



Balancing Game Rules for Improving Creative Output of Group Brainstorms

Niko Vegt *, Valentijn Visch, Arnold Vermeeren, Huib de Ridder, and Zsolt Hayde

Delft University of Technology, Delft, the Netherlands

This article describes a user-centered design experiment investigating positive and negative effects of adding game rules to brainstorming. We studied effects on brainstorm output and user experience and behavior. A coin-based gamification was developed with rules intended to improve brainstorm output in relation to quality and quantity of ideas. However, the invasiveness of a gamification can be expected to affect users both positively and negatively. To find an optimum between positive and negative effects of gamification invasiveness, we tested 5 different rule-sets with varying quantity and quality of rules. The results demonstrated that game rules stimulating competitive game behavior improved the quantity and quality of brainstorm output. Yet the invasiveness of the gamification also hindered this positive effect, due to discussions about rules and mandatory game behavior. From these results we deduced 3 types of invasiveness evoked by the rules' qualities: a) governing rules led to negative cognitive invasiveness, b) forcing rules caused positive as well as negative behavioral invasiveness, and c) adding coins may have led to positive affective invasiveness (i.e., a playful attitude). We conclude our study with recommendations on designing and researching gamification invasiveness in real-life contexts.

Keywords – Gameful Design, Game Rules, Group Brainstorming, Invasiveness, User-centered Gamification.

Relevance to Design Practice – Our research results give insight to user behaviors and experiences within a gamified context, demonstrating dos and don'ts for gamification design. Additionally, the research addressed group brainstorming, providing a method for improving brainstorm output.

Citation: Vegt, N. J. H., Visch, V., Vermeeren, A. P. O. S., de Ridder, H., & Hayde, Z. (2019). Balancing game rules for creative output in group brainstorming. *International Journal of Design*, 13(1), 1-19.

Introduction

Recent studies suggest that teamwork could benefit from gamification, i.e., the use of game elements in non-game contexts (Deterding, Dixon, Khaled, & Nacke, 2011). For example, game-like incentive mechanisms are introduced that award participants with points for certain activities, which lead to positive effects of increased productivity (Mekler, Brühlmann, Opwis, & Tuch, 2013) and collaborative behavior (Moccozet, Tardy, Opprecht, & Léonard, 2013). These interventions are promising, yet they adopt a rather narrow research approach, only demonstrating the effect of single game elements, like awarding points for particular activities.

Gamification research should broaden its scope to understand and optimize its potential effects properly. Moreover, unintended negative effects should be addressed. Knaving and Björk (2013) suggest that an inappropriate underlying model or mandatory game actions can easily overshadow the non-game goal of a gamification. Yet gamification research generally does not address negative effects (Hamari, Koivisto, & Sarsa, 2014). Most experiments take a systemic perspective in which the introduced game element and its intended effect are leading. This often leads to no or weak results because games are complex systems in which player behavior is hard to predict.

To explore the full potential of gamification, a user-centered perspective (Huotari & Hamari, 2012; Visch, Vegt, Anderiesen, & van der Kooij, 2013) seems more fruitful because the user's

behaviors and experiences essentially define its effect. Hence, we take user experience and behavior as the starting point for the development of a gamification. To understand users in a gamified situation better, Visch et al. (2013) suggest that a user's experience lies on a continuum between the extremes of a real world experience and game world experience (see Figure 1). By adding game elements to a real world situation, such as winning coins in a brainstorm meeting, the user's experience may be transported (Green, Brock, & Kaufman, 2004) towards a game world experience (i.e., the feeling of playing a game). In gamified contexts, the resulting gameful experience (McGonigal, 2011) could then affect the non-game situation, which is referred to as the transfer effect. For example, in the case of adding coins to a business meeting, the playfulness of shoving coins to each other could affect the tone of a conversation, which could possibly lead to a transfer effect of more creative outcomes.

Received June 19, 2017; Accepted August 5, 2018; Published April 30, 2019.

Copyright: © 2019 Vegt, Visch, Vermeeren, de Ridder, & Hayde. Copyright for this article is retained by the authors, with first publication rights granted to the *International Journal of Design*. All journal content, except where otherwise noted, is licensed under a *Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License*. By virtue of their appearance in this open-access journal, articles are free to use, with proper attribution, in educational and other non-commercial settings.

*Corresponding Author: n.j.h.vegt@tudelft.nl

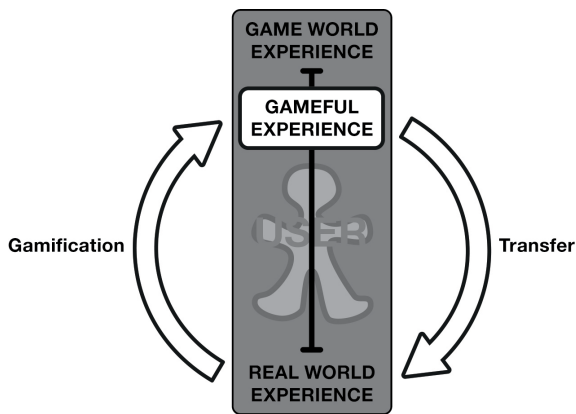


Figure 1. Model based on (Visch et al., 2013).

Niko Vegt is postdoctoral researcher at the faculty of Industrial Design Engineering of the Delft University of Technology, the Netherlands. His research typically follows a research through design approach with a focus on social interaction. His current research project is concerned with interactive storytelling design for obesity prevention. In June 2018, Niko received a PhD on the design and application of game elements in teamwork contexts. As part of this research, he developed gamification designs for management consultancy firms and a steel galvanizing factory. Moreover, he developed and facilitated gamified workshops at companies and conferences. Next to the research, Niko also works as an educator and coach in a variety of design courses at the faculty of Industrial Design Engineering at Delft University of Technology.

Valentijn Visch is associate professor at the faculty of Industrial Design Engineering of the Delft University of Technology, the Netherlands. He conducts and supervises theoretical, empirical, and design research in the area of Persuasive Game Design, often with mental healthcare as application domain. His design research is creative, user-experience centered and typically involves multidisciplinary teams like creative industry, healthcare institutions, patients, medical practitioners, and researchers from various domains. Valentijn chairs the eHealth Design lab at the TU Delft and has a background in literature (MA), art theory (MA), animation (NIAf), cognitive film studies (PhD), and emotion research (PD). Valentijn recently coordinates a project on storytelling as a persuasive design tool for eHealth.

Arnold Vermeeren was educated as an industrial designer and received a PhD on methodological issues of usability testing at TU Delft, the Netherlands. He is Associate Professor at the Faculty of Industrial Design Engineering (TU Delft) and Director of the Faculty's MuseumFutures lab. He was a member of the management teams of two EU-Cost actions focussing on methods for usability and user experience, and supervised PhD candidates in Dutch nationally funded projects on user experience in various application areas such as crowd experiences, gamification of work, and involvement of children in experience design activities. His current research in the MuseumFutures lab focuses on experience design for creating relevance to new museum audiences.

Huib de Ridder is full professor of Informational Ergonomics at the Faculty of Industrial Design Engineering, Delft University of Technology, Delft, the Netherlands. His main research objective is understanding human behavior in context, with a focus on human perception and human information processing. May 2008, he co-founded the Delft Perceptual Intelligence Lab (π -Lab) to investigate (multi-) sensory fluency. He is (co-) author of more than 250 scientific publications in such diverse fields as information presentation, image quality, picture perception, light, space and material interaction, vision-based gesture recognition, user understanding, dual tasking, assessment methodologies, ambient intelligence, social connectedness, crowd well-being, gamification of work, involvement of children in experience design activities. Recently, he got involved with the Holland Proton Therapy Center, a new Delft-based clinic/research institute for cancer treatment, to develop human/patient-centric technologies for healthcare settings.

Zsolt Hayde is currently technical support engineer at Mimaki Europe B.V. In 2017, he received his Masters degree on developing a modular controller for Xbox One to increase lifespan and personal attachment, at the faculty of Industrial Design Engineering at the Delft University of Technology, the Netherlands. During his study he worked as a research intern for the study that is described in this article.

In the present article, we examine the transfer effect of a structured series of gamifications for group brainstorming. Group brainstorming is common practice in many disciplines and generally the process is still based on the four original rules introduced by Osborn (1963): 1) withhold criticism, 2) welcome unusual ideas, 3) combine and improve ideas, and 4) focus on quantity. However, since then, research gained many new insights, suggesting that Osborn's rules are not sufficient for productive brainstorm sessions (Sutton & Hargadon, 1996). Throughout the years, many methods have been developed to improve group brainstorms. For example, gamestorming is a method that generates game worlds in which participants can jointly produce ideas (Gray, Brown, & Macanuso, 2010). Instead of preparing a gamestorm session, just adding game elements to a brainstorm might lead to an equally beneficial game world experience.

To evoke a game world experience, rules are commonly mentioned as an essential element (Caillois, 1961; Juul, 2003; McGonigal, 2011). By adding game rules to a group brainstorm, group members get an increased feeling of playing a game (i.e., gameful experience). This could lead to enhanced creativity since creativity is commonly related to intrinsic motivational states (Amabile, 2012) and game rules generally derive their motivational power by tapping into intrinsic needs, i.e., autonomy, competence and social relatedness (Ryan, Rigby, & Przybylski, 2006). Moreover, rules can stimulate behavior that is beneficial for brainstorm output (see Figure 2). However, according to Figure 1, if game rules move a user's experience towards the game world, they move it away from the real world. Increased engagement with the gamification could therefore be expected to reduce the user's engagement with the brainstorm (see Figure 2). Consequently, an optimum between positive and negative effects of adding invasive game rules needs to be found, and this will be researched in the present study.

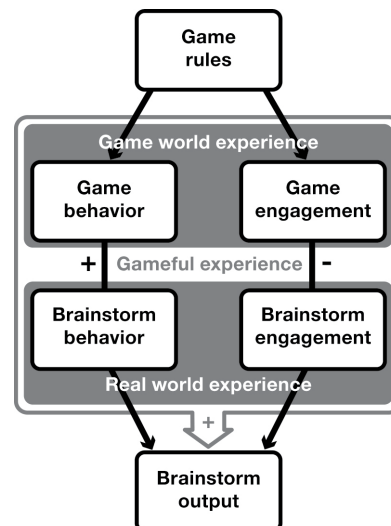


Figure 2. Initial model for the gamification of group brainstorming.

In order to find such an optimal configuration of rules, this article describes an experiment examining the transfer effect of varying rule-sets in a coin-based gamification for brainstorm

sessions. The article is organized as follows. First we look into group brainstorming and game design literature to specify the initial model further (Figure 2) in a research framework that defines possible transfer effects (see Figure 3). Based on the framework, an experiment was set up in which rules were clustered into different rule-sets with varying levels of invasiveness. The results of the experiment are used to revise the research framework (Figure 14) and to develop a model that addresses positive and negative effects of invasive game rules in brainstorm sessions (Figure 15).

Theoretical Background

Game Rules

In relation to invasiveness, game design literature commonly distinguishes explicit and implicit rules (Bergström, 2010; Björk & Holopainen, 2003; Salen & Zimmerman, 2004; Sniderman, 2006). Explicit rules are the *official* rules of a game, whereas implicit rules are referred to as *unwritten rules*. When distinguishing game rules from a systems perspective, explicit rules are more invasive than implicit rules, because explicit rules ask for conscious processing and enforcement. However, from a player perspective, implicit rules can define the invasiveness of a game as well, for example in Hagoo, a social game in which two players need to frown while walking towards each other with continuous eye contact (Fluegelman, 1976). In this game, explicit rules define the winning condition and positioning of the players. Yet breaking the implicit social rule of *staring is impolite* makes Hagoo interesting and engaging. In computer games, the official rules generally are not perceived as rules because they are mostly embedded into virtual objects (Järvinen, 2008). Yet these embedded rules, such as physical obstacles or a withdrawing progress bar, can strongly influence game invasiveness.

Thus rather than distinguishing rules from a systems perspective, adopting a user-centered approach seems more useful for understanding and optimizing the effects of gamification. From a user perspective, a large quantity of rules might already cause stronger invasiveness as to coping with them. Moreover, the quality of rules could influence the invasiveness of a gamification. Knaving and Björk (2013) stress that users should be free to interact with the gamification and mandatory actions should be particularly meaningful. Thus the level of obligation of rules might influence the invasiveness of a gamification as well.

Avedon (1971) describes two categories of rules that vary in obligation from a designer perspective: *procedures for action* and *rules governing action*. Procedures for action force actions upon players, such as drawing a *go to jail* card in Monopoly. Conversely, rules governing action just define conditions, such as the rules in chess that prescribe the movements chess pieces can make. These examples demonstrate that governing rules often require a player to make decisions, whereas procedures are just to be followed by the player. As a result, procedures (i.e., forcing rules) are generally less cognitively demanding than governing rules and can be expected to be less invasive in a group-brainstorming task. Moreover, forcing rules may directly stimulate behavior that is beneficial for brainstorm output, so they

seem most suitable for the gamification of group brainstorming. On the other hand, unlike forcing rules, governing rules can be flexibly used in favor of attaining a real world goal, such as producing many ideas. Governing rules provide players with playing freedom that is commonly assumed to be crucial for a game experience (Caillois, 1961; Juul, 2003; McGonigal, 2011; Suits, 1978), thereby evoking intrinsic motivation that is beneficial for creativity (Amabile, 2012).

In conclusion, when designing a gamification, the obligation of rules can be used to achieve varying rule qualities that we assume could be beneficial for the outcome of a brainstorm meeting. Governing rules could positively influence brainstorm output by evoking a more gameful experience in participants and forcing rules could evoke beneficial behavior.

Group Brainstorming

As explained in the introduction, the original group brainstorming method (Osborn, 1963) contained four rules to achieve maximal creativity and creative productivity through interaction between individuals. Recent studies on group brainstorming suggest that these rules are not sufficient. Thus to develop game rules that benefit brainstorm output, we needed to gain a better understanding of the elements that influence the quantity (i.e., productivity) and quality of brainstorm output.

Productivity

As Osborn assumed, exposure to ideas of others indeed positively influences brainstorm productivity (Dugosh & Paulus, 2005). However, many studies on group brainstorming consistently observe that groups are less productive than individuals when generating ideas. Literature suggests several reasons for this group productivity gap (Sutton & Hargadon, 1996; Brown, Tumeo, Larey, & Paulus, 1998). One explanation is that group members avoid expressing ideas because they worry about the opinion of others (i.e., evaluation apprehension). It is also suggested that individuals in groups do not feel accountable for producing ideas and as a result devote less effort to it (i.e., free-riding). Another explanation is that team members tend to match their productivity to members that generate fewer ideas (Brown & Paulus, 1996). Moreover, research suggests that individuals overestimate their productivity in group-brainstorms. They tend to claim more ideas than they actually produce (Paulus, Dzindolet, Poletes, & Mabel Camacho, 1993). The strongest support exists for the fact that, compared to working alone, waiting for your turn to talk as well as listening to others can hamper idea generation (i.e., production blocking). In fact, Nijstad, Stroebe, and Lodewijkx (2006) demonstrate that production blocking partly explains overestimation of one's productivity. Waiting for and listening to others leads to fewer new ideas, yet also to fewer failures, and reduction of failures mostly influences one's satisfaction.

In order to benefit from group brainstorming, *attention to each other's ideas* and *performance feedback* are mostly mentioned as crucial factors (Brown et al., 1998). Paulus, Putman, Dugosh, Dzindolet, and Coskun (2002) suggest that the group process needs

to maximize exchange of ideas and minimize distracting effects of task-irrelevant discussions. Members should be encouraged to be attentive to ideas of other group members while they are sharing them. Performance feedback can reduce free riding and overestimation (Harkins & Szymanski, 1989). Monitoring each other's tasks was found to improve coordination and feedback processes (Marks & Panzer, 2004). Yet individual performance feedback could also lead to evaluation apprehension. Jung, Schneider, and Valacich (2010) suggest that adopting a temporary identity overcomes evaluation apprehension while keeping the benefits from social comparison. Another solution to maximize idea generation could be to provide feedback on behavioral factors, such as individual speaking time or information exchange. Yet such types of feedback generally lead to moderate performance, because group members tend to adapt their effort to underperforming members and high performing teams can get distracted (DiMicco, Pandolfo, & Bender, 2004; Tausczik & Pennebaker, 2013).

In conclusion, to stimulate productivity in a group brainstorm, Osborn's rules of *combining and improving ideas* and *focusing on quantity* should be refined. To combine and improve ideas, team members should be encouraged to listen carefully to each other's ideas. To assure a focus on quantity, team members should be allowed to stimulate one another by giving positive as well as negative feedback on each other's productivity.

Quality of Ideas

Next to producing many ideas, Osborn's rules (1963) were meant to increase the quality of ideas. Yet rather than *withholding criticism*, conflict is often mentioned as a positive factor for the quality of ideas (Jehn, 1995). Conflicts stimulate team members to think more creatively to resolve the problem that interferes with their goal achievement (Jung & Lee, 2015). Cognitive conflict (i.e., conflict about tasks and goals) is generally found to improve creative team performance, whereas affective conflict (interfering with relationships) is generally detrimental (Wu et al., 2015). Yet even affective conflict can be beneficial. Yong, Sauer, and Mannix (2014) demonstrate that if one team member experiences affective conflict and the others do not, it positively influences novelty of ideas. However, controlling affective conflict is difficult, thus it is generally recommended to avoid it.

Avoiding affective conflict and encouraging cognitive conflict may be achieved by framing discussions towards learning instead of performance goals (Huang, 2010) and by emphasizing on collective instead of individual goal achievement (Deutsch, 2006). The extent to which cognitive conflict is beneficial depends on the way groups deal with it. Generally, when groups adopt a collaborative conflict behavior style (i.e., striving for a win-win solution) they are found to arrive at the most successful outcomes (Weingart & Jehn, 2003). However, in creative group processes, adopting a pure collaborative style seems less effective. Badke-Schaub, Goldschmidt, and Meijer (2010) demonstrate that design teams that exhibit relatively more competitive conflict behavior produce more new ideas and are more associative, leading to high innovation and functionality in design concepts.

In conclusion, Osborn's rule of *withholding criticism* requires refinement to lead to high quality brainstorm output. Indeed, personal or emotional criticism should be avoided, yet criticism in relation to the collective output should be allowed, because competition among group members can stimulate the quality and quantity of brainstorm output.

Research Framework

Based on the above-described insights, we developed a framework for the gamification of group brainstorms (see Figure 3). Group brainstorming literature provides a variety of ways to improve brainstorm processes and two factors stand out that could be addressed well by gamification: individual performance feedback and competitive behavior. By introducing a game-like feedback mechanism in which participants judge each other on their contribution to the brainstorm, they may become more attentive to each other's ideas. Moreover, games often are competitive systems of conflict (Salen & Zimmerman, 2004) providing many mechanics that could stimulate competitive conflict behavior, such as opponents competing for *the best idea* or a zero-sum score distribution in which only one idea *survives*.

As described by the framework (Figure 3), forcing rules can be used to achieve a positive transfer effect on the quality of brainstorm output by evoking competitive game behavior that stimulates competitive conflict behavior. For a positive transfer effect on productivity, governing rules need to provide a system of individual performance feedback that stimulates attention to each other's ideas. However, as explained in the introduction, we also expect a negative transfer effect of the quantity of rules within a gamification layer. With an increasing number of game rules, participants are expected to pay more attention to the gamification and reduce their attention to the brainstorm, leading to reduced output quality and quantity. In other words, increased engagement in the game world probably weakens the intended positive transfer effect in the real world.

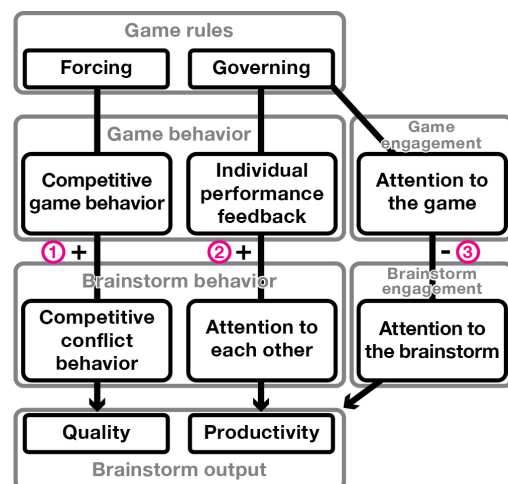


Figure 3. Expected transfer effects of game rules on brainstorm output. 1) Positive effect on quality by evoking competition, 2) positive effect on productivity by evoking attention to each other, 3) negative effect on brainstorm output due to distraction by the game.

Study Design

To find the optimal invasiveness of a gamification layer, we developed a game with gold-colored coins (see Figure 5) in which participants reward and punish each other for their contribution to the brainstorm by giving and taking coins. Based on the research framework, the game rules were designed to stimulate behavior that would improve brainstorm output. After designing the rules of the brainstorm gamification (described below), 5 rule-sets were defined to achieve variance in invasiveness (see Table 1). The rule-sets increased in quantity of rules, with rule-set 1 only containing the basic rules and rule-set 5 containing all rules. Moreover, they varied in quality as to the number of governing, forcing, and embodied rules. The increasing quantity of rules across rule-sets was expected to increase the invasiveness of the gamification. In particular rule-set 2 was expected to raise attention to the gamification due to the large number of additional governing rules. Only the additional rule in rule-set 5 was expected to reduce invasiveness.

The rule-sets were randomly assigned to 10 groups, such that each rule-set was played in 2 sessions. As shown in Figure 4, we measured invasiveness of the gamification, brainstorm behavior and engagement, and brainstorm output. During the brainstorm, video recording and observations captured the participants' behavior. At the end of a session, participants filled in a questionnaire about their behavior and engagement. To measure brainstorm output, two experts assessed the number of different ideas, the originality of ideas, and the feasibility of ideas (Amabile & Pillemer, 2012; Christiaans, 2002). One researcher conducted the study by observing the brainstorms live on a monitor and assessing brainstorm output afterwards. A second researcher first assessed the output and observed videos afterwards to avoid biased judgment. To increase the participants' commitment, the five best performing groups would be rewarded with prize money.

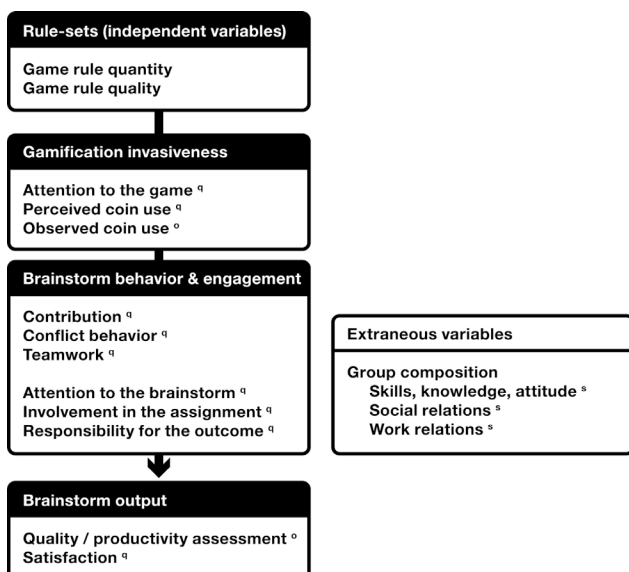


Figure 4. Study variables with rule-sets as independent variable and invasiveness, behavior & engagement and output as dependent variables
(°: questionnaire, °: observation, *: pre-selection).



Figure 5. Design students in a brainstorm session with the coin game (rule-set 4).

Procedure

Ten groups of four design students were asked to participate in a 30-minute brainstorming challenge. One session, including introduction and filling in questionnaires, took approximately 1 hour. Each group was separately welcomed in a neutral room and sat around one table to work on. They received the game material, an instruction sheet and explanation from the researcher about the assessment procedure, distribution of prize money, and the rules of the game. Groups with rule-set 1 received instructions explaining that one could give each other coins for useful contributions to the brainstorm and take coins for criticizing each other's contributions. Groups with subsequent rule-sets received the rules as described in Table 1. The rule of taking coins every 5 minutes was announced through an intercom by the researcher sitting in a separate room. In rule-set 5 participants received cups instead of sheets as pools, so they would not see each other's coins.

After the introduction, participants were asked to sign a consent form, stating that they understood the rules and procedure. Next, they would receive several colored pencils, large flip-over sheets, and a description of the assignment. When participants started to read the assignment, the researcher started a 30-minute timer (not visible for the participants). The assignment consisted of two phases: 1) domain selection, and 2) idea generation. To avoid spending much time on discussing the domain, the groups had to choose a user (elderly, young children, teenagers, professionals), context (hospital, shopping mall, airport, outdoor), and product category (mobility, clothing, tools, entertainment). We assumed that they would choose a domain that would be easy for them, thereby equalizing the influence of prior experiences and interests. After 30 minutes, the participants had to put down their pencils and make final coin transactions if necessary (in rule-sets 4 and 5). Next, they would fill in the questionnaire. When all participants were done, they were debriefed about the purpose of the experiment and again reminded of the procedure for assessing and rewarding their output.

Brainstorm Game Rules

Participants started with 3 coins (rule 1 in Table 1). Pilot tests demonstrated that this number was low enough to make participants carefully consider their transactions and large enough to keep

Table 1. Rules of the brainstorm gamification.

Rules	Rule-set 1	Rule-set 2	Rule-set 3	Rule-set 4	Rule-set 5
1. 3 Coins per participant					
2. Give and take coins					
3. Most coins wins					
a. Most coins in profit wins					
4. Stock pool / profit pool					
a. Receive is profit / take is stock					
b. Own stock to profit prohibited					
c. Take from profit, otherwise stock					
5. Take 2 coins every 5 minutes					
6. Stock pool size limit (4 coins)					
7. Stock must be empty at the end					
8. Pools are blinded					

Note: Light grey: embodied rules; medium grey: governing rules; dark grey: forcing rules.

exchange going. The rules explicitly allowed taking and giving (rule 2), and the winning condition was to own the most coins in the end (rule 3). To retain participants from excessively taking coins from each other in order to win, the winning condition was restricted to coins that were given to someone (rule 3a). To clarify this rule, a stock and profit pool was introduced (rule 4). Only received coins go into the profit pool. The start-coins and taken coins had to go into the stock pool (rule 4a). Consequently, the profit pool resembled one's *score* and the stock pool contained coins that could be used for exchange but would not count as score. Transporting one's own stock coins to one's profit pool was prohibited (rule 4b). Moreover, to keep the exchange going, one had to take coins from the other's profit pool, unless it was empty (rule 4c).

Rule 5 forced the group every five minutes to designate one participant to take two coins. This rule was added because individuals tend to reward good contributions more frequently than they punish bad contributions (Wang, Galinsky, & Murnighan, 2009). As rule 5 would lead to a large increase of coin taking, a size limit was introduced for the stock pool (rule 6) to force participants to give away coins when they had taken many (see Figure 6). Moreover, rule 7 obligated participants to give away all stock coins before the end of the meeting. To avoid

that participants would become too strategic in exchanging coins (e.g., exchange coins to equalize scores or only take coins from high performing participants), they could not see the state of each other's pools (rule 8). In this way, participants were expected to focus more on rewarding or punishing brainstorm contributions instead of winning the game.

Participants

Eight groups consisted of Masters students at the design faculty of Industrial Design Engineering at Delft University of Technology. Two groups contained one and two students with another background. Four groups consisted of only male participants; the other groups were mixed with either predominantly male participants or an equal gender distribution. All participants were familiar with group brainstorming, thus individual experience level and group coherence, important factors for group productivity (Bottger & Yetton, 1987), were expected to be similar. Moreover, all groups consisted of participants that already knew each other, because social factors were expected to play a role. Three groups had worked together before, four groups consisted of friends, and three groups had worked together and were friends.

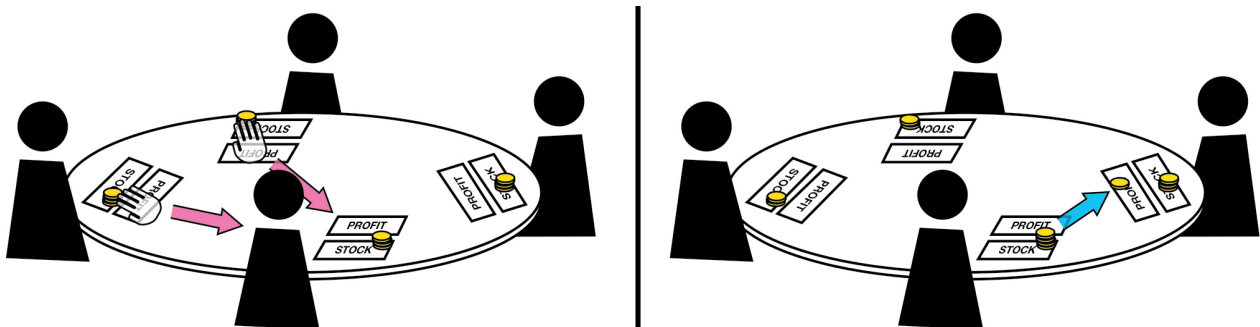


Figure 6. Taking 2 coins obliges the taker to give 1 coin because he owns more than 4.

Measures

Brainstorm Output

Brainstorm output measurement reflected the assessment criteria for the challenge. As described above, two researchers performed the assessment. The number of varying ideas assessed productivity. The quality of ideas was assessed by rating the originality and feasibility (scale 1-10) of the total set of ideas that were drawn on the paper sheets (see Figure 7). For example, brainstorm output containing just one highly original idea would be rated higher than five moderately original ideas. In this way, quality and productivity were separate measures of brainstorm output.

Observations

As explained earlier, one researcher observed the sessions live on a monitor in a separate room. A second researcher did observations through the recorded videos. The main purpose of the observations was to count coin transactions, marked by direction (giving and taking) and obligation (voluntary and obligatory). While counting coin transactions, the researchers also made qualitative notes on other behavior that stood out, such as long discussions about the rules or the vividness of brainstorm discussions.

Questionnaire

After the brainstorm, a questionnaire examined perceived behavior and engagement (see Appendix I). In this questionnaire, first the participants were asked about brainstorm behavior in terms of one's own contribution and contribution of the others. Next, the questionnaire inquired the participants' perception of conflict behavior (criticizing and complementing) and attention to each other (ignoring and noticing). Brainstorm engagement was measured through statements regarding own and others' involvement with the assignment and responsibility for the outcome. Gamification invasiveness was measured through statements about the number of coins used and one's attention to the game. For comparison, the questionnaire also inquired attention to the brainstorm. Next to the main variables, the questionnaire inquired the coin game's added value, satisfaction with brainstorm output, and quality of the teamwork.

Results

The data was analyzed in three steps (see Figure 8). The first step examined the variation in brainstorm output between rule-sets to find effects of the quantity and quality of rules. Secondly, we investigated to what extent the rule-sets led to variation in invasiveness of the gamification and its subsequent effect on brainstorm output. Thirdly, the mediating effect through brainstorm behavior and engagement was analyzed. Before describing the results of each step we describe the data and observations.

Data and Observations

Altogether, the groups produced 112 ideas. The most productive group came up with 19 ideas and the least productive group generated four ideas (see Figure 9). The former group (group 2) quickly started drawing, whereas the latter group (group 7) spent a considerable amount of time on exploring the domain (see mind map in Figure 7a). The participants seemed aware of their own approach, as participants of the productive group reported strong engagement with brainstorm output and participants in the unproductive group reported stronger engagement with the process (see Table 4 in Appendix II).

The variation in quality of ideas among the groups was relatively small. Groups scored on average 6.0 on originality, with the best performing group scoring 8.0 and the worst performing group 4.0. The average score regarding feasibility was 6.6, ranging between 8.5 and 4.0. The two assessments per group generally matched or just varied by 1 point. Only regarding the originality of ideas of group 4, and the originality and feasibility of group 9, assessments deviated more. In both cases, the drawings were unclear and prone to different interpretations.

Observations suggest that most groups used the coins meaningfully. In the first five minutes, half of the groups already exchanged coins voluntarily with comments such as: "good argument, you deserve a coin". On average, ten coins were exchanged voluntarily in a session. In sessions with *forcing* rules, seven obligatory coin transactions were carried out additional to the voluntary transactions. Figure 10 shows the total number of coin transactions during each session, categorized by obligation

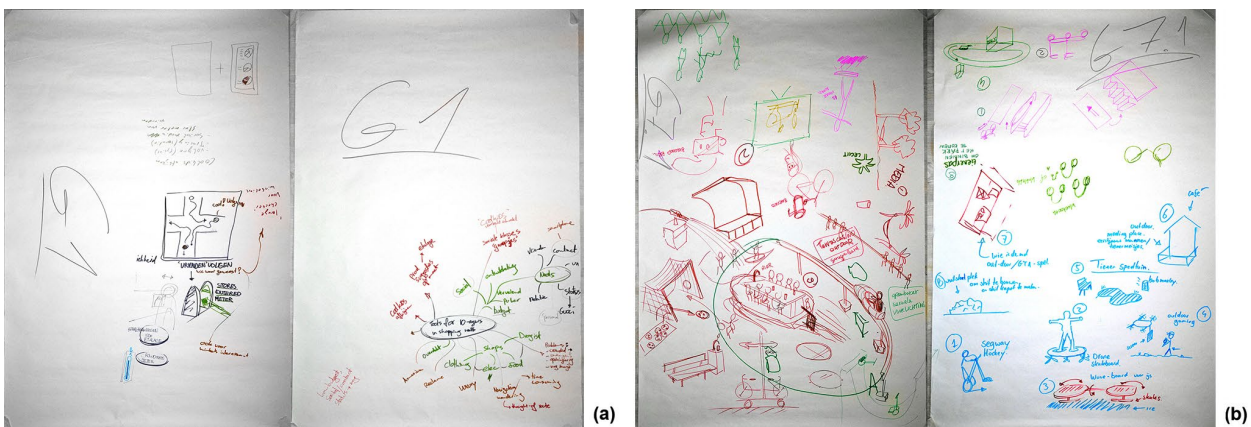


Figure 7. Brainstorm output. a) Least productive group, b) most productive group.

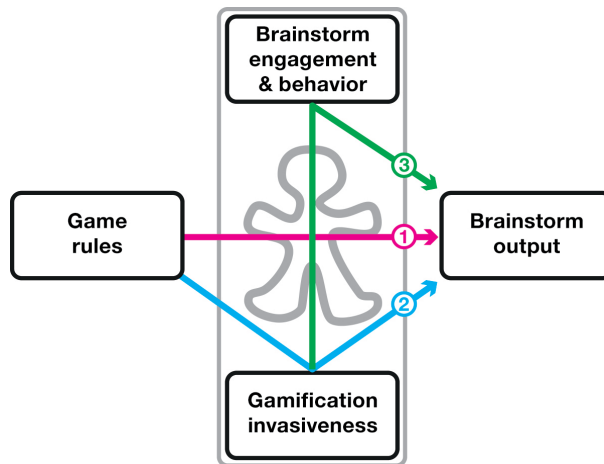


Figure 8. Steps of data analysis. 1) The effect of rule-sets on brainstorm output, 2) the mediating effect of gamification invasiveness, 3) the mediating effect through brainstorm engagement & behavior.

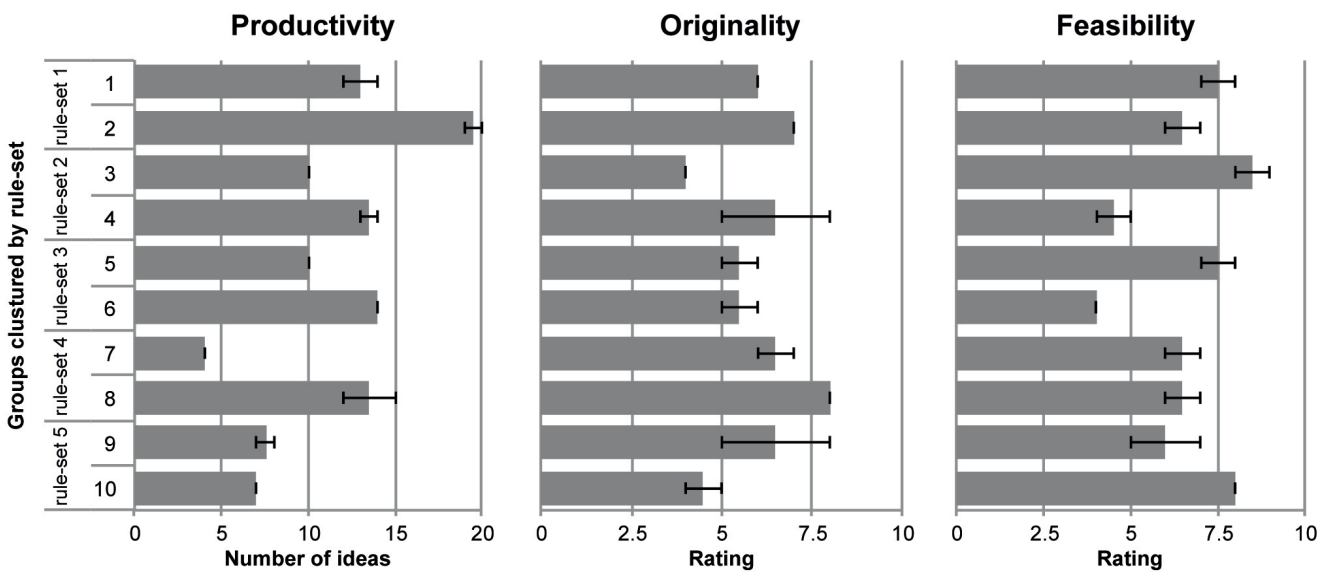


Figure 9. Average brainstorm output per group (error bars: SE).

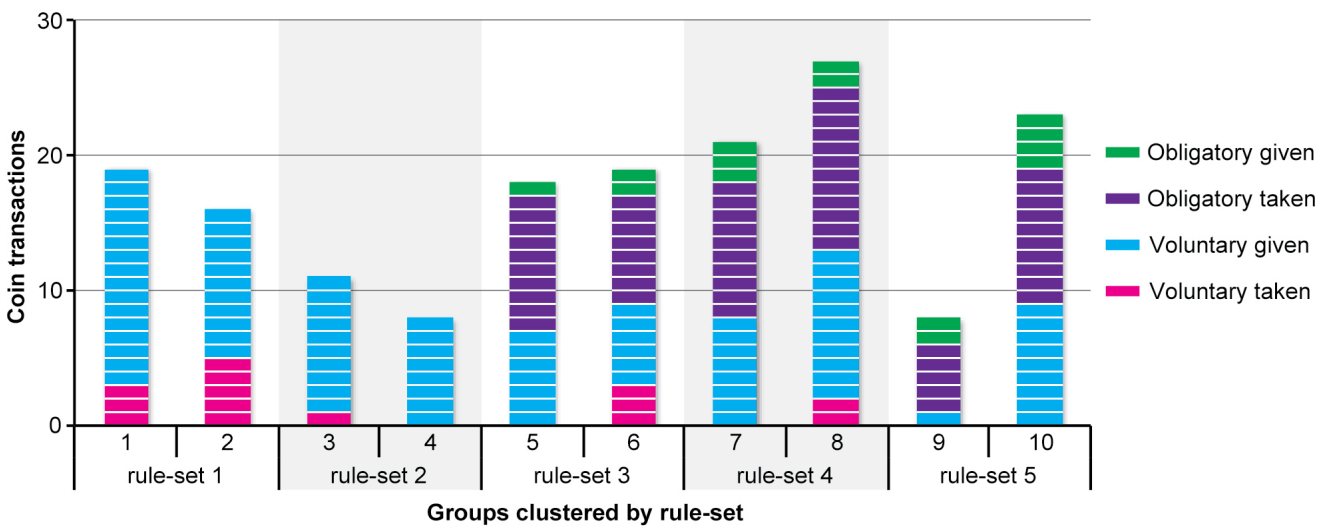


Figure 10. Observed coin use per group.

and direction. Overall, groups with the same rule-set exchanged approximately the same number of coins and thus seem to have followed the given rules. Only the groups with rule-set 5 differ strongly. This could be explained by the fact that group 9 interpreted the obligatory taking moment as the only moment to exchange coins. After a taking round, one participant commented: “*this [coin exchange moment] feels like standing in front of a council.*”

Table 4 in Appendix II shows the data of the questionnaire. One questionnaire from group 10 was not filled in properly (all questions were rated equally low) and thus discarded from analysis. To get an initial indication for possible effects we used a MANOVA with all questionnaire elements as dependent variable and the groups as independent variable. The results revealed that groups only varied significantly in perceived coin use $F(9, 39) = 16.94, p = .036, \eta_p^2 = .43$ and attention to the game $F(9, 39) = 13.94, p = .004, \eta_p^2 = .53$. Post hoc comparisons revealed significant differences in attention to the game between groups 3 and 6 ($p = .005$), 3 & 7 ($p = .021$), and 6 and 9 ($p = .021$), which indicate an effect of the rule-sets. The participants of group 3 ($M = 1.8$) and 9 ($M = 2.0$) had less attention for the game than participants from group 6 ($M = 3.8$) and 7 ($M = 3.5$). For the other items, we found no clear patterns that indicate an effect of the rule-sets. Instead, they reflect common practice in group-brainstorms. Participants were generally positive about one’s own ($M = 3.5, SD = .82$) and others’ contribution ($M = 3.7, SD = .75$). They complemented each other more ($M = 3.5, SD = .68$) than that they criticized each other ($M = 2.6, SD = 1.03$) and the participants reported more noticing ($M = 3.7, SD = .83$) than ignoring ($M = 2.4, SD = 1.10$).

Analysis Step 1:
Effect of Rule-Sets on Brainstorm Output

To examine the effect of the quantity and quality of rules on brainstorm output, we compared output assessments per rule-set (see Figure 11). We used a separate one-way ANOVA for each output category with productivity, originality, and feasibility scores as dependent variables and the rule-sets as independent variable. Levene’s tests revealed unequal variance in productivity ($p < .001$) and feasibility ($p = .003$), as such Welch’s tests were used for productivity and feasibility. Trends in brainstorm output

across rule-sets could indicate an effect of quantity of rules and variations between individual rule-sets could indicate effects of rule qualities. We checked for an influence of group composition by analyzing the effect of gender distribution and internal relationships on brainstorm output. The results did not reveal any significant effects of group composition.

The analysis of brainstorm output per rule-set only revealed a significant main effect of rule-sets on productivity Welch $F(4, 6.34) = 9.88, p = .007, est. \omega^2 = .78$. There was a significant linear trend $F(1, 15) = 15.66, p = .001$, indicating that an increasing quantity of rules reduces brainstorm productivity. The productivity graph in Figure 11a mainly shows drops between rule-sets 1 and 2 and between rule-sets 4 and 5. This suggests a negative effect of additional governing rules and of blinding the pools. Yet post hoc comparisons, using the Games-Howell post hoc procedure, indicated no significant differences between these individual rule-sets. Thus effects of rule qualities could not be deduced. Consequently, these results only indicate a negative effect of game rule quantity on brainstorm productivity. Brainstorm output quality (i.e., originality and feasibility) was not directly affected.

Analysis Step 2: Gamification Invasiveness
Effect of Rule-Sets on Gamification Invasiveness

As an increasing number of game rules reduced productivity, we expected that participants experienced increased invasiveness of the gamification across rule-sets. To define and analyze gamification invasiveness; we combined attention to the game, perceived coin use, and observed coin use ($\alpha = .91$; see Figure 12). Observed coin use was transformed to a 1-5 scale (as shown in the y-axis on the right) to resemble the other two items. Moreover, the average rating of attention to the game and perceived coin use was taken because they both seem to have measured perceived invasiveness of using the coins. The variation in gamification invasiveness between rule-sets (see outlined bars in Figure 12) was examined using a one-way ANOVA with the rule-sets as independent variable and invasiveness as dependent variable. The results revealed a significant variation in invasiveness as a result of the rule-sets $F(4, 19) = 5.01, p = .009, \eta^2 = .57$. Post-hoc comparisons, using Tukey’s test, revealed significant variation

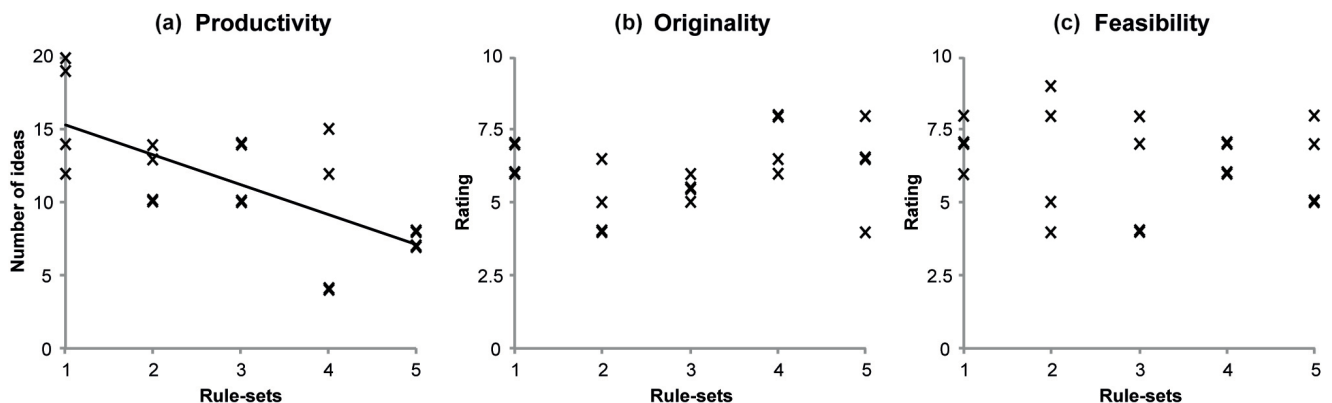


Figure 11. Brainstorm output assessment per rule-set (line: significant trend).

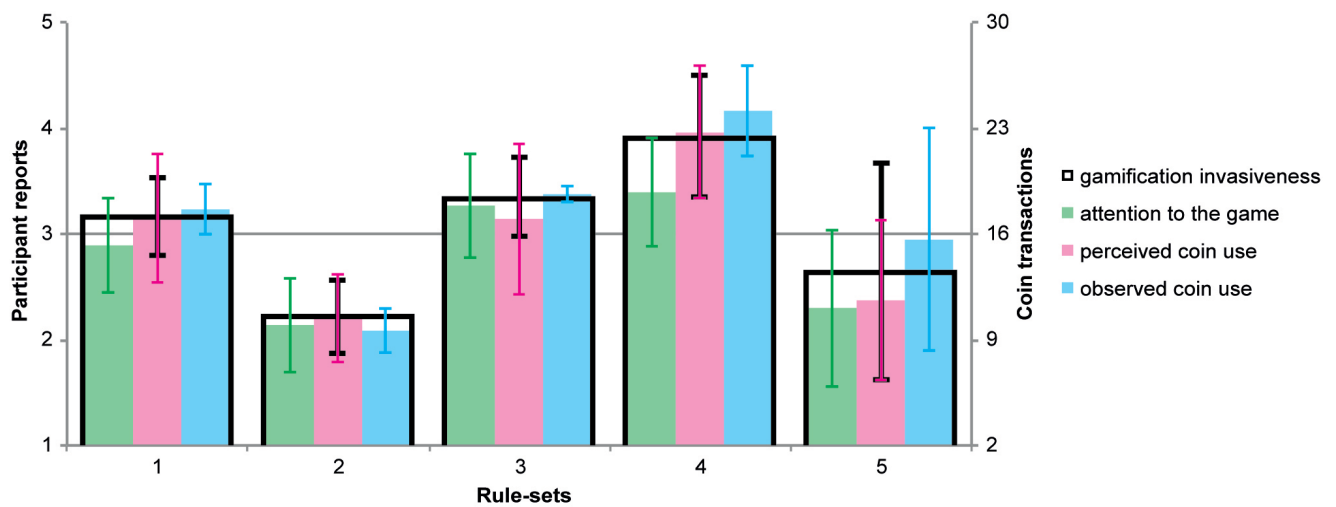


Figure 12. Gamification invasiveness per rule-set (error bars black: *SD*, green/red: 95% *CI*, blue: *SE*).

in invasiveness of the gamification between rule-sets 2 and 4 ($p = .004$), indicating that the forcing rules led to a statistically significant increase of gamification invasiveness.

The separate invasiveness items were analyzed to investigate the users' game behavior and experience in more detail. We analyzed the effect of rule-sets on perceived coin use and attention to the game using a MANOVA with the rule-sets as independent variable and all questionnaire elements as dependent variables. The results revealed a significant multivariate effect $F(80, 72) = 1.69, p = .012, \eta_p^2 = .65$ with significant variation between rule-sets in perceived coin use $F(4, 34) = 5.39, p = .002, \eta_p^2 = .39$ and attention to the game $F(4, 34) = 5.04, p = .003, \eta_p^2 = .37$. Post-hoc comparisons revealed significant variation in perceived coin use between rule-sets 2 and 4 ($p = .002$) and 4 and 5 ($p = .008$), and in attention to the game between rule-sets 2 and 3 ($p = .022$), 2 and 4 ($p = .009$), and 4 and 5 ($p = .036$). Hence, next to the positive effect of forcing rules, blinding the pools reduced perceived invasiveness.

The effect of rule-sets on observed coin use was analyzed using a one-way ANOVA with the number of transactions as dependent variable. The analysis revealed no significant variation, probably due to the low number of brainstorm sessions per rule-set. We did find significant variation when separating coin exchange in directions (i.e., giving and taking). The results revealed a significant variation between rule-sets in the number of taken coins *Welch* $F(4, 2.37) = 28.95, p = .021, \text{est. } \omega^2 = .92$. Post hoc comparisons, using the Games-Howell post hoc procedure, indicated that rule-set 3 led to significantly more coin taking than rule-set 2 ($p = .015$), which is explained by the forcing coin taking rule in rule-set 3 (rule 5 in Table 1).

Effect of Gamification Invasiveness on Brainstorm Output

As the invasiveness of the gamification did not vary strongly, analyses of the effect of invasiveness on output did not reveal significant results. To get an indication of possible transfer effects,

we exploratively examined the relationship between invasiveness and brainstorm output by calculating correlations. The results, shown in Table 2, indicate that overall invasiveness did not relate to brainstorm output. Yet the separate invasiveness items do reveal significant correlations. We found a strong and significant positive relationship between perceived coin use and originality. Groups that perceived more coin use generally scored higher on the originality of their ideas (see Figure 13a). Moreover, the observed number of voluntary taken and obligatory given coins correlated significantly with productivity. Voluntary taking related positively to productivity, suggesting that competitive game behavior may have stimulated idea generation. Obligatory giving correlated negatively with productivity, which indicates reduced productivity due to obligatory transactions.

Analysis Step 3: Brainstorm Engagement and Behavior

To examine the mediating effect of brainstorm engagement and behavior, correlations were calculated as well (see Table 3). The correlations with attention to the game and perceived coin use were calculated using the questionnaire data per participant ($N = 39$). The other correlations used averages per group ($N = 10$). The correlations with perceived brainstorm behavior revealed significant relationships with contribution, criticizing, and collaboration. Self-reported contribution to the brainstorm related positively to feasibility, indicating that the participants assessed their contribution mainly on feasible output rather than on original output. Criticizing correlated positively with originality (Figure 13a). Thus in sessions where participants perceived more critique, the output was more original. This supports the expected positive effect of competitive conflict behavior on brainstorm output (Badke-Schaub et al., 2010). Interestingly, criticizing also correlated significantly with perceived coin use (Figure 13b), indicating that increased coin exchange evoked competitive conflict behavior. Moreover, perceived collaboration correlated negatively with voluntary taking, suggesting that coin exchange reduced collaborative conflict behavior as well.

Table 2. Correlations between gamification invasiveness and brainstorm output (N = 10).

	Productivity		Originality		Feasibility	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Gamification invasiveness	-.018	.961	.378	.282	-.133	.714
Attention to the game	.176	.626	.416	.231	-.466	.174
Perceived coin use	.169	.640	.692* Fig.13a	.027	-.262	.464
Observed coin use	-.046	.900	.201	.577	.213	.554
Voluntary given	.394	.260	.036	.921	.409	.240
Voluntary taken	.829**	.003	.292	.413	-.139	.702
Obligatory given	-.646*	.044	-.066	.856	.016	.965
Obligatory taken	-.468	.172	.140	.700	-.004	.992

** *p* < .01; * *p* < .05.

Table 3. Correlations (*p* < .20) of perceived brainstorm behavior & engagement with brainstorm output and gamification invasiveness.

	Brainstorm behavior					Brainstorm engagement			
	Contribution	Complement	Criticize	Ignore	Notice	Attention to the brainstorm	Own engagement with the process	Own engagement with the output	Satisfaction with brainstorm output
Productivity					.478				
Originality	-.464		.633* Fig.13c						
Feasibility	.797**				-.621				
Gamification invasiveness						-.465			-.660*
Attention to the game									
Perceived coin use			.478** Fig.13b				.241	-.212	
Observed coin use		-.451							-.826**
Voluntary given	.587					.650*			
Voluntary taken				.603	-.726*				
Obligatory given		-.498				-.510	-.558		-.666*
Obligatory taken								-.512	-.619

** *p* < .01; * *p* < .05; italic = individual-level correlations (N = 39); non-italic = group-level correlations (N = 10).

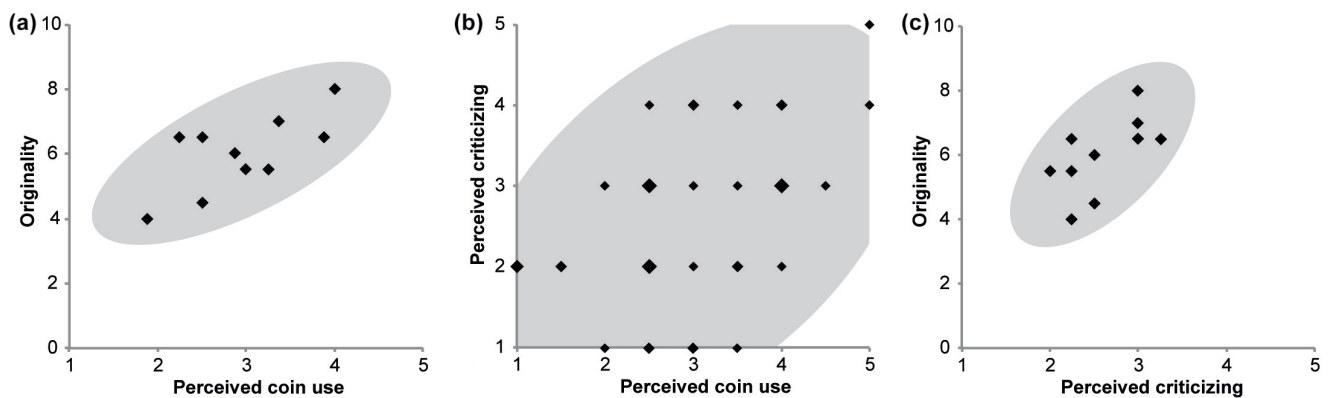


Figure 13. Significant correlations signifying a positive transfer effect on brainstorm output quality. a) Perceived coin use and originality (N = 10), b) perceived coin use and perceived criticizing (N = 39), c) perceived criticizing and originality (N = 10).

The correlations with brainstorm engagement only revealed significant relationships with gamification invasiveness. Attention to the brainstorm correlated positively with voluntary giving, suggesting that groups that were more occupied with the brainstorm gave each other more coins without being forced by the rules. Conversely, obligatory transactions only revealed negative correlations, suggesting that obligatory game behavior distracted participants from brainstorming. Moreover, significant negative correlations with output satisfaction suggest that the gamification made participants feel less satisfied with the output.

Summary of the Results

In summary, we found a significant linear negative trend in productivity across rule-sets, indicating a negative effect of rule quantity on brainstorm productivity. The analysis of gamification invasiveness revealed that the forcing rules, in contrast to governing rules, significantly increased the invasiveness of the gamification. Subsequently, significant correlations between separate invasiveness items and brainstorm output suggest that the invasiveness of the gamification did affect brainstorm output. Perceived coin use correlated positively with originality and voluntary coin taking correlated positively with productivity. Conversely, obligatory coin giving correlated negatively with productivity.

The analysis of brainstorm behavior revealed significant positive correlations between perceived criticizing and perceived coin use as well as originality. Moreover, contribution related positively to the feasibility of brainstorm output. Correlations between brainstorm engagement and gamification invasiveness mainly revealed negative relationships between obligatory coin exchange and brainstorm engagement as well as negative relationships between overall gamification invasiveness and satisfaction with brainstorm output.

Reconsidering the Research Framework

To understand the effects of game rules better, we used the experimental data to reconsider and revise the initial research framework (Figure 3). As explained below, the revised framework (Figure 14) describes positive and negative transfer effects based on the above-described experimental results.

Positive Effects of the Gamification

The forcing rules seem to have positively influenced the originality of brainstorm output by evoking competitive behavior in the brainstorm (see nr.1 Figure 14). Forcing rules significantly increased perceived coin use and perceived coin use was on its turn positively related to originality (Figure 13a). The analysis of questionnaire data revealed that this positive effect was mediated by perceived criticizing because it correlated positively with perceived coin use and originality (Figure 13b-c).

Additionally, coin taking related to reduced collaboration and improved productivity (nr.2 Figure 14). This was, however, not caused by the forcing rules, because only the number of voluntarily taken coins correlated significantly with collaboration (see Table 3) and productivity (see Table 2). The data does not show a significant relationship between collaboration and productivity, yet literature suggests that intense collaboration can reduce productivity because participants need to wait for their turn to talk (Paulus et al., 2002). Hence, we assume that coin taking led to increased productivity by reducing collaborative behavior.

Negative Effects of the Gamification

The negative trend in productivity across rule-sets (Figure 11a) could only partly be explained by the research results. The linear reduction indicates a negative effect as a result of the quantity of rules. The correlations between observed coin use and brainstorm

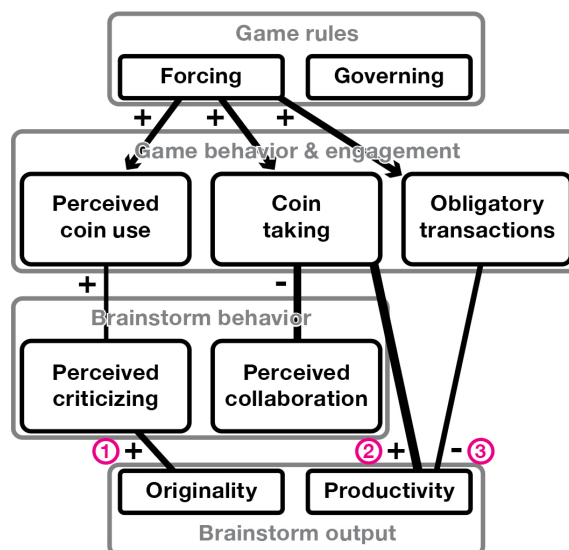


Figure 14. Revised framework based on experimental results (line thickness indicates correlation strength).

- 1) Significant positive transfer effect of forcing rules on originality, 2) significant positive transfer effect of voluntary competitive game behavior on productivity, 3) significant negative transfer effect of forcing rules on productivity.

output (Table 2) suggest that obligatory transactions and thus the forcing rules may have hindered idea generation (nr.3 in Figure 14). Yet the data did not fully support this, because only voluntary taking and obligatory giving correlated significantly with productivity (see Table 2).

When inspecting the reduction of productivity across rule-sets in Figure 11a more closely, we conclude that obligatory transactions were not the only cause for production blocking. The data mainly shows drops in rule-sets 2 and 5, where obligatory transactions had no or reduced influence (see Figure 10). Instead, in both rule-sets, gamification invasiveness dropped along with productivity. This suggests that the rules that were introduced in rule-sets 2 and 5 not only hindered participants in producing ideas but also in using the coins.

Discussion

This article presented an experiment that investigated positive and negative effects of adding game rules to group brainstorming by examining the participants' behavior and engagement during gamified brainstorming. In these gamified brainstorming participants could reward or punish each other by giving or taking golden coins. Game rules were designed to stimulate competitive conflict behavior and attention to each other by forcing and governing game behavior. According to literature, this would be beneficial for the quality and quantity of brainstorming output. Yet we also expected that increased engagement with the gamification would probably weaken brainstorming performance. In a between-group experiment, brainstorming groups received different rule-sets to investigate to what extent the invasiveness of the gamification affected their originality, feasibility, and productivity. Moreover, the behavior and engagement of participants were measured to gain a better understanding of how the rules influenced brainstorming output.

The results of the experiment demonstrate that voluntary competitive game behavior (i.e., coin taking) related to reduced collaboration in the brainstorming and increased productivity. This suggests that rules that simply allow for game-like competition could already be beneficial for brainstorming output quantity. Additionally, the quality of brainstorming output improved as a result of rules that forced competitive game behavior. The forcing rules significantly increased invasiveness of the gamification and were related to competitive brainstorming behavior and a higher originality of ideas.

The forcing rules also seem to have had a negative effect on brainstorming productivity. In general, the quantity of rules had a negative effect on the number of produced ideas and further analysis suggested that this was caused by mandatory game behavior (i.e., behavior that was forced by the rules). However, we observed the major productivity drops in brainstorming sessions where the forcing rules had no or reduced influence. Instead, the invasiveness of the gamification dropped along with productivity. These drops occurred in sessions where the rule-sets contained governing rules (rule-set 2) and a rule that blinded the participants' score (rule-set 5). The latter rule was indeed intended to reduce invasiveness, yet with the assumption that it would increase engagement with the brainstorming and improve productivity.

Oppositely, the governing rules were indeed expected to reduce productivity, yet as a result of an increased invasiveness of the gamification. Instead, in both cases the rules led to discussions that were irrelevant to the brainstorming, thereby hindering idea generation as well as coin exchange.

Regarding the positive effects of the gamification, the results support our initial assumptions. The found negative effects, however, contradict our expectations. Based on Visch et al. (2013), we initially assumed that increased invasiveness of the gamification would reduce engagement with the brainstorming (see Figure 1). Yet the experiment suggests that a user's game world and real world experience are not directly linked. Instead, gamification may lead to a game world experience that exists next to or independent from a user's real world experience and the strongest positive as well as negative effects were found where both types of experience were high.

To achieve a strong game world experience as well as real world experience, we need to gain a better understanding of gamification invasiveness. This was an important secondary aim of the experiment because it is generally overlooked in gamification literature (Hamari et al., 2014). Most gamification studies raise the issue of user consent, i.e., the extent to which users accept a gamification layer. The results of our experiment suggest that the invasiveness of a gamification layer could be the underlying mechanism for the consent of users. Users seem to abandon the gamification to reduce its invasiveness if the game rules do not fit their real world goal, such as high quantity and quality of ideas during brainstorming sessions. However, this not only reduces gameplay but also hinders real world performance.

To avoid this dual negative effect, the game rules need to fit the real world context and be as little invasive as possible. In the present experiment we mainly found evidence for behavioral invasiveness, as described by the revised framework (Figure 14). Yet the data suggests that other types than behavioral invasiveness are experienced as well. As described by the model in Figure 15, the distracting effect of governing rules and blinding scores indicates cognitive invasiveness (see Figure 15). Additional governing rules demand understanding and decision making from the participants and blinding the scores demands memorization. Moreover, the positive effect of forcing rules on brainstorming quality may be explained by affective invasiveness (e.g., a playful attitude). In our experiment, all positive effects were related to competition, which in the collaborative context of a group brainstorming made participants probably more playful, leading to increased creativity (Amabile & Pillemer, 2012).

As to affective invasiveness, framing (Bateson, 1972; Goffman, 1974) a brainstorming as a game may already have an effect on creativity. The game-like nature of giving each other feedback through coins may have reduced evaluation apprehension just like a temporary identity reduced it (Jung et al., 2010). Moreover, coin exchange may have framed personal criticism, thereby reducing the negative impact of affective conflict (Jung & Lee, 2015). In the current setup of our experiment, the effect of framing could not be examined, yet the positive correlations between gamification invasiveness and output originality (see Table 2) do hint towards a framing effect.

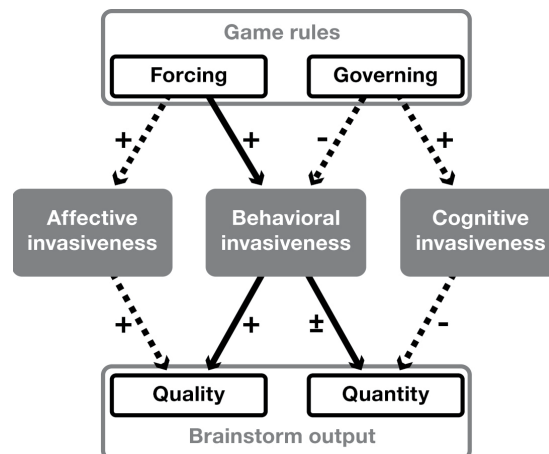


Figure 15. The invasiveness of game rules and its effect on brainstorm output
(straight lines: supported by data, dashed lines: not supported by data).

Limitations

Due to the low number of groups per rule-set, direct effects of particular rule-sets on brainstorm output remain speculative. Regarding their effect on gamification invasiveness we can be more certain, as the effects were coherent within rule-sets and significantly different between rule-sets. Accordingly, the correlations with gamification invasiveness items (i.e., attention to the game, perceived coin use, observed coin use) can be regarded to reflect true relations. The effect of the rules through specific coin exchange categories should be interpreted more carefully, because the variation between groups and rule-sets was small and generally not significant.

Moreover, the causal effect of coin exchange categories remains questionable. For example, a group might have been more competitive by nature and therefore have taken more coins from each other. Consequently, coin exchange may have reflected a group's brainstorm process, rather than influence it. Thus, again, effects from coin exchange categories should be interpreted carefully and subsequently the assumed influence that the game rules may have had.

The fact that the participants in the present experiment were design students may have weakened the positive effect of the gamification on brainstorm output due to a ceiling effect. The participants probably already adopted a playful attitude because they knew they were going to have a brainstorm. Thus the positive affective invasive effect may have been negligible. A group of participants without a design background, such as managers in a company, might benefit more from the gamification because the affective effect will be stronger.

Moreover, the goal, task, and timeframe of the brainstorm may have influenced the effect of the gamification. In the present experiment, time was short and the brainstorm's goal was relatively ambiguous, i.e., focusing on quantity as well as on quality. The short timeframe might explain the fact that we only found significant effects of the straightforward forcing rules, as opposed to the governing rules that require more learning time. The ambiguous brainstorm goal may have reduced the positive transfer effect of

giving each other feedback through coins, because the attribution of the feedback was often not clear. If the brainstorm would have been aimed at only generating feasible ideas, for example, coin exchange may have improved the feasibility of brainstorm output.

Conclusion

In conclusion, we could deduce two negative and two positive transfer effects from the experimental data: 1) competition in gamification increased the quality of brainstorm output, 2) voluntary competition led to increased brainstorm output quantity, 3) obligatory behavior in the gamification reduced the quantity of brainstorm output, and 4) discussion about the rules reduced quantity as well. Overall, in the present coin game, competitive game behavior improved the quality of brainstorm output at the cost of productivity. Hence, to arrive at an optimum for brainstorm output, a gamification layer should stimulate competition, yet avoid discussable governing rules and limit mandatory game behavior.

To design an effective gamification, the quality of game rules seems more important than the quantity. Forcing rules were found to be beneficial because they were easy to learn. This explains their positive effect in the present experiment, because groups played the gamification for the first time. The governing rules apparently required too much learning for this particular experiment.

Lessons Learned for Gamification Design and Research

To reduce the learning curve, the rules should be defined clearly. One way to achieve this is by embedding all rules into objects. For example, the coins could be visually marked as given or taken in order to ease enforcement of the rules related to the stock and profit pool (rule 4a-c in Table 1). Another solution could be to transform these rules into forcing rules (i.e., procedures) by digitalizing the game and automatically place given and taken coins in their respective pool.

The distinction between behavioral, cognitive, and affective invasiveness not only provides a model for anticipating on positive and negative transfer effects of a gamification, they can also be used to define the level on which you want to impact a non-game situation. As explained before, simply putting game-like artifacts on the table can lead to positive affective invasiveness. Adding a game system, such as rules governing coin exchange, makes a gamification cognitively invasive. In brainstorming this is detrimental, yet in other situations cognitive invasiveness might have a positive effect, such as making conscious decisions about your health. Steering users towards, for example, competition increases behavioral invasiveness in a generally collaborative setting such as a brainstorm. Yet a gamification may be much less behaviorally invasive if game and non-game behavior are aligned.

In gamification research, we would recommend to measure the different types of invasiveness separately. Ideally, behavioral invasiveness should not only be measured through behavioral data. Cognitive and affective invasiveness are more difficult to separate, because both generally rely on self-reporting. Yet cognitive invasiveness may also be measured through transcriptions of communication or by counting speaking time about the rules of a gamification.

Future Research

Based on our research framework, we see opportunities for gamification research on three levels: game rules, user experience a behavior, and real world output. Regarding game rules, our findings suggest that rule qualities should receive more attention in gamification research. In this article we distinguished rules in their quality of obligation, yet in other contexts other qualities may be relevant, such as being open-ended or goal-driven.

Regarding brainstorm output, our research demonstrated an improvement in the originality of ideas. Yet within the process of product development originality is mainly beneficial at the beginning. An interesting follow-up question would be: do gamified brainstorming eventually lead to better products? In some cases, the feasibility of ideas may be more important, which was not influenced or even reduced by the gamification in our experiment. To gamify a full design process one should probably design separate gamifications for each part that requires a different attitude from the user.

The main take-away of this article is probably that understanding the dynamics within a gamified situation provides a much broader and deeper understanding of the transfer effects of a gamification. Behavioral, cognitive, and affective invasiveness allowed us to capture the behaviors and experiences of users that are relevant for understanding and optimizing a gamification to achieve its intended real world goal.

References

- Amabile, T. M. (2012). Componential theory of creativity. In E. H. Kessler (Ed), *Encyclopedia of management theory* (pp. 134-139). London, UK: Sage.
- Amabile, T. M., & Pillemer, J. (2012). Perspectives on the social psychology of creativity. *The Journal of Creative Behavior*, 46(1), 3-15.
- Avedon, E. M. (1971). The structural elements of games. In E. M. Avedon & B. Sutton-Smith (Eds.), *The study of games* (pp. 419-426). New York, NY: Wiley.
- Badke-Schaub, P., Goldschmidt, G., & Meijer, M. (2010). How does cognitive conflict in design teams support the development of creative ideas? *Creativity and Innovation Management*, 19(2), 119-33.
- Bateson, G. (1972). *Steps to an ecology of mind* (1987 ed.). London, UK: Jason Aronson.
- Bergström, K. (2010, October). The implicit rules of board games: On the particulars of the lusory agreement. In *Proceedings of the 14th International Academic MindTrek Conference on Envisioning Future Media Environments* (pp. 86-93). New York, NY: ACM.
- Björk, S., & Holopainen, J. (2003). Describing games: An interaction-centric structural framework. In *Proceedings of the DiGRA International Conference: Level Up*. Retrieved October 18, 2018, from <http://www.digra.org/digital-library/publications/describing-games-an-interaction-centric-structural-framework/>
- Bottger, P. C., & Yetton, P. W. (1987). Improving group performance by training in individual problem solving. *Journal of Applied Psychology*, 72(4), 651-657.
- Brown, V., & Paulus, P. B. (1996). A simple dynamic model of social factors in group brainstorming. *Small Group Research*, 27(1), 91-114.
- Brown, V., Tumeo, M., Larey, T. S., & Paulus, P. B. (1998). Modeling cognitive interactions during group brainstorming. *Small Group Research*, 29(4), 495-526.
- Caillois, R. (1961). *Man, play, and games* (2001 ed.). Chicago, IL: University of Illinois.
- Christiaans, H. H. C. M. (2002). Creativity as a design criterion. *Creativity Research Journal*, 14(1), 41-54.
- 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 on Envisioning Future Media Environments* (pp. 9-15). New York, NY: ACM.
- Deutsch, M. (2006). Cooperation and competition. In M. Deutsch, P. T. Coleman, & E. C. Marcus (Eds.), *The handbook of conflict resolution: Theory and practice* (pp. 23-42). San Francisco, CA: Jossey-Bass.
- DiMicco, J. M., Pandolfo, A., & Bender, W. (2004). Influencing group participation with a shared display. In *Proceedings of the Conference on Computer Supported Cooperative Work* (pp. 614-623). New York, NY: ACM.
- Dugosh, K. L., & Paulus, P. B. (2005). Cognitive and social comparison processes in brainstorming. *Journal of Experimental Social Psychology*, 41(3), 313-320.
- Fluegelman, A. (1976). *The new games book*. New York, NY: Doubleday.

18. Galton, M. (2010). Assessing group work. In P. Peterson, E. Baker, & B. McGraw (Eds.), *International encyclopedia of education* (pp. 342-347). Oxford, UK: Elsevier.
19. Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. New York, NY: Harper & Row.
20. Gray, D., Brown, S., & Macanuso, J. (2010). *Gamestorming: A playbook for innovators, rulebreakers, and changemakers*. Sebastopol, CA: O'Reilly.
21. Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, 14(4), 311-327.
22. Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025-3034). Piscataway, NJ: IEEE.
23. Harkins, S. G., & Szymanski, K. (1989). Social loafing and group evaluation. *Journal of Personality and Social Psychology*, 56(6), 934-941.
24. Huang, J. C. (2010). Unbundling task conflict and relationship conflict: The moderating role of team goal orientation and conflict management. *International Journal of Conflict Management*, 21(3), 334-355.
25. Huotari, K., & Hamari, J. (2012). Defining gamification: A service marketing perspective. In *Proceedings of the 16th International Academic MindTrek Conference* (pp. 17-22). New York, NY: ACM.
26. Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, 40(2), 256-282.
27. Järvinen, A. (2008). *Games without frontiers: Theories and methods for game studies and design*. Tampere, Finland: Tampere University.
28. Jung, E. J., & Lee, S. (2015). The combined effects of relationship conflict and the relational self on creativity. *Organizational Behavior and Human Decision Processes*, 130, 44-57.
29. 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.
30. Juul, J. (2003). The game, the player, the world: Looking for a heart of gameness. In *Proceedings of the DiGRA International Conference: Level Up*. Retrieved October 18, 2018, from <http://www.digra.org/wp-content/uploads/digital-library/05163.50560.pdf>
31. Knaving, K., & Björk, S. (2013, October). Designing for fun and play: Exploring possibilities in design for gamification. In *Proceedings of the 1st International Conference on Gameful Design, Research, and Applications* (pp. 131-134). New York, NY: ACM.
32. Marks, M. A., & Panzer, F. J. (2004). The influence of team monitoring on team processes and performance. *Human Performance*, 17(1), 25-41.
33. McGonigal, J. (2011). *Reality is broken*. London, UK: Jonathan Cape.
34. Mekler, E. D., Brühlmann, F., Opwis, K., & Tuch, A. N. (2013, April). Disassembling gamification: The effects of points and meaning on user motivation and performance. In *Proceedings of the Extended Abstracts on Human Factors in Computing Systems* (pp. 1137-1142). New York, NY: ACM.
35. Moccozet, L., Tardy, C., Opprecht, W., & Léonard, M. (2013, September). Gamification-based assessment of group work. In *Proceedings of the International Conference on Interactive Collaborative Learning* (pp. 171-179). Piscataway, NJ: IEEE.
36. Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. M. (2006). The illusion of group productivity: A reduction of failures explanation. *European Journal of Social Psychology*, 36(1), 31-48.
37. Osborn, A. F. (1963). *Applied imagination: Principles and procedures of creative problem solving*. New York, NY: Charles Scribner's Sons.
38. Paulus, P. B., Dzindolet, M. T., Poletes, G., & Mabel Camacho, L. (1993). Perception of performance in group brainstorming: The illusion of group productivity. *Personality and Social Psychology Bulletin*, 19(1), 78-89.
39. Paulus, P. B., Putman, V. L., Dugosh, K. L., Dzindolet, M. T., & Coskun, H. (2002). Social and cognitive influences in group brainstorming: Predicting production gains and losses. *European Review of Social Psychology*, 12(1), 299-325.
40. Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, 30(4), 344-60.
41. Salen, K., & Zimmerman, E. (2004). *Rules of play: Game design fundamentals*. Cambridge, MA: MIT.
42. Sniderman, S. (2006). Unwritten rules. In K. Salen & E. Zimmerman (Eds.), *The game design reader: A rules of play anthology* (pp. 476-502). Cambridge, MA: MIT.
43. Suits, B. (1978). *The grasshopper: Games, life, and Utopia*. Toronto, Canada: University of Toronto.
44. Sutton, R. I., & Hargadon, A. (1996). Brainstorming groups in context: Effectiveness in a product design firm. *Administrative Science Quarterly*, 41(4), 685-718.
45. Tausczik, Y. R., & Pennebaker, J. W. (2013, April). Improving teamwork using real-time language feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 459-468). New York, NY: ACM.
46. Visch, V., Vegt, N. J. H., Anderiesen, H., & van der Kooij, K. (2013). *Persuasive game design: A model and its definitions*. Paper presented in CHI 2013 Workshop on Designing Gamification: Creating Gameful and Playful Experiences. Retrieved from <http://gamification-research.org/chi2013/papers/>.
47. Wang, C. S., Galinsky, A. D., & Murnighan, J. K. (2009). Bad drives psychological reactions, but good propels behavior: Responses to honesty and deception. *Psychological Science*, 20(5), 634-644.

48. Weingart, L., & Jehn, K. A. (2003). Manage intra-team conflict through collaboration. In E. A. Locke (Ed.), *Handbook of principles of organizational behaviour* (pp. 327-346). Hoboken, NJ: John Wiley & Sons.
49. Wu, L. Z., Ferris, D. L., Kwan, H. K., Chiang, F., Snape, E., & Liang, L. H. (2015). Breaking (or making) the silence: How goal interdependence and social skill predict being ostracized. *Organizational Behavior and Human Decision Processes*, 131, 51-66.
50. Yong, K., Sauer, S. J., & Mannix, E. A. (2014). Conflict and creativity in interdisciplinary teams. *Small Group Research*, 45(3), 266-289.

Appendix I. Questionnaire

To **what extent** were the following aspects present during the design assignment?

	Not much	Little	Somewhat	Much	A great deal
1. Your contribution to the end result.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. The contribution to the end result of your fellow designers.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Criticize each other's contribution.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Complement each other's contribution.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Ignore each other's contribution.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Taking notice of each other's contribution.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Your involvement in the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. The involvement of your fellow designers in the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Your feeling of responsibility for the outcome of the design assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. The feeling of responsibility of your fellow designers for the outcome of the design assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. The use of coins during the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. The use of coins at the end of the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. The added value of the coin game to the progress of the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. The added value of the coin game to the outcome of the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. A proportional contribution of all participants to the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. Your attention to the design challenge.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
17. Your attention to the coin game.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Could you indicate the **quality** of the following aspects of the design assignment?

	Very low	Low	Average	High	Very high
18. The collaboration during the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
19. The end result of the assignment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
20. Your chance of winning the challenge.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
21. The solidarity within the group.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix II. Questionnaire Data

Table 4. Average self-reported behavior and engagement per group with ranking in brackets.

Group	Engagement				Perceived coin use*	Brainstorm behavior					
	Brainstorm process	Brainstorm output	Brainstorm attention	Game attention*		Own contribution	Other's contribution	Criticize	Complement	Ignore	Notice
1	[3] 4.1	[4] 3.5	[2] 3.8	[5] 2.8	[6] 2.9	[1] 4.0	[2] 4.0	[5] 2.5	[3] 3.5	[5] 2.5	[4] 3.8
2	[5] 3.9	[1] 4.0	[1] 4.0	[4] 3.0	[3] 3.4	[6] 3.3	[6] 3.5	[2] 3.0	[3] 3.5	[1] 3.3	[7] 3.5
3	[7] 3.8	[8] 2.8	[6] 3.5	[10] 1.8	[10] 1.9	[5] 3.5	[6] 3.5	[7] 2.3	[7] 3.3	[3] 2.8	[10] 2.8
4	[2] 4.3	[1] 4.0	[2] 3.8	[8] 2.5	[7] 2.5	[4] 3.8	[2] 4.0	[7] 2.3	[1] 4.3	[8] 2.0	[1] 4.3
5	[1] 4.4	[4] 3.5	[2] 3.8	[5] 2.8	[5] 3.0	[1] 4.0	[2] 4.0	[7] 2.3	[2] 3.8	[10] 1.0	[1] 4.3
6	[5] 3.9	[10] 2.5	[10] 2.8	[1] 3.8	[4] 3.3	[6] 3.3	[5] 3.8	[10] 2.0	[7] 3.3	[6] 2.3	[3] 4.0
7	[3] 4.1	[7] 3.0	[6] 3.5	[2] 3.5	[2] 3.9	[6] 3.3	[10] 3.3	[2] 3.0	[7] 3.3	[9] 1.8	[8] 3.3
8	[8] 3.6	[9] 2.6	[2] 3.8	[3] 3.3	[1] 4.0	[1] 4.0	[1] 4.3	[2] 3.0	[3] 3.5	[2] 3.0	[4] 3.8
9	[8] 3.6	[6] 3.1	[9] 3.0	[9] 2.0	[9] 2.3	[10] 3.0	[6] 3.5	[1] 3.3	[3] 3.5	[3] 2.8	[4] 3.8
10	[10] 3.3	[3] 3.7	[8] 3.3	[7] 2.7	[7] 2.5	[6] 3.3	[6] 3.5	[5] 2.5	[7] 3.3	[6] 2.3	[8] 3.3

* Significant variation between groups ($p < .05$); grey highlights: groups that vary significantly ($p < .05$).