect and cognition. Once again however, the impact on behavioral measures remains unclear.

Another important property of games is their presentation mode and their ability to incorporate different forms of sensory input, such as visualizations and audio tracks [36]. This combination of different sensory perceptions in one presentation mode is referred to as multimodality [4]. Past research has shown that the use of multimodality can impact the way information is processed. For example, using multiple modalities, distributed over separate sensory channels in parallel, as by supplementing a lecture with descriptive images, can improve information processing [26, 40]. However, relaying two pieces of information simultaneously over the same channel can lead to inferior processing [4]. Simply changing the presentation mode can also have consequences on attitudes, as for example presenting a political debate either in a single-screen view or a split-screen view can significantly impact the way viewers process the debate issues and evaluate candidates [10]. Findings on the impact of presentation mode and modality in the context of serious games have been mixed. For example, Ritterfeld et al. [37] found that using combined images and sound instead of text in a game for education significantly increased knowledge gain and interest in learning. Conversely, Peng et al. [34] found no significant difference between the recording of the cartoon-like animation supplemented by sound in the game Darfur is Dying and a text telling the same story regarding their impact on role-taking and willingness to help.

While several studies have examined the effects of interactivity and presentation mode and some, such as Peng et al. [34], have even included them in the same design, the effect of interactivity has to our knowledge never been systematically examined across different presentation modes. Peng et al. [34] for instance did not include an interactive equivalent of the text, meaning that any findings on the effectiveness of the interactive game could not be completely attributed to interactivity, as it might have been an effect of the interaction between interactivity and the presentation mode of the game.

Aim of this study

In this study, our goal was to firstly examine the individual effects of interactivity and presentation mode on a measurable behavioral variable, namely donating, considering the importance of donations for humanitarian aid [39]. The effects of interactivity and presentation mode were examined by com-

paring interactive and noninteractive versions of three different presentation modes. As research on games for change has so far lacked examination of the impact of game properties on behavior, it will be interesting to see whether the link found between interactivity and willingness to donate [34], can also be found between interactivity and donating behavior. As attitudes are generally connected to corresponding behavior [1, 18], we hypothesize that as with willingness to donate:

H1: Interactivity will lead to increased donations.

As findings on the effectiveness of presentation mode in the context of serious games have been mixed [34, 37], no specific hypotheses will be proposed regarding its impact. As the study design will however allow a systematic examination of the impact of presentation mode, as with interactivity, results regarding the effectiveness of presentation mode will nevertheless be of interest.

Of further interest is the question of how subjective player ratings are impacted by interactivity and presentation mode and how they relate to donating. This will allow the findings of this study to be compared to previous research examining subjective ratings. In the following sections we will therefore focus on four factors shown to be of importance in recent serious game and media research.

Role-Taking and Willingness to Help

Empathy has long been associated with prosocial behavior [17]. Role-taking refers to a specific form of empathy whereby a person temporarily imagines themselves as another person and takes on their perspective [14]. In the context of media, role-taking is closely related to identification, which effectively describes role-taking specifically with a mediated character, such as a character in a book, movie, or game [13, 34].

In their study, Peng et al. [34] argued that due to the game's interactivity, playing as a character in a game could lead to an increase in role-taking with that character more than noninteractive forms of media could. In line with this expectation, they found that interactivity led to an increase in role-taking. Similar results are expected for this study:

H2: Interactivity will lead to more role-taking.

Furthermore, Peng et al. [34] also found that an increase in role-taking led to an increase in self-reported willingness to help a cause that would benefit people like the character they had just played. It remains to be seen how role-taking and willingness to help translate into donating behavior, however, considering the well-established links between empathy and prosocial behavior [17] as well as research linking attitudes with related behavior [1, 18], we hypothesize that:

H3: Role-taking will be positively correlated with donating behavior.

H4: Willingness to help will be positively correlated with donating behavior.

Peng et al. [34] did not find an impact of presentation mode on role-taking or on willingness to help. As their design however did not allow for a clear distinction between the impacts of interactivity and presentation mode, it is possible that isolating the effects of presentation mode may yield different findings.

Enjoyment and Appreciation

A key aspect of games, which should not be neglected, is their capacity to be entertaining and enjoyable [27]. Of special interest in the context of games for change is the concept of eudaimonic entertainment [33]. This describes entertainment, which leads to gratification not necessarily by being fun, but by being thought-provoking and meaningful [6, 25]. Examples of media that are in this way appreciated, but not necessarily enjoyed, are films such as Schindler's List, which may not be considered fun, but is widely appreciated for its ability to make the audience think [33]. It is important to bear in mind that media can be simultaneously appreciated and enjoyed [33]. Accordingly, while the game may be appreciated for its message, the gameplay may still be enjoyable. Specifically, the interactivity of a game has in the past been linked to enjoyment [23]. We therefore assume that:

H5: Interactivity will lead to more enjoyment.

