Gamified Individual Brainstorming

Reaching for the stars

Jon Spruyt
Abstract

Ideation is a critical step in the creation of new solutions to problems. Brainstorming, being the most commonly known ideation method, is mostly done in groups. However, research points out that brainstorming individually is more effective than brainstorming in groups. This paper investigates if negative aspects (reduced enjoyment, motivation and self-rated performance) associated with brainstorming alone can be overcome by using a gamified approach to brainstorming. Gamified and non-gamified prototypes have been made and used in combination with a questionnaire to measure differences in these negative aspects. From the within-subject study including 20 participants, it appears that the implemented gamification approach was not different from the non-gamified approach for most of the negative aspects. Self-rated performance was however significantly higher in the non-gamified version. Using the gamified or non-gamified approach seems to largely come down to the preference of the user.

Keywords: Ideation, Gamification

1. Introduction

In innovation, ideation is often a first, but always a critical step in coming with new solutions to problems. Ideation is the process towards attaining a desired goal through generating ideas or concepts (Briggs & Reinig, 2007). It is a key step in developing a final solution or product to a problem or opportunity. Because it plays such a major role many ideation techniques have been developed, for example the Morphological Analysis (Zwicky, 1969), the Synectics method (Gordon, 1961), the 6-3-5 Method (Rohrbach, 1969) but most commonly known is the Brainstorm method originally proposed by Osborn (Osborn, 1953). Osborn proposed that groups thinking of ideas together, building on each other's ideas while being open for wild ideas and not criticizing them would lead to an increased quantity and quality of ideas.

In their paper “Does Group Participation When Using Brainstorming Facilitate or Inhibit Creative Thinking?” Taylor, Berry, & Block (1958) test this claim and find that nominal groups, consisting of the same amount of individuals as real groups, but who brainstorm individually, outperform real groups of equal size. In their experiment these nominal groups generated nearly twice as many ideas as the real group did. Many other researchers replicated this experiment and found that nominal groups were mostly outperforming real groups and sometimes no difference was measured between the groups. For example Lamm and Trommsdorff (1973) confirm this claim, but indicate that no such claims can be made about the quality of the ideas, since it is unclear to the research community how to appropriately measure quality ideas.

Diehl and Stroebbe (1987) uncovered that this loss in productivity is caused by three main factors: production blocking, evaluation apprehension and free riding. Production blocking occurs when only one person in a group speaks, other people might have an idea in their
head, but can’t form it or can’t speak up about it at that time. This idea might get lost or never formed. Evaluation apprehension occurs when people have formed their ideas, they might be too afraid to propose that to the group for fear of negative feedback from the others. Free riding is the last factor and it means people lifting on other people’s ideas and not bringing in original ideas. For these reasons, Diehl and Stroebe suggest to first let people brainstorm individually and then come together in groups for evaluation of the ideas.

Researchers have thought that brainstorming electronically might remove these inhibitions, however, these seemed to only make a difference in really large brainstorming groups (Aiken, Krosp, Shirani, & Martin, 1994). Paulus, Dzindolet, Poletes, & Camacho (1993) in their experiment look at why group brainstorming is still so prevalent if so much research indicates it is less effective than brainstorming alone. They found that people perceive their performance as better while brainstorming in groups compared to brainstorming alone. Diehl and Stroebe (1991) concur with this and state that people simply enjoy working in groups more than working individually and are more motivated to participate when there is social interaction.

In the field of gamification many researchers try to increase motivation, enjoyment or engagement of the users by adding game elements to their applications or services (Hamari, Koivisto, & Sarsa, 2014). In their literature review Hamari, Koivisto, & Sarsa look at the different gamification papers, their methods and their outcomes. They summarize that much research had positive results or partly positive results when using a wide array of motivational affordances such as points, achievements, challenges, et cetera. They also show that gamification is prevalent in many different fields including innovation and ideation.

One explorative study that was mentioned there looked at how gamification worked with ideation in online idea competitions (Witt, Scheiner, & Robra-Bissantz, 2011). Their study indicates game mechanics may play a solution to improving motivation to participate, enjoyment and task involvement. Other games for ideation have been made, for example based on the Synectics method (Hauge, Duin, & Thoben, 2008; Duin, Baalsrud Hauge, & Thoben, 2009). Or based on brainstorming where the main objective is disrupting the brainstorm session to generate disruptive ideas (Duin, Hauge, Thoben, & Bierwolf, 2009).

However, very few of these games are directly aimed towards ideation or brainstorming individually. Some of them are played alone, but are multiplayer through the internet or are aimed towards a facilitator. For this it is left to question how a gamified brainstorm game can overcome the issue of perceived performance when playing individually raised by Paulus, Dzindolet, Poletes & Camacho. Additionally how introducing game mechanics to brainstorming can improve the motivation and enjoyment of doing the activity alone. The question this paper will be concerned with is thus:

*Can a gamified tool be used to facilitate brainstorming alone, and overcome negative aspects typically associated with brainstorming individually?*

This research will investigate if there is a difference between enjoyment, motivation, self-rated performance and the number of ideas when brainstorming individually with or without gamified elements in the brainstorm session. For this two versions of a prototype will be made where people can brainstorm on their own, either with gamified elements or without.
2. Related Research

2.1 Brainstorming

After Osborn introduced the concept of brainstorming in 1953 much research has been done to the claims he has made and, in turn, to the claims against his. In his original idea a brainstorm session consisted of more than three people and ideas should be spoken out while adhering to four rules:

1. To not criticize each other’s ideas; wait with judgement and evaluation for later.
2. To "free-wheel" as much as possible; It's easier to reduce wild ideas to realistic ones than to think up realistic ideas.
3. Quantity is wanted; The more ideas the higher the chance of a winning idea.
4. Combination and improvement are sought; Not only should participants come up with their own ideas, they should seek to improve others' ideas.