Appreciation has been linked to the degree to which media is moving and thought-provoking [5, 6, 33]. A recent study by Bopp et al. [8] found games to be an effective medium for inspiring both strong emotions and reflective thoughts. Interestingly, players often especially appreciated game situations where both positive and negative emotions were elicited, such as when the player won, but only after having made a sacrifice. This may be especially relevant for games for change, such as Darfur is Dying, where the player is likely to be confronted with negative emotions when contemplating the humanitarian crisis the game is illustrating. While the resulting game experience may not necessarily be fun, this does not have to make it a bad experience or one that game designers should avoid designing for. As Marsh and Costello [25] have argued, other forms of media such as literature or film are often acclaimed for their ability to portray suffering and adversity and that limiting their storytelling to be only positive or fun would be considered a serious restriction. Marsh and Costello advocate that there is no reason why games should not similarly aim to be moving and thought-provoking. Considering the above-mentioned research by Bopp et al. [8] and other recent findings on the ability of games to promote affective learning [38] and stronger affective reactions [24], it could be postulated that games may even be uniquely qualified to facilitate moving experiences. As a game property, interactivity in particular has been associated with stronger cognitive and affective reactions [24]. We therefore hypothesize that:

H6: Interactivity will lead to more appreciation.

Lastly, moving media, distinguished by the presence of both positive and negative emotions, has been associated with an increased likelihood of participants performing prosocial behavior, such as sharing an informational video with others to spread awareness around skin cancer prevention [30]. Additionally, past research has highlighted the importance of meaningfulness for people engaging in prosocial behavior,

such as donating blood [2, 3]. Considering this, we propose that:

H7: Appreciation will be positively correlated with donating behavior.

It is yet to be seen how presentation mode affects enjoyment and appreciation or how enjoyment impacts donating behavior. Finally, it will be interesting to explore whether any of these subjective ratings for player attitudes and experiences will be able to mediate the effect of interactivity and presentation mode on donating behavior.

METHOD

The experiment had a 2x3-between-subject design. The independent variables were interactivity with two levels (interactive, noninteractive) and presentation mode with three levels (text, text with pictures, (recorded) gameplay). The primary dependent variable was percentage donated. Further dependent variables were role-taking, willingness to help, enjoyment, and appreciation. To control for confounding effects, the covariables empathic concern, general involvement with international humanitarian affaires, and previous knowledge of the crisis in Darfur were also included.

Participants

Participants were recruited on the crowdsourcing platform Crowdflower. Participants were only allowed to participate once, any repeated participations were excluded from the dataset (73 participants). We also excluded participants who did not complete the survey (19 participants) or had technical difficulties, which led to them being unable to experience the experimental condition they were assigned to (3 participants). We also excluded participants who were unable to correctly answer an open-ended question about what had happened in the experimental condition (6 participants) or obviously had randomly answered multiple choice questions (8 participants). As only the game condition included the option of winning, we excluded participants, who indicated they had won (5 participants), to keep the outcome consistent across conditions.

After data cleanup, a sample of 234 participants (121 female) remained: 29 in the gameplay condition, 31 in the interactive text with pictures condition, 43 in the interactive text condition, 40 in the noninteractive recorded gameplay condition, 39 in the noninteractive text with pictures condition, and 52 in the text condition. The mean age was 38 with a range from 16 to 79. After conducting a pilot study, we realized that good English skills were essential for understanding the questionnaires and the text conditions. Therefore for the main study we restricted recruitment to countries with English as an official language. Despite this constraint, participants came from a fairly broad range of nationalities; 35% identifying as American, 23% as British, 22% as Canadian, and the remaining 20% identifying as from one of 28 other nationalities. 49% were full-time employed, 19% were unemployed, 13% were part-time employed, 8% were students, and the remaining 11% identified as either stay-at-home parents, self-employed, retired, or preferred not to say. Participants received $1 for their participation, which they were assigned after entering a code on Crowdflower that they were awarded at the end of the study.

Materials

Stimuli

We partially replicated the design of Peng et al. [34], using the same two presentation modes as they had. These were the interactive web-based video game Darfur is Dying, which has been previously used in research on games for change (e.g., [12, 31, 34]), a recorded gameplay video of Darfur is Dying, and a text, recounting the narrative of the game. While there were several characters available, we asked participants – similar to Peng et al. [34], – to play as the little girl Poni, for the sake of consistency across the other conditions, which only offered the option of playing as Poni. Likewise, we excluded participants,who won (i.e., successfully brought water back to their camp without getting caught) to keep the outcome consistent across conditions.

We furthermore supplemented Peng et al.'s study design with an additional three conditions. The first was an interactive version of the text adapted from Peng et al. ([34], p. 741). This interactive text was a simple form of interactive fiction or text adventure, which allowed the reader to make choices as to how the story would progress. The interactive text condition was created by modifying the noninteractive text in Twine, a software, which allows the creation of hypertext-based interactive stories. Where the noninteractive text condition described the decisions Poni made when running across the landscape to get to the well, the interactive text condition let the participants choose in which direction Poni should run. The player would make their choice by clicking on their preferred answer and were then taken to a new page in the browser with a text reflecting their choice (see also Figure 1).

To keep the experience consistent with the noninteractive text, the interactive text told the same story, independently of the choices the player made, although the players did not know this. While the player would choose a direction in which to run and the next page would give feedback about the direction they ran in (e.g., "Poni runs east, away from the oncoming jeep"), the rest of the text would be the same for each option. The only exception was if participants chose an option that took them towards the jeep, in which case they were captured immediately. If they did not run towards the jeep, players went through eight pages, on seven of which they were given a section of the story and had to choose which way to run. To keep the story consistent across conditions, the final choice always lead to Poni being captured on the eighth page.

While interactive fiction may be visually very different to a video game such as Darfur is Dying, it can nonetheless be defined as a form of game [29]. Considering the substantial visual difference between an interactive text and a video game, a third presentation mode of a text with pictures using screenshots from the game (see Figure 1) was included, to allow a more nuanced examination of the impact of different forms of presentation mode. The three presentation modes also varied in their use of modality, as the gameplay offered visual information, as well as auditory information (e.g., Poni's footsteps, motor sounds of nearby militia jeeps),

Poni runs east, away from the oncoming jeep. She successfully avoids the Jeep. Perhaps there will not be any more patrols, she thinks, but that familiar whine is heard in the distance to the north once again. Now she is in a bad position, as she has gone further away from the rocks and has no place to hide.

What should Poni do?

Run north, towards the well
Run east, away from the well
Run south, towards camp
Run west, towards the well

Figure 1. A section of the interactive text with pictures condition. The image is a screenshot of *Darfur is Dying* taken by the first author.

making it multimodal. Meanwhile, the text offered information only through written language as a single modality and the text with pictures offered information through written language and images, meaning that two pieces of information had to be transferred over the same processing channel.

Measures

To measure donating behavior, participants were given a $1 bonus in addition to the $1 that they were already receiving as compensation for taking part in the study. While $1 may not seem to be a large amount, several studies have previously employed this or similar amounts (e.g., [9, 16, 41]). Participants had to choose which percentage of this $1 they wanted to have paid to them and which percentage should be donated for them to the charity Save Darfur. Using a dropdown menu, participants selected the amount to be donated in 10-percent increments between 0% and 100%.

Cohen's identification scale [13] was used to measure role-taking, with the name of the character of the game, Poni, inserted in the item statements (Cronbach's α = .90). Participants were asked to use a 7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree) to rate six statements, such as "I was able to understand the events in a manner similar to that in which Poni understood them" or "When Poni failed, I was sad; when Poni succeeded I felt joy".

While role-taking is a measure for the empathy felt towards a specific individual, or in this case character, it is reasonable to assume that a person's general tendency towards empathy may also impact prosocial behavior such as donating. To control for this potentially confounding factor, the Empathic Concern subscale developed by Davis [15] was utilized (Cronbach's α = .86). Participants rated seven statements, such as "When I see someone taken advantage of, I feel kind of protective towards them" on a 7-point Likert scale from 1 (does not describe me well) to 7 (describes me very well).

Besides empathic concern, it is also plausible that a person with an interest in following news about humanitarian issues or someone with previous knowledge of the crisis in Darfur would be more likely to donate money to this cause. To measure general involvement with international humanitarian affaires, participants were asked the same four questions used in Peng et al. [34], such as "I pay attention to news about human rights", on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree, Cronbach's α = .92). To measure previous knowledge of the crisis in Darfur, one item was used similar to Peng et al., asking whether the participant had heard of the crisis in Darfur.

Four dependent variables focusing on willingness to help were measured using the same four questions used by Peng et al. [34]. The participants were asked to rate how likely on a 7-point Likert scale from 1 (very unlikely) to 7 (very likely) it was that they would (a) donate money to help fund crucial awareness and advocacy programs needed to end the crisis in Darfur; (b) sign a petition to build the political pressure needed to end the crisis in Darfur; (c) discuss the situation in Darfur with their friends or family; and (d) forward the link of the game/video/text/interactive text to their friends to disseminate the message about Darfur. It is important to note that, as in Peng et al. [34], these four questions were not used as a scale measuring one variable, but as four separate variables measuring four separate forms of willingness to help.

Enjoyment and appreciation were measured using the scale developed by Oliver and Bartsch [33], which consists of three items each for enjoyment (Cronbach's α = .90) and appreciation (Cronbach's α = .87). To accommodate the different media used in this study, statements were slightly modified depending on the condition. For example, a statement for appreciation in the game condition was formulated as "I found this game to be very meaningful", while in the recorded gameplay condition it was written as "I found this video to be very meaningful" (7-point Likert scale, 1 = strongly disagree, 7 = strongly agree).

Procedure

Participants began the online survey after following a link from Crowdflower. After being informed about the rough procedure and length of the study, participants filled out questionnaires for empathic concern, as well as general involvement with international humanitarian affaires, and knowledge of the humanitarian crisis in Darfur. Next, participants were randomly assigned to one of the six experimental conditions, each of which told Poni's story through a different form of presentation mode (text, text with pictures or (recorded) gameplay) in either an interactive or noninteractive version. Participants completed the various experimental conditions (e.g., playing the game) in between five and seven minutes.

Immediately after the experimental condition, participants filled out questionnaires for role-taking, willingness to help, appreciation, and enjoyment. Then, participants were told that they would be receiving a bonus of $1, in addition to the $1 they were already receiving for participating in the study. They were then given the choice to keep the entire bonus for themselves or to donate up to 100 percent to a charity called Save Darfur. Optionally, they could click on a link to find out more about Save Darfur before making their decision. After choosing the amount to donate, participants were asked

Table 1. Descriptive statistics for all dependent variables for all conditions.

	Non-interactive			Interactive		
	Text	Text with Pictures	Recorded Gameplay	Text	Text with Pictures	Gameplay
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Percentage donated	49.42 (41.46)	52.56 (41.15)	50.75 (41.16)	65.12 (35.48)	55.16 (41.38)	66.55 (43.12)
Role-taking	5.11 (1.25)	5.12 (1.36)	4.85 (1.26)	5.75 (1.08)	5.07 (.99)	4.80 (1.27)
Enjoyment	2.73 (1.44)	2.64 (1.29)	2.65 (1.25)	3.93 (1.92)	4.15 (1.73)	3.77 (1.54)
Appreciation	5.31 (1.41)	5.05 (1.58)	4.92 (1.51)	5.90 (1.13)	5.40 (1.24)	5.39 (1.33)
Willingness to help	4.45 (1.62)	4.29 (1.64)	4.47 (1.51)	4.95 (1.62)	4.56 (1.51)	4.47 (1.49)
Empathic concern	3.39 (.62)	3.02 (.86)	3.13 (.63)	3.40 (.79)	3.32 (.77)	3.43 (.75)
Humanitarian involvement	4.88 (1.48)	4.72 (1.47)	4.89 (.97)	5.42 (1.20)	4.94 (1.25)	4.98 (1.35)

to briefly recount what had happened in the story they had read, watched, or played through. This was followed by some quality-check questions to ensure all media in their condition had been presented correctly (e.g., for the text with pictures condition: "In the story you just read, did you see pictures illustrating the story?"). Finally, participants answered demographic questions and three validation questions about the content of the study, which when answered correctly gave them two codes, which when entered on Crowdflower, led to them receiving their compensation and the chosen bonus.

RESULTS

An alpha level of .05 was used for all statistical tests. Across all conditions participants donated an average 56% of their bonus, this amounted to a total of $131 that we consequently paid to Save Darfur. The average percentage donated in each condition is shown in Table 1.

Percentage donated

To examine the effects of interactivity and presentation mode on percentage donated, the data were analyzed using a two-way analysis of variance (ANOVA) for unrelated samples. There was a significant main effect for interactivity ($F(1, 228) = 4.427$, $p = .036$, $\eta_p^2 = .019$). Percentage donated was significantly higher in the interactive conditions ($M = 62.52$, $SD = 39.45$) than in the noninteractive conditions ($M = 50.76$, $SD = 40.98$), supporting H1. Neither the main effect for presentation mode ($p = .77$), nor the interaction effect ($p = .53$) were significant.

Next, analyses were performed to examine whether empathic concern, humanitarian involvement, or knowledge of the crisis in Darfur might be confounding the effects of interactivity and presentation mode on percentage donated. The results indicated that empathic concern was not significantly correlated with percentage donated (see Table 2). The same analysis for humanitarian involvement similarly revealed no significant relationship between humanitarian involvement and percentage donated. To examine whether participants who had previous knowledge of the humanitarian crisis in Darfur had donated differently than those who had not, an independent t test for equal variances was conducted. Results were significant ($t(232) = 2.061$, $p = .040$), indicating that participants who had known about the crisis in Darfur donated significantly more ($M = 62.35$, $SD = 39.73$), than those who had not ($M = 51.32$, $SD = 40.82$). A t test was conducted to examine whether the interactive and the noninteractive conditions differed in their previous knowledge of Darfur. The difference between the groups was however not significant ($p = .812$).

Role-taking and willingness to help

An analysis of the impact of interactivity and presentation mode on role-taking revealed a significant main effect for presentation mode ($F(2, 228) = 5.25$, $p = .005$, $\eta_p^2 = .049$). As listed in Table 1, role-taking was highest in the text conditions, followed by the text with pictures conditions, and lowest in the (recorded) gameplay conditions. Planned contrasts further revealed that role-taking was significantly higher in the interactive text condition compared to the other five conditions, ($t(228) = 3.68$, $p < .001$). Lastly, neither the main effect for interactivity ($p = .302$), nor the interaction effect for interactivity and presentation mode ($p = .068$), on role-taking were significant. H2 was therefore not supported.

To examine the relationship between role-taking and percentage donated, the data were analyzed using Pearson's *r*. As can be seen in Table 2, role-taking was significantly positively correlated with percentage donated. Thus supporting H3. Interestingly, when the data were split by interactivity, the significant positive correlation remained for the noninteractive conditions between percentage donated and role-taking ($r(131) = .26$, $p = .003$), but disappeared for the interactive conditions ($r(103) = -.05$, $p = .607$).

To allow comparisons with the results reported by Peng et al. [34], a two-way multivariate analysis of variance (MANOVA) was conducted to examine the impact of interactivity and presentation mode on willingness to donate, willingness to sign a petition, willingness to discuss with friends and family, and willingness to forward message. Against our expectations and in contrast to the findings of Peng et al. [34], no significant effects were found for any of the four ratings (p-values between .11 and .85). For this reason and since all four items were moderately to strongly correlated ($r(234) = .56$ - $.81$, $p < .001$), we decided to collapse the four individual items into a single factor "willingness to help" for subsequent analyses (Cronbach's $\alpha = .69$).

Table 2. Pearson's Correlation for dependent variables and covariables over all conditions.

	Percentage donated	Role-taking	Enjoyment	Appreciation	Willingness to help	Empathic concern
Role-taking	.15*					
Enjoyment	-.03	.29**				
Appreciation	.25**	.76**	.23**			
Willingness to help	.22**	.67**	.30**	.69**		
Empathic concern	.08	.41**	-.01	.50**	.40**	
Humanitarian involvement	.10	.42**	.12	.48**	.58**	.40**

* Significant at $p < .05$. ** Significant at $p < .01$.

To examine the relationship between willingness to help and percentage donated, the data were analyzed using Pearson's *r*. In support of H4, the results indicated positive significant correlations for willingness to help with percentage donated, as can be seen in Table 2. Just as had been the case with role-taking, when the data were split by interactivity, the significant positive correlation remained for the noninteractive conditions between percentage donated and willingness to help ($r(131) = .33$, $p < .001$), but not for the interactive conditions ($r(103) = .06$, $p = .532$).

Enjoyment and appreciation

An ANOVA was conducted to examine the impact of interactivity and presentation mode on enjoyment, revealing a significant main effect for interactivity ($F(1, 228) = 33.99$, $p < .001$, $\eta_p^2 = .13$), but no significant effects for presentation mode ($p = .860$), or the interaction between interactivity and presentation mode ($p = .779$), thereby supporting H5. To examine the relationship between enjoyment and percentage donated, the data were analyzed using Pearson's *r*. No significant correlation was found.

An analysis of the impact of interactivity and presentation mode on appreciation likewise revealed a significant main effect for interactivity ($F(1, 228) = 6.05$, $p = .015$, $\eta_p^2 = .026$), but no significant effects for presentation mode ($p = .071$), or the interaction between interactivity and presentation mode ($p = .763$), thereby supporting H6. An analysis of the relationship between appreciation and percentage donated revealed a significant positive correlation ($r(234) = .25$, $p < .001$). Thus, H7 was also supported.

Since appreciation, as the only one of the subjective player ratings, was significantly associated with both interactivity and percentage donated, a mediation analysis was performed to explore whether appreciation mediated the effect of interactivity on percentage donated. To this end, two path-models were set up, as seen in Figure 2. The first path model examined the direct effect of interactivity on percentage donated, while the second path model included appreciation as a mediator variable. As had already been found in the ANOVA, the first path model revealed a significant direct effect of interactivity on percentage donated ($\beta = .14$, b = 11.76, SE = 5.298, t = 2.20, $p = .026$). The second path model revealed significant paths from interactivity to appreciation ($\beta = .17$, b = .49, SE = .182, t = 2.71, $p = .007$) and appreciation to percentage donated ($\beta = .23$, b = 6.69, SE = 1.86, t = 3.561, $p < .001$), while the path from interactivity to percentage donated

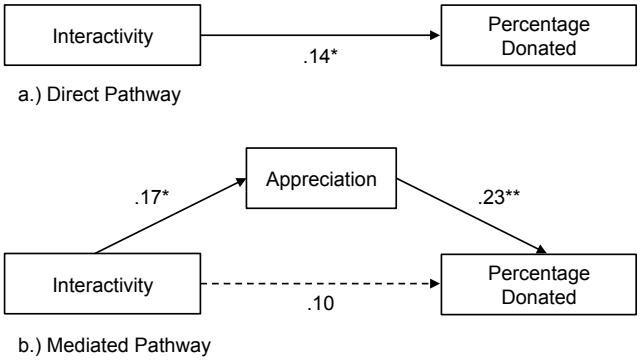

Figure 2. The relationship between interactivity and percentage donated, fully mediated by appreciation. * $p < .05$. ** $p < .01$.

was now no longer significant ($\beta = .104$, b = 8.46, SE = 5.24, t = 1.62, $p = .106$), indicating that the effect of interactivity on percentage donated is fully mediated by appreciation.

DISCUSSION

The results of the present study offer further support for the findings of previous research (e.g., [31, 34, 38]) on the effectiveness of games for change, while additionally providing insight into the individual contributions of specific game properties. Furthermore, for the first time effectiveness was examined using both a behavioral measure and ratings of player attitudes and experience, shedding light on the relationship between subjective ratings and donating behavior, as influenced by games for change.

With 63%, participants in the interactive conditions donated a significantly higher percentage of their one dollar bonus than the 51% donated by participants in the noninteractive conditions. Presentation mode on the other hand did not have a significant impact on the percentage participants chose to donate. However, presentation mode did significantly impact role-taking. Specifically, participants showed significantly higher role-taking in the interactive text condition than in the other conditions. Role-taking was also significantly correlated with willingness to help. This is in line with previous findings by Peng et al. [34].

Additionally, the interactive conditions led to significantly more enjoyment and appreciation than the noninteractive conditions, whereas presentation mode had no significant impact on either enjoyment or appreciation. While enjoyment was

not associated with a higher percentage donated, appreciation was. Furthermore, appreciation fully mediated the higher percentage donated in the interactive over the noninteractive conditions.

Due to the systematic experimental manipulation of interactivity across the three presentation modes, conclusions may now be drawn as to their direct effects on the examined dependent variables. Across all dependent variables save for willingness to help and role-taking, interactivity was the relevant game property, while presentation mode had no significant impact. This means that percentage donated, enjoyment, and appreciation were all significantly increased by making participants interact with the media they were consuming. This is in line with previous research on the importance of interactivity as a game property [19, 24, 34, 37, 38], however, this was the first study to examine the effect of interactivity while controlling for presentation mode.

In further support of the findings of Peng et al. [34], higher willingness to help correlated with a higher percentage donated, indicating that willingness to help is related to actual donating behavior. However, the correlation was fairly weak, indicating that other factors besides willingness to help may be involved in the decision to donate. While there was no significant main effect for interactivity on either willingness to help or role-taking, interactivity did have an interesting impact, in the respect that there was a significant positive correlation between role-taking and percentage donated and willingness to help and percentage donated for the noninteractive conditions, but not for the interactive conditions. At the same time, participants in the interactive conditions donated almost 12% more than in the noninteractive conditions. In other words, interactivity seemed to invoke a higher percentage donated regardless of participants' reported willingness to help or role-taking. Beyond further establishing the importance of interactivity, this further implies that role-taking and willingness to help are not the only relevant factors related to increasing percentage donated.

One of these relevant factors appears to be appreciation, which was not only increased by interactivity, but also fully mediated the relationship between interactivity and percentage donated. While the nature of the present study does not allow for any causal inferences, this may suggest that participants found a narrative that they could actively participate in more meaningful than a narrative they were passively consuming and this then possibly encouraged them to donate a larger percentage of their bonus. This is in line with previous research on the ability of games to be thought-provoking [8, 21, 25] and findings on the relationship between meaningfulness and prosocial behavior [2, 3, 30]. However, this is to our knowledge the first study to find evidence for a potential connection between game properties, appreciation, and prosocial behavior. These findings indicate the importance of including appreciation in the examination of the effectiveness of games for change, as well as highlighting its potential for encouraging prosocial behavior, such as donating.

Enjoyment was increased by interactivity, but was not directly associated with a higher percentage donated. This means that while participants clearly enjoyed the interactive conditions more than the noninteractive conditions, this did not necessarily make them donate a higher percentage. This recalls Cohen's findings [12], who found that while enjoyment increased the intention to share a game for change, it was not associated with a higher likelihood of later (self-reported) sharing. A caveat for this study however, is that while interactivity increased enjoyment, the primary goal of the game makers had most likely not been to make the narrative of Darfur is Dying enjoyable [31]. It would be interesting to examine games for change with more enjoyable narratives and explore whether for these games enjoyment might be more likely to impact behavior, as well as investigating how more enjoyable narratives impact appreciation.

Presentation mode did not have a significant impact on any of the dependent variables, except for role-taking, where the interactive text increased role-taking more than the other conditions. That role-taking would be higher for a text condition than for other, more visual conditions may point to the ability of a text to convey the character's thoughts and emotions better than the cartoon-like animation of Darfur is Dying, making it easier to identify with the character when reading the text. Furthermore these findings also indicate that the multimodality of the presentation mode does not play a significant role in increasing factors relevant to the effectiveness of games for change, such as willingness to help, enjoyment, appreciation, or donating. This sets games for change apart from games for education where multimodality had been found to significantly impact knowledge gain and interest in learning [37]. Considering research on the positive impact of multimodality on information processing [40], it is perhaps not surprising that multimodality was more effective in the context of learning.

Implications, limitations and further research

For game designers and organizations aiming to create games for change, the main takeaway from this study is that while interactivity is crucial for the effectiveness of games for change to encourage donating, presentation mode is seemingly less important. Strikingly, in this study this meant that using a simple interactive text was almost exactly as effective at motivating participants to donate, as the video game Darfur is Dying. However, an important limitation of this study is that the effectiveness of the game only refers to the behavior of players after being prompted to play the game. It is very possible that while presentation mode may not be important for increasing donating behavior, it may increase the likelihood of a player noticing or seeking out a game, as a video game may look more interesting than a text-based game. Likewise, participants were instructed to play, watch, or read through the conditions until the end. While none of the conditions took longer than seven minutes to complete, it is possible that without the context of the study, participants may have been more likely to stop playing in the text or text with pictures conditions than in the (recorded) gameplay conditions. Finally, the short play time utilized in this game means that these findings may not be generalizable to games played over a longer course of time. However, the finding that even such

a short play time could lead to significant changes in participant attitudes, experiences, and behavior, is an interesting finding in and of itself. Together, these findings suggest that while further research is necessary to understand the potential and limitations of text-based interactive fiction as games for change, this may be an area worth exploring both in future research and practical work.

Perhaps more importantly, the present study identified appreciation as a potential component of the effectiveness of games for change, as showcased by its mediating the relationship between interactivity and donating behavior. It remains to be seen what game properties other than interactivity may potentially inspire player appreciation and how this subsequently relates to various prosocial behaviors, including but not limited to donating. Iacovides and Cox [21] mention narrative, gameplay, and audio as factors that helped create a meaningful and thought-provoking experience in a game illustrating the dilemmas facing health professionals. Bartsch et al. [6] highlight the role that moving music can play in evoking appreciation for a film. Another interesting approach was highlighted in recent research by Gerling et al. [20], who utilized embodied interaction, which had participants controlling a digital game about living with disabilities by sitting in and operating a wheelchair themselves. This embodied interaction lead participants to reflect more on real-world challenges facing people with disabilities than participants controlling the same game by traditional gamepad. Considering the current findings, further research could investigate how these and other game properties impact appreciation and prosocial behavior in the context of games for change.

A further limitation is that this study only examined results for participants who lost the game. Past research has found that success and failure can lead to considerably different affective responses [32]. In sports, winning has been associated with more positive affect, while losing is more likely to lead to a negative affective response [43]. Considering past findings that media is especially appreciated when it evokes mixed emotions (e.g., [5, 8, 21]), it would be interesting to see how appreciation differs depending on whether participants win or lose at a game for change. Further research on how game outcome and the consequent emotional response impact appreciation and its connection to prosocial behavior could help improve understanding of its importance for games for change.

Having participants donate their bonus was successful in showing differences in prosocial behavior depending on the experimental condition experienced. However, while Clark [11] found that participants tend to give similar amounts of their own money in comparison to an unexpected sum of money given to them during a study, people may still take other criteria into consideration when donating their money to an organisation in a real-world setting than when donating a bonus in a study. Therefore future research should strive to examine donating behavior and other forms of prosocial behavior using more realistic measurements and settings.

Finally, an obvious limitation of this study was the use of only three modes of presentation. These represented only a very small spectrum of the presentation modes possible in the design of games. While using an animated cartoon-like gameplay may not increase appreciation, enjoyment, willingness to help, or percentage donated compared to an interactive text conveying the same information, these same variables might behave quite differently for other presentation modes, not examined in this study, such as photo-realistic graphics. Similarly, it is possible that cartoon-like presentations might not necessarily be the presentation mode best suited for the context of games for change, but might have a different impact on factors such as enjoyment or other behavioral variables in another context such as games for education or pure entertainment games. Therefore, to explore if these findings can be generalized to other presentation modes and game genres, further research is necessary.

CONCLUSION

The findings indicate that between presentation mode and interactivity, interactivity is the more important property in driving the effectiveness of games for change in increasing enjoyment, appreciation, and donating behavior. Interactive conditions also increased donating independently of role-taking or willingness to help in comparison to the noninteractive conditions. Interestingly, interactivity increased the appreciation players felt for the story being told, which in turn fully mediated the effect of interactivity on donating behavior. Role-taking was the only one of the variables studied, which was significantly affected by presentation mode. It is important however, to keep in mind that these findings do not allow inferences as to the effectiveness in winning or holding players' attention, as well as how these factors may change over a longer play time. It is also possible that other more effective presentation modes exist that were not included in this study. More research is necessary to examine how far these results can be generalized. Finally, the crucial role that appreciation played in mediating the relationship of interactivity on donating behavior, suggests the potential of appreciation as a promising addition to future research on games for change.

ACKNOWLEDGEMENTS

We would like to thank the reviewers for their very helpful comments. Special thanks also go to the members of our research department for providing valuable discussions and feedback throughout the course of this research project.

REFERENCES

1. Ajzen, I. *Attitudes, personality and behaviour*. McGraw-Hill International, 2005.
2. Andersen, M. H., Mathisen, L., Øyen, O., Wahl, A. K., Hanestad, B. R., and Fosse, E. Living donors' experiences 1 wk after donating a kidney. *Clinical transplantation 19*, 1 (2005), 90–96.
3. Apter, M. J., Spirn, N., Sveback, S., and Apter, M. Motives for donating blood. *Stress and health: A reversal theory perspective* (1997), 145–156.
4. Baddeley, A. Working memory. *Science 255*, 5044 (1992), 556–559.

5. Bartsch, A., and Hartmann, T. The role of cognitive and affective challenge in entertainment experience. *Communication Research* (2015), 0093650214565921.
6. Bartsch, A., Kalch, A., and Oliver, M. B. Moved to think: The role of emotional media experiences in stimulating reflective thoughts. *Journal of Media Psychology: Theories, Methods, and Applications 26*, 3 (2014), 125.
7. Bogost, I. *Persuasive games: The expressive power of videogames*. Mit Press, 2007.
8. Bopp, J. A., Mekler, E. D., and Opwis, K. It was sad but still good: Gratifications of emotionally moving game experiences. In *Ext. Abstracts CHI '15*, ACM (2015), 1193–1198.
9. Chatterjee, P., Rose, R. L., and Sinha, J. Why money meanings matter in decisions to donate time and money. *Marketing Letters 24*, 2 (2013), 109–118.
10. Cho, J. Disentangling media effects from debate effects: The presentation mode of televised debates and viewer decision making. *Journalism & Mass Communication Quarterly 86*, 2 (2009), 383–400.
11. Clark, J. House money effects in public good experiments. *Experimental Economics 5*, 3 (2002), 223–231.
12. Cohen, E. L. What makes good games go viral? the role of technology use, efficacy, emotion and enjoyment in players' decision to share a prosocial digital game. *Computers in Human Behavior 33* (2014), 321–329.
13. Cohen, J. Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication & Society 4*, 3 (2001), 245–264.
14. Coutu, W. Role-playing vs. role-taking: An appeal for clarification. *American Sociological Review* (1951), 180–187.
15. Davis, M. H. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology 44*, 1 (1983), 113.
16. DeScioli, P., Massenkoff, M., Shaw, A., Petersen, M. B., and Kurzban, R. Equity or equality? moral judgments follow the money. *Proceedings of the Royal Society B: Biological Sciences 281*, 1797 (2014).
17. Eisenberg, N., and Miller, P. A. The relation of empathy to prosocial and related behaviors. *Psychological bulletin 101*, 1 (1987), 91.
18. Fazio, R. H. Multiple processes by which attitudes guide behavior: The mode model as an integrative framework. *Advances in experimental social psychology 23* (1990), 75–109.
19. Gee, J. P. Deep learning properties of good digital games: How far can they go. *Serious games: Mechanisms and effects* (2009), 67–82.
20. Gerling, K. M., Mandryk, R. L., Birk, M. V., Miller, M., and Orji, R. The effects of embodied persuasive games on player attitudes toward people using wheelchairs. In *Proc. CHI '14*, ACM (2014), 3413–3422.
21. Iacovides, I., and Cox, A. L. Moving beyond fun: Evaluating serious experience in digital games. In *Proc. CHI '15*, ACM (2015), 2245–2254.
22. Klimmt, C. *Serious games and social change: Why they (should) work*. Routledge, 2009, 248–270.
23. Klimmt, C., Hartmann, T., and Frey, A. Effectance and control as determinants of video game enjoyment. *Cyberpsychology & behavior 10*, 6 (2007), 845–848.
24. Lin, J.-H. Do video games exert stronger effects on aggression than film? the role of media interactivity and identification on the association of violent content and aggressive outcomes. *Computers in Human Behavior 29*, 3 (2013), 535–543.
25. Marsh, T., and Costello, B. Experience in serious games: between positive and serious experience. In *Serious Games Development and Applications*. Springer, 2012, 255–267.
26. Mayer, R. E. Multimedia learning and games. *Computer games and instruction* (2011), 281–305.
27. Mekler, E. D., Bopp, J. A., Tuch, A. N., and Opwis, K. A systematic review of quantitative studies on the enjoyment of digital entertainment games. In *Proc. CHI '14*, ACM (2014), 927–936.
28. Mitgutsch, K., and Alvarado, N. Purposeful by design?: a serious game design assessment framework. In *Proc. FDG '12*, ACM (2012), 121–128.
29. Montfort, N. *Twisty Little Passages: an approach to interactive fiction*. Mit Press, 2005.
30. Myrick, J. G., and Oliver, M. B. Laughing and crying: Mixed emotions, compassion, and the effectiveness of a youtube psa about skin cancer. *Health communication*, ahead-of-print (2014), 1–10.
31. Neys, J., and Jansz, J. Political internet games: Engaging an audience. *European Journal of Communication 25*, 3 (2010), 227–241.
32. Nummenmaa, L., and Niemi, P. Inducing affective states with success-failure manipulations: a meta-analysis. *Emotion 4*, 2 (2004), 207.
33. Oliver, M. B., and Bartsch, A. Appreciation as audience response: Exploring entertainment gratifications beyond hedonism. *Human Communication Research 36*, 1 (2010), 53–81.
34. Peng, W., Lee, M., and Heeter, C. The effects of a serious game on role-taking and willingness to help. *Journal of Communication 60*, 4 (2010), 723–742.
35. Ratan, R., and Ritterfeld, U. Classifying serious games. *Serious games: Mechanisms and effects* (2009), 10–24.

36. Ritterfeld, U., Cody, M., and Vorderer, P. *Serious games: Mechanisms and effects*. Routledge, 2009.

37. Ritterfeld, U., Shen, C., Wang, H., Nocera, L., and Wong, W. L. Multimodality and interactivity: Connecting properties of serious games with educational outcomes. *Cyberpsychology & Behavior 12*, 6 (2009), 691–697.

38. Ruggiero, D. The effect of a persuasive social impact game on affective learning and attitude. *Computers in Human Behavior 45* (2015), 213–221.

39. Stoianova, V. Private funding: An emerging trend in humanitarian donorship. *Global Humanitarian Assistance* (2012).

40. Sweller, J., Van Merrienboer, J. J., and Paas, F. G. Cognitive architecture and instructional design. *Educational psychology review 10*, 3 (1998), 251–296.

41. Tsvetkova, M., and Macy, M. W. The social contagion of generosity. *PloS one 9*, 2 (2014), e87275.

42. Van Wassenhove, L. N. Humanitarian aid logistics: supply chain management in high gear. *Journal of the Operational Research Society 57*, 5 (2006), 475–489.

43. Wilson, G. V., and Kerr, J. H. Affective responses to success and failure:: a study of winning and losing in competitive rugby. *Personality and Individual Differences 27*, 1 (1999), 85–99.

44. Wouters, P., Van Nimwegen, C., Van Oostendorp, H., and Van Der Spek, E. D. A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology 105*, 2 (2013), 249.