The aim of these rules is to free the participants of any criticism or inhibitions they have to come up with the most creative ideas. Additionally participants should enforce these rules upon others if someone seems to break one of them.

Osborn provides some indicative research that this method works to get more ideas. Taylor, Berry, & Block (1958) first empirically tested this method. In their experiment, they compared real groups to nominal groups. These nominal groups consisted of the same amount of individuals as the real groups, however they brainstormed alone and their ideas were gathered as if they would have brainstormed in a group. Their experiment measured the number of ideas, the uniqueness of ideas and three measures respectively to the quality of ideas. Their experiment found that the nominal groups outperform the real groups by a near 2:1 ratio. Lamm & Trommsdorff (1973) in their review also found that research they analyzed indicated that the number of ideas in nominal groups are clearly higher than with real groups. They also indicate that there is less clear evidence in the difference for the quality, uniqueness and variety of ideas. They and Diehl and Stroebe (1991) account this mainly due to there being no coherent and similar measure for quality and uniqueness for ideas. Some relatively new research has suggested using one metric for quality, uniqueness and variety, but it has not gained much traction (Nelson, Wilson, Rosen, & Yen, 2009).

As mentioned before, Diehl and Stroebe (1987), found three major aspects of why group brainstorming is not as effective as individual brainstorming: production blocking, evaluation apprehension and free riding. They indicate that production blocking is the largest factor from the three and their follow-up research tries to uncover where this phenomenon comes from (Diehl and Stroebe, 1991). They found that it is not the amount of speaking time the real groups have in comparison to the nominal groups that accounts for this, but rather the waiting time for a participant to tell her own idea. In this waiting time, the participant might either forget the idea that she had or to avoid this, rehearse her idea. This prevents her from producing additional ideas in that time.

In their concluding section, Diehl and Stroebe indicate that the reason why brainstorming in groups is still so prevalent is due to the illusion of group effectivity. First, they found that people enjoy working with groups more than working alone. So they are more motivated to
do the brainstorm when they can work with others. And the second they call the baseline fallacy:

"The baseline fallacy is reflected in such sayings as "two heads are better than one." Because more ideas are produced by a four person group than by an individual, members of such groups are under the impression that they can achieve much more as group members than they can on their own." (Diehl and Stroebe 1991, p. 403).

These reasons, they argue, are probably prevalent since nominal groups are not part of the experience of everyday life.

As mentioned in the introduction Paulus, Dzindolet, Poletes, & Camacho (1993) also investigated why group brainstorming was more prevalent. They discovered that while individuals brainstormed individually, they thought they would be more performant in groups and while brainstorming in groups they think they performed better than while alone. They indicate that this may in part be because of social comparison in groups and that individuals tend to credit themselves with a disproportionate amount of the results.

2.2 Gamification

Gamification is defined as enhancing services with game/motivational affordances to support user engagement and enhancing positive patterns in service use. For example to increase social engagement, user or task involvement and increase the quality or productivity of tasks (Hamari, Koivisto, & Sarsa, 2014). Or in words of Huotari and Hamari: "Gamification refers to: a process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation." (Huotari & Hamari, 2012, p18).

Hamari, Koivisto & Sarsa conceptualize gamification in three steps:

1. The motivational affordances that are implemented
2. The psychological outcomes resulting from these affordances
3. The behavioural outcomes resulting from the psychological changes.

The motivational affordances that are used in the literature are of cognitive (progress, badges or achievements), emotional (narrative, avatars, theme), or social (leaderboards, challenges, feedback) nature. Motivational affordances is a concept that ties need satisfaction theories of motivation with perceived opportunities for action, commonly called affordances. Need satisfaction theories state that people will engage in activities if these satisfy motivational needs, like relatedness or competence. So an object with a motivational affordance will somehow give the user a form of satisfaction upon completing or interacting with it (Deterding, 2011). Psychological outcomes may refer to the enjoyment, attitude or motivation to complete a task, whereas behavioural outcomes may refer to quality or quantity when completing tasks or engagement in certain activities (Hamari, Koivisto, & Sarsa, 2014).

Additionally they found that gamification is applied in many different fields, and have mostly positive results indicating that gamification does not only work in a particular setting but can be applied to many different settings. For example Eickhoff, Harris, de Vries, & Srinivasan (2012) where they gamified a crowdsourcing service. They used cognitive and
social motivational affordances in the form of points and leaderboards. While playing the game a tree grows based on the decisions the player makes, every grown tree gives additional points and acts as a milestone. With this they looked at the participant’s intrinsic and extrinsic motivation as psychological outcomes. More specifically if they wanted to be on top of the leaderboard, or if they wanted to complete another tree. The behavioural outcomes they looked at were the speed of task completion and the quality of the task being completed. In their study they found that all the results were all positive and that gamification seemed to really increase both the psychological and behavioural outcomes.

2.3 Brainstorming Games

Considering their near-playful nature and the need for high motivation and enjoyment throughout the entire brainstorm session, it is not surprising that researchers have applied gamification to brainstorm practices. For example Kultima, Niemelä, Paavilainen, & Saarenpää (2008) made several games to generate names for new games to be used by game developers. These games were all meant for game developers to sit down in groups and think of names for the games they were creating. In their study they found that using games as the brainstorming method, ideas started flowing immediately and that it made the participants more focused on the goal of generating ideas. They also mentioned that when introducing points to the brainstorm, it was sometimes unclear if participants were playing for the points or for creativity. Lastly, they concluded that keeping track of the ideas, or documenting them, is harder and sometimes burdensome while playing a game, so they suggest aiming to include the documentation in the game itself.

Toubia (2006) generated a different ideation game, one where he was interested in how incentives can play a role in ideation effectiveness. He had three experimental groups, one which he would reward regardless of the ideas proposed, a second where he would reward people based on the unique ideas they produced on their own, and lastly a group he rewarded for how many people generated descendant ideas from their ideas. He found that this third group performed a lot better on many metrics, like breadth, depth, novelty, total contribution and things like quantity and uniqueness of ideas. The second group, in turn, outperformed the first group significantly as well. This gives an indication that the type of incentive matters to generate an increasing quantity and improved quality of ideas.

Similarly, Hauge, Duin and Thoben (2008) made a brainstorming game where groups had to generate ideas based on a perspective, for example a technology perspective or a user-driven perspective. In addition this game introduced ‘events’, where the facilitator could trigger a new perspective on the problem to renew motivation and ideas. Initial results show that users felt more motivated using this game than traditional brainstorming and that these different perspectives and events had a positive effect on that. In a follow-up paper, they propose a different game where disruption is taken one step further by suggesting that a company has ‘taken over’ the current company and forcing people to think in values of this new company (Duin, Hauge, Thoben, & Bierwolf, 2009). This game proposes that people work alone on a computer, with many individuals working at the same time. There is one facilitator that can monitor the ideas of every individual and add disruptive events to steer thinking into another direction. These disruptive events are shared among all individuals. At the end of this stage the evaluation of all ideas will be held by the entire group.
3. Methods

This study looked at the differences in enjoyment, motivation, self-rated performance and the number of ideas generated while brainstorming individually between a gamified tool and a non-gamified tool. Measurements were done with a questionnaire as a follow up to a brainstorming task. This brainstorming task was performed within a prototype that has been made specifically for this study. The study consisted of two parts, a gamified brainstorming task with questions afterwards and a normal brainstorming task with questions afterwards. The following sections will go in depth on what the prototype, the questionnaire and the process consist of.

3.1 Prototype

A prototype was developed for this study to control variables that should be in- or excluded in the two versions of the brainstorm session. These two versions are nearly identical except for the gamification elements present in one. They were made so that participants can use their own computer and perform the task. The basic version uses a few design and brainstorming principles. As a brainstorming method, it uses associative brainstorming, where ideas can be generated with the help of associations made by various stimuli (Malaga, 2000). In this case, word stimuli were used in the form of adjectives. These adjectives were randomly generated for each participant and are presented above “sticky-note” like cards that are presented in a tray on the top of the screen (see illustration 1). The design metaphor used in the prototype is that of a whiteboard with coloured sticky notes on them, much like actual analog group brainstorms would work. Once participants fill out an idea on a card, they can use drag and drop functionality to put the ideas on the ‘board’ and group them by placing them together (see illustration 2).

As motivational affordances for the gamified version, progress has been combined with points and clear goals, as per definition of Hamari, Koivisto, & Sarsa (2014). Progress could be observed with progress bars (see illustration 3) which also indicated what the goal was for the game. For each four ideas, the participant would get a point, or in this case a star (see illustration 4). The total amount of stars can be seen in the top right of the screen, both illustrated in illustration 1 and 2. The cards resized themselves based on the amount of text, however users had the freedom to resize the cards themselves as they’d like. Lastly, a context menu is available on each card where an option is presented to delete ideas.

3.1.2 Implementation

The prototype has been created over a span of 40 to 50 hours and was implemented in a full frontend manner in HTML, CSS and Javascript. It has been tested for modern and most used versions of Google Chrome, Safari and Mozilla Firefox (data taken from StatCounter¹). It has only been optimized for desktop/laptop usage, as mobile devices are not suited for creating an overview of ideas. Tablets have not been considered as their market share is too little (see also StatCounter).

The prototype does not use server communication, this is for two reasons. First of all, the participants should have the freedom to quit their questionnaire and brainstorming task and

Illustration 1: Empty board with three opened adjective cards.

Illustration 2: Completed tool from a pilot test, 4 ideas have been filled in for six out of seven adjectives.

Illustration 3: progress bars

Illustration 4: Scoring a star mid-animation
not send in their answers at any time. This way, the server does not save any data from their tasks if they would quit somewhere halfway. A disadvantage of this is that if the participants’ browser crashes, all their data will be lost, which happened with one participant that was therefore excluded from the study. The other reason is that the researcher now does not get to see the ideas of the participants, negating the risk that participants might feel evaluation apprehension and write down fewer ideas.

3.2 Questionnaire

The main form of evaluation was a questionnaire. Participants of the study were sent a questionnaire containing all the instructions for the study (see Appendix 1 for an example), but if anything remained unclear they had the opportunity to contact the researcher directly. The questionnaire consisted of two brainstorm sessions, one gamified and one normal, and various questions afterwards. The brainstorm topics had to be of a somewhat similar depth and challenge, but at the same time so different that participants could not use the same ideas for both sessions. The two questions were:

1) How can a technology company make smartphones more accessible to the elderly?
2) What can universities do to introduce student life and all its aspects to new students?

These are deemed equal in the sense of that the participant would need to identify needs of a target group and have to apply those needs to a product or service. Both the target group and the service were made distinct from each other, so that ideas could not be reused, but the process should be similar.

An introductory text to explain the study was given at the start of the questionnaire as well as an introductory text to each brainstorm session (also see Appendix 1). For each session the brainstorm rules according to Osborn (1953) were given, however slightly adjusted so that they do not refer to ideas of other people any longer, but only to ideas of the person brainstorming alone. Each introduction text also had an image of the prototype with an explanation of how it functions and an explanation of how the prototype timed the session. For each gamified session an explanation of the stars was also given and participants were told to try to score as many stars as possible.

A 2x2 factorial design has been made for this study, resulting in four questionnaires. The two variables were the order of the brainstorming topic and whether the first test included gamification or not. For example, the first questionnaire contained the first topic the first round with the gamified version. The second round included the second topic and the non-gamified version of the prototype. This design was chosen because it might happen that people become more comfortable with brainstorming the second round, or alternatively become mentally exhausted from the task and perform less well the second round. Additionally, it needs to prove that the topic has no significant effect on the dependent variables.

To segment the participants, and see how certain variables influenced the results, three introductory questions were asked. They were asked about their gender, their age and finally whether they had brainstormed before. In this last question participants could indicate if they brainstormed alone, in a group, both or neither. At the end of the survey there was an open question about if participants had any additional thoughts, comments or ideas.
concerning either the prototype or the study. All other questions in the questionnaire were asked twice, once after each brainstorm session.

The questions in the questionnaire were based on various sources. The study aims to measure a difference in enjoyment, motivation, perceived performance and the number of ideas. According to Deci & Ryan, Intrinsic motivation is measured purely by enjoyment whereas extrinsic motivation, on the other hand, refers to the performance the participant shows regarding a separate outcome (Deci & Ryan, 1985; Deci & Ryan, 2000). The measurement tool for both these types of motivation that has been used is the Intrinsic Motivation Inventory (Ryan, 1982). In this tool various subscales for motivation, each including multiple statements is measured on a scale from 1 (strongly disagree) to 7 (strongly agree). Based on this inventory, the first two sections of the questionnaire have been formed (see appendix 1). Enjoyment will be measured by the enjoyment subscale in the Intrinsic Motivation Inventory, extrinsic motivation will be measured with the value/usefulness scale in combination with goal achieving questions (Deci & Ryan, 2000).

The last category of questions were those for perceived performance. In their paper Paulus, Dzindolet, Poletes, & Camacho (1993) asked the question: “If you had to brainstorm in a group with three others, do you think you would personally have generated more ideas?”. The same question was asked regarding the quality of ideas. To adhere these questions to the questionnaire format they have been rephrased to be a statement instead (see Appendix 2). The other statements in this subscale are based on how people view their performance in relation to the task instructions.

Lastly, the number of ideas also were measured. At the end of the brainstorm session the prototype generated a code that participants could copy and paste into the questionnaire. A code might look like: “adj1-8/adj2-5/adj3-0/T-13”. This indicates that the participant generated 8 ideas for the first adjective, 5 for the second adjective, 0 for the third adjective and thus had 13 ideas total. The adjectives that are missing from this code mean that the participant has not activated these adjectives. So in this case he or she only had three adjective cards active. This code can both provide a comparison between the total ideas, and the ideas generated per adjective.

### 3.3 Participants

The participants for this study have been recruited on convenience. No selection was made based on academic background or field of study. In total 22 participants between the age of 20 and 33 were recruited and 20 participants finished the study. One participant could not complete the study as his browser crashed near to the end. Another finished the study but contacted the researcher afterwards that he misunderstood the assignment and brainstormed about something else, the results have therefore not been taken into account. Out of the remaining 20 participants, 11 were male and 9 were female.

### 3.4 Limitations

Using a within-subject design for the control group and the experimental group has the benefit of reducing the error variance associated with individual differences; All ‘personal’ factors between the two groups are the same. However, this introduces the problem that the two sessions, mainly the brainstorm topics, need to be similar in the way of depth, fun and
Although the two questions have been made somewhat similar on a theoretical level, it is to be seen how equal they actually are in practice, as people might find it easier to imagine the thought of being a student rather than an elderly, or the other way around.

Another issue is that participants might rate themselves in the second round with the first round as a benchmark. For example they might feel that their performance is lower or higher than the first round, and will make their ratings based on that rather than on the round itself. It is however assumed that if that behaviour occurs, it should occur approximately the same in both the groups that have gamification first and gamification second. This would make the effect the same on both sides, but the effect should still be taken into consideration.

Letting participants do the tasks and the questionnaire on their own time has many advantages and disadvantages. This is not only useful for participants to be able to choose their own time of completing the task, but also makes sure they do not feel evaluated while doing the task. This might have been the case if the researcher was present. A disadvantage, however, is that there was no control over if the participants truly read and understood the assignment. One participant, for example, had to be disqualified because he brainstormed solely based on the adjectives. With this study setup, there is no way of telling if this happened to other participants as well. Another disadvantage of having no control is that participants might have done it in different settings, they might have done it at different times of the day or in different environments.

Because some adjectives might spark more ideas for one person than the other, the adjectives were completely randomized. This randomization was done from a pool of 60 adjectives. Randomization was the preferred choice to overcome this bias towards certain adjectives. In addition the study was not concerned in comparing the number of ideas per specific adjective (e.g. "Impolite"), so there was no need for using the same adjectives for each participant. Lastly, associative brainstorming itself also advocates using random stimuli.

Using a factorial design resulting in four questionnaires has the advantage that it can account for more variables, such as getting tired or becoming trained in the activity. But it also means that each group gets smaller, with five participants in each. This number is too small to perform any significant statistical tests. However, since the assignment of participants in each group was pseudo-random (each participant was assigned randomly to a group until the groups were the size of a quarter of the sample size) it is assumed that both groups from the 2x2 design can be used when testing statistics, resulting in 10 people in the same condition. For example, when testing the significance of the topic order, both questionnaires with topic one first can be compared against both questionnaires with topic two, regardless of the gamification order.

A quantitative approach has been chosen to analyze the differences in gamification for brainstorming. However, by introducing a prototype that is unknown to people, it could introduce variables that can not be accounted for in a comparative study. It would, therefore, have been good to do a qualitative study beforehand, aiming at what people think of brainstorming alone with the prototype and how it fits their brainstorming habits. Due to the limited amount of time available for this study, the choice has been made not to do this, but it is critical to consider it might have an influence on the study outcomes.
4. Results

This chapter will explore the results gathered by the study. The analysis consisted of multiple steps. Section 4.1 shows an overview of the results followed by a detailed analysis in section 4.2. Each subsection in 4.2 will address a different analysis. Section 4.3 then concludes and summarizes the most important findings.

4.1 Result overview

The results of the average scores for each dependent variable can be seen in illustration 5 and 6. Illustration 5 shows the means for all the dependent variables that have been rated on a scale from one to seven, including their standard deviations. Both the normal and gamified versions are shown in the graph. It can be seen that the normal condition is always slightly higher than the gamified condition with the biggest gap existing in the performance metric. Illustration 6 shows the mean and median for the number of ideas in both the gamified and normal versions of the prototype. It shows that the mean of the normal version is again higher than the gamified version, but the median is exactly the same. The standard deviations are big for this variable, indicating a big variation in the number of ideas across participants.

Illustration 7 and 8 show the averages for the dependent variables per round and their respective standard deviations. It can be seen that people rate their performance slightly higher in the second round, but all the other variables are rated lower, including how many ideas people wrote down on average.

Illustration 5: Averages of the dependent variables Enjoyment, Motivation and Performance for the gamified and normal version of the brainstorming prototype.
Illustration 6: Mean and median of the number of ideas for the gamified and normal version of the brainstorming prototype.

Illustration 7: Means of the dependent variables for each round.
4.2 Data analysis

This subsection contains all the detailed analysis of the results. Section 4.2.1 looks at the analysis of how gamification affected the dependent variables. 4.2.2 looks at whether the order of gamification mattered to the results. Similarly, 4.3.2 looks at whether the topic had a significant influence on the results. Section 4.2.4 analyses the differences in rounds regardless of whether the experimental version was used. 4.2.5 investigates whether people brainstorming before had a different result than people who have not. Section 4.2.6 looks at the number of ideas per adjective in the game. Lastly, section 4.2.7 gives a brief overview of the comments and attitudes people gave towards the tool.

4.2.1 How did the gamified version influence the dependent variables

The outcomes in mean differences between the gamified and non-gamified version show that only performance is significant, favoring the normal version over the gamified version. All other variables are non-significant. For this analysis all the gamified version results will be compared against the normal version results. Testing for normality the Kolmogorov-Smirnov test shows all variables are non-significant with a \( p > .05 \), except for performance in the normal condition with \( p = .008 \). Paired sample \( t \)-tests will be calculated for all except for performance where a Wilcoxon signed ranked test will be done.

The average enjoyment for the gamified version of the prototype (\( M = 4.95, SE = .28 \)) does not differ significantly with the normal version of the prototype (\( M = 5.14, SE = .27 \)), \( t(19) = -.877, p = .195, r = .20 \). However not significant it does show a mean difference of .19 in favor of the normal condition.
The same difference can be observed for motivation where there is another .19 in favor of the normal condition. Motivation for the gamified version ($M = 5.2, SE = .33$) was however also not significantly different in relation to the normal version ($M = 5.4, SE = .28$), $t(19) = -.748, p = .232, r = .17$.

The self-rated performance in the gamified condition ($Mdn = 4.00$) is significantly less good than performance in the normal condition ($Mdn = 4.67$), $T = 31.5, p = .004, r = .57$. The mean difference of .76 in favor of the normal condition indicates that people favored their performance in the normal condition.

The average number of ideas in the gamified condition ($M = 13.8, SE = 1.67$) again shows no significant difference from the average number of ideas in the normal condition ($M = 14.1, SE = 2.46$), $t(15) = -.193, p = .425, r = .05$. The difference mean is again in favor of the normal condition with .30 ideas more in this condition.

### 4.2.2 Did the order of gamification make a difference?

Both the gamified and the normal version have comparable means across all variables when grouped by whether gamification was first. This indicates that the order of gamification had no effect on the results. For the analysis all the summarized dependent variables will be taken and grouped by gamification. So regardless of in which round the participant encountered the gamification or normal condition, their results for the dependent variables will be combined. As shown in the previous section, all variables are approximately normally distributed apart from performance in the normal version. An independent sample $t$-test will be done, with the grouping variable set to gamification. A Mann-Whitney U test will be done for the performance variable.

The test shows that on average, enjoyment of the gamified version with the gamified first condition ($M = 5.35, SE = .24$) did not significantly differ from the gamified second condition ($M = 4.55, SE = .50$), $t(18) = -1.45, p = .163, r = .32$. Enjoyment for the normal version also did not significantly change between the gamified first condition ($M = 5.28, SE = .32$) and the gamified second condition ($M = 5.00, SE = .45$), $t(18) = -.498, p = .624, r = .12$. Enjoyment therefore seems to not have been influenced by the order of gamification. The same hold true for motivation, in the gamified version the motivation for the gamified first condition ($M = 5.55, SE = .40$) is not significantly different from the gamified second condition ($M = 4.85, SE = .51$), $t(18) = -1.07, p = .297, r = .25$. Motivation for the normal version for the gamified first condition ($M = 5.28, SE = .35$) were also not significantly different from the gamified second condition ($M = 5.50, SE = .46$), $t(18) = .389, p = .702, r = .09$.

Performance for the gamified version with the gamified first condition ($M = 3.90, SE = .36$) did not diverge from the gamified second condition ($M = 4.30, SE = .45$), $t(18) = .49, p = .531, r = .11$. Since performance for the normal condition was not normally distributed a Mann-Whitney-U test has been done. It shows that the gamified first condition ($Mdn = 4.67$) did not significantly differ from the gamified second condition ($Mdn = 4.67$), $U = 42, p = .579, Z = -.610$. Lastly also the number of ideas did not seem to be significantly different in the gamified version, where the gamified first condition ($M = 14.60, SE = 2.10$) and the gamified second condition ($M = 12.70, SE = 2.36$) showed a mean difference of 1.9 ideas, $t(15) = -.561, p = .583, r = .15$. The normal version had an average difference of ideas of 1.33.
where the gamified first condition ($M = 15.40, SE = 1.80$) did not significantly differ from the gamified second condition ($M = 12.60, SE = 4.20$), $t(16) = -.291, p = .775, r = .07$.

So the order of gamification does not seem to have much effect on the dependent variables for the different versions of the prototype. Both the gamified version and the normal version each have comparable means for the two situations.

4.2.3 Did the topic make a difference?

The topic did not matter for the outcomes but shows a non-significant trend towards topic 2. For this analysis all the summarized dependent variables will be taken and grouped per topic. So regardless of in which round the participants encountered topic one or two, their enjoyment, motivation, performance and number of ideas will be compared. Each dependent variable per topic seems to follow an approximate normal distribution with $p > .05$ on the Kolmogorov-Smirnov test except for motivation on topic one with $p = .340$. A paired sample $t$-test shows that on average, the ideas for topic one ($M = 13.30, SE = 1.72$) did not significantly differ from the ideas of topic two ($M = 14.60, SE = 2.42$), $t(15) = -.826, p = .422$; and it represented a small-sized effect $r = .20$. The mean in enjoyment for topic one ($M = 5.03, SE = .28$) hardly differed anything from the enjoyment for topic two ($M = 5.06, SE = .27$) and was also not significant, $t(19) = -.172, p = .864, r = .04$. The performance between topic one ($M = 4.40, SE = .31$) also did not significantly differ from the performance in topic two ($M = 4.60, SE = .25$), $t(19) = -.759, p = .457, r = .17$. Since motivation was not approximately normally distributed a Wilcoxon signed ranks test was done. Motivation for topic one ($Mdn = 5.5$) were not significantly different from the motivation for topic two ($Mdn = 5.38$), $T = 55.5, p = .536, r = -.15$. So the topic did not matter significantly across all the dependent variables. However, the mean seems to be slightly higher for topic two with each dependent variable, showing a non-significant trend leaning towards topic two.

4.2.4 Differences in rounds

When analyzing the results in 4.2.1 it shows that the means are always higher in the normal condition (see illustration 5 and 6), however, 4.2.2 shows that the means for gamification are higher in the gamification first condition. This shows signs of exhaustion and therefore an analysis of differences in rounds was performed (see illustration 7 and 8). The analysis shows that performance was higher in the second round, but all other variables were lower in the second round, where motivation was significantly lower. For this analysis all totals for the dependent variables for round one were compared to round two. Each variable seems to be approximately normally distributed, with an insignificant Kolmogorov-Smirnov test for each. Each dependent round will therefore be analyzed with a paired sample $t$-test.

Enjoyment for round one ($M = 5.18, SE = .25$) has a slightly bigger mean than enjoyment for round two ($M = 4.91, SE = .30$), but the difference is not significant, $t(19) = .426, p = .113, r = .10$. Motivation for round one ($M = 5.53, SE = .31$) is significantly different from motivation of round two ($M = 5.06, SE = .30$) with a mean difference of .46, $t(19) = 2.00, p = .03$, and has a medium effect size with $r = .42$. Performance in the first round ($M = 4.40, SE = .28$) is the only dependent variable that has a lower mean than round two ($M = 4.60, SE = .28$), however not significant, $t(19) = -.647, p = .263, r = -.15$. Lastly the number of ideas in round one ($M = 14.31, SE = 2.47$) was also higher than the number of ideas in round two ($M = 15.40, SE = 1.80$).
= 13.6, SE = 1.65) with a mean difference of nearly .70. This difference is however also not significant, \( t(15) = .426, p = .338, r = .11. \)

It seems that self-rated performance was higher in the second round, but all other variables were on average lower in the second round. Of these, only the difference in motivation was significant.

### 4.2.5 Does brainstorming before affect the results?

In the questionnaire a question was asked whether the participant brainstormed before, and if yes if he or she did that alone, in groups, or both. This section will look at how these groups performed in the results. None of the participants only brainstormed alone, leaving three groups. Table 1 shows the mean averages for each group, their standard deviations, and how many participants were in each group.

<table>
<thead>
<tr>
<th>Brainstormed Before</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never brainstormed before</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 1 Gamified enjoyment</td>
<td>5.0833</td>
<td>6</td>
<td>.98319</td>
<td>.40139</td>
</tr>
<tr>
<td>Enjoyment Normal</td>
<td>5.6250</td>
<td>6</td>
<td>.78661</td>
<td>.32113</td>
</tr>
<tr>
<td>Pair 2 Motivation Gamified</td>
<td>5.3333</td>
<td>6</td>
<td>.99382</td>
<td>.40654</td>
</tr>
<tr>
<td>Motivation Normal</td>
<td>6.2083</td>
<td>6</td>
<td>.57715</td>
<td>.29444</td>
</tr>
<tr>
<td>Pair 3 Performance Gamified</td>
<td>4.0657</td>
<td>6</td>
<td>.36893</td>
<td>.15978</td>
</tr>
<tr>
<td>Performance Normal</td>
<td>5.1117</td>
<td>6</td>
<td>1.02638</td>
<td>.41520</td>
</tr>
<tr>
<td>Pair 4 Number of Ideas Gamified</td>
<td>12.0000</td>
<td>6</td>
<td>4.14729</td>
<td>1.69312</td>
</tr>
<tr>
<td>Number of ideas Normal</td>
<td>10.5000</td>
<td>6</td>
<td>3.08221</td>
<td>1.25831</td>
</tr>
<tr>
<td>Brainstormed only in groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 1 Gamified enjoyment</td>
<td>4.8500</td>
<td>5</td>
<td>.97788</td>
<td>.43732</td>
</tr>
<tr>
<td>Enjoyment Normal</td>
<td>4.5500</td>
<td>5</td>
<td>.54199</td>
<td>.24238</td>
</tr>
<tr>
<td>Pair 2 Motivation Gamified</td>
<td>5.5500</td>
<td>5</td>
<td>1.00623</td>
<td>.45000</td>
</tr>
<tr>
<td>Motivation Normal</td>
<td>4.9000</td>
<td>5</td>
<td>.57722</td>
<td>.24495</td>
</tr>
<tr>
<td>Pair 3 Performance Gamified</td>
<td>3.6660</td>
<td>5</td>
<td>1.10276</td>
<td>.49318</td>
</tr>
<tr>
<td>Performance Normal</td>
<td>3.9660</td>
<td>5</td>
<td>.69226</td>
<td>.30599</td>
</tr>
<tr>
<td>Pair 4 Number of Ideas Gamified</td>
<td>15.5000</td>
<td>4</td>
<td>8.96289</td>
<td>4.48144</td>
</tr>
<tr>
<td>Number of ideas Normal</td>
<td>14.2500</td>
<td>4</td>
<td>6.55106</td>
<td>3.27554</td>
</tr>
<tr>
<td>Brainstormed both in groups and alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pair 1 Gamified enjoyment</td>
<td>4.0187</td>
<td>9</td>
<td>1.64112</td>
<td>.54804</td>
</tr>
<tr>
<td>Enjoyment Normal</td>
<td>5.1368</td>
<td>9</td>
<td>1.59644</td>
<td>.53215</td>
</tr>
<tr>
<td>Pair 2 Motivation Gamified</td>
<td>4.9167</td>
<td>9</td>
<td>1.94454</td>
<td>.68418</td>
</tr>
<tr>
<td>Motivation Normal</td>
<td>5.1111</td>
<td>9</td>
<td>1.63989</td>
<td>.54663</td>
</tr>
<tr>
<td>Pair 3 Performance Gamified</td>
<td>4.4007</td>
<td>9</td>
<td>1.68814</td>
<td>.56005</td>
</tr>
<tr>
<td>Performance Normal</td>
<td>5.2967</td>
<td>9</td>
<td>1.04716</td>
<td>.34605</td>
</tr>
<tr>
<td>Pair 4 Number of Ideas Gamified</td>
<td>14.5000</td>
<td>6</td>
<td>7.86756</td>
<td>3.21996</td>
</tr>
<tr>
<td>Number of ideas Normal</td>
<td>17.9567</td>
<td>6</td>
<td>14.98888</td>
<td>5.11919</td>
</tr>
</tbody>
</table>

Table 1: Dependent variables for gamified and normal split by whether a participant brainstormed before.

Looking at the average means from table 1 it appears that both the "never brainstormed before" group and the "brainstormed both in groups and alone" group preferred the normal version on each of the variable. The remaining group, which only brainstormed in groups before, seemed to have a higher rating on both enjoyment, motivation and the number of ideas. This group was however also the smallest group with 5 participants. When looking at
the results of the paired test, none are significant at the $p < 0.05$ level and thus brainstorming before did not affect the results.

4.2.6 Does the amount of ideas per adjective differ?
Because the prototype was based around associative brainstorming with different adjectives, it is interesting to see how the scores between adjectives changed between the two version of the prototype. The number of ideas per adjective are roughly similar for the gamified and normal versions, much like the total of ideas as seen in illustration 4. However, as seen in illustration 7, the median scores for the gamified version were higher for the first two adjectives, the same up to adjective six and one lower for the last two adjectives. As the scoring mechanism of the gamified prototype had a cutoff point of four, this could indicate that people have spent more time writing ideas for the earlier adjectives in the gamified version to get points. In the end, they might not have had time left, resulting in a lower median.

![Median score per adjective](image)

*Illustration 9: Median scores per adjective for both the normal and gamified version of the prototype.*

4.2.7. Participants comments and attitudes
Half of the participants left comments concerning how the prototype worked, sometimes against their expectations, or how they liked or did not care for the gamified version. People that expressed a negative attitude towards the gamification said, for example:

- "I found it interesting that it was easier to pick adjectives without the addition of the stars. I feel it was easier to pick and choose while [the stars] weren't there, as it made it more of a game of reaching as many stars as possible, as opposed to generating as many ideas as possible."
- "The tool is very interesting, I don't think the stars are extremely important though."
"The stars did not matter for me, because I noticed that once I got these stars I couldn’t actually do anything with them. It sold me hot air!"

The people expressing a positive attitude towards the gamification mentioned:

- "Setting goals with the stars certainly induce a "just one more idea" behavior per adjective."
- "I prefered the second one [gamified version]. I think it forces you to [write down] ideas you wouldn’t say. Regardless of the quality.
- "For me the [gamified] tool felt more "efficient" since it "forced me" to stick with an adjective longer"

Some comments concerned the prototype itself:

- "Very interesting idea and I found it to be a lot of fun, though slightly stressful with the time limit, but I guess that is just my competitiveness striking out""
- "I felt like the tool in itself gave me restrictions. The words that were supposed to give me direction, gave more of a sense that I "should" write something that had to do with the given word. The layout of the “idea boxes” was a bit frustrating. It would be good to be able to color coordinate or sort it in another way."
- "You should extend the app to allow drawing circles around a group of ideas and name the circle"
- "It was easier to just pick a new adjective and go with ideas there than to think of 4 complete new ones for the previous adjective."

From the comments it seems the gamified tool was preferred by some, but neither was preferred by others. Some participants found the prototype fun to use, regardless of the version, others felt like it gave the restrictions on how they would actually want to use it. Lastly, the idea of having "finished" adjectives was mentioned in the last quote, so always getting a fresh adjective rather than getting new ones.

4.3 Summary of findings

Mostly the analyses have shown no significant differences across multiple tests. However, there was a significant difference in self-rated performance, favoring the normal condition over the gamified condition. This indicates participants favored their performance in the normal condition over the gamified condition. The order of gamification, so whether the gamification or normal prototype was used first, did not significantly influence the results. The brainstorming topic also did not significantly influence the results. When analysing the differences in round one and round two, there was a significant difference in motivation, where the motivation in round two was significantly lower. This could indicate an exhaustion effect during the test. The test whether participants brainstormed before influenced the results, concluded in that it had no significant effect. Lastly, when looking at the ideas per adjective, the gamified version shows a higher median for the first two adjectives, equalling the score needed to gain a point. This could indicate that participants spend more time at the start trying to get points.
5. Discussion

The results indicate that there is no significant difference in enjoyment, motivation and the number of ideas in the gamification and normal versions. For all of these outcomes, the average scores for the normal version were slightly higher than the gamified version. Performance was the only significant difference measured, favoring the normal condition. There might be several reasons that measures on three out of four dependent variables showed no significant differences. First, it might be that gamification really does not make any difference, or it might have been the sample, both the size and the selection. Thirdly it might have been the way the prototype was introduced in the study and the design of the prototype itself. It could have also been how the study implemented gamification.

If the first reason is true, gamification does not improve the experience of brainstorming. Since gamification has been applied successfully to brainstorming in the past, however, this seems unlikely and one or a combination of the other reasons are probably the right explanation. The sample size was 20 people in a within-subject design, making it 40 observations. While this should be enough to get statistically significant data, it is still considered a small sample and having a larger sample would have drastically increased the likelihood of finding any underlying significant difference. The selection could have also impacted the results. Since the study gathered participants that did not have the need to brainstorm themselves, they were likely less motivated than people who would need to brainstorm. If the extrinsic motivation was higher, so people would want to get something out of the session, the results might have been a lot different. Additionally, the study included six participants that had never brainstormed before, and they might have a completely different view of the tool than people who have brainstormed before. Also, for segmenting participants, a question could have been asked about whether the participants were gamers, and how much games they play. Perhaps the gamification elements work better on hardcore gamers than on nongamers.

The third possible reason is the implementation and introduction of the prototype. Many respondents found the tool fun to use, as the average enjoyment rating was above the middle on the one to seven scale. The drag and drop interface it uses has some playful aspects to it already. So the prototype being fun to start with, might have made it hard to measure a difference between gamification and non-gamification for enjoyment or the other dependent variables. A partial solution for this would have been to start the study by introducing the prototype first, without measuring anything. This way the user could have gotten a feel for how the prototype worked and the novelty might have worn off, making the study more accurately measure the brainstorming activity and how the prototype supported this.

Another argument to have introduced the prototype first was that participants seemed to have a different approach on how to use the prototype. Some participants, for example, used idea cards as topics to cluster ideas, instead of clustering ideas by moving them around on the board. This sometimes resulted in unexpected behaviour from the prototype. However this could have also been mitigated by performing the study in a room where the researcher was present or doing a pre-study to develop the prototype based on the wants and needs for an electronic brainstorming tool that lets you brainstorm alone.
The fourth possible reason mentioned as to why the results are mostly insignificant is the way the study implemented gamification. The prototype used a 4 idea milestone to score points, using progress bars to indicate how far the participant was towards scoring this point. The reason these motivational affordances were chosen is that it would motivate people to write down all their ideas, regardless of their quality to reach the milestone. Also it was thought to help with overcoming self-evaluation of ideas, as people would dispose of their ideas less quickly if they got a reward for them. However, this might have had the opposite effect, where people could not reach the milestone, thus not thinking their performance was adequate enough. This could possibly also explain why people did not favor their own performance in the gamified condition over the normal condition. Another approach would have been to reward the participants for each idea they would write down, maybe not in the form of a point, but in some other way. This way the motivation to reach for the point can stay but each idea would be its own reward.

Another point of interest in this study is the non-significance of the topic and the differences in rounds. The topic seemed not to have mattered concerning the results, however, there was a non-significant trend favoring the second topic, which was about students. Some participants wrote that they had a much easier time writing ideas for this second topic too. A bigger sample size could improve the understanding of this trend. The differences in rounds show that motivation was significantly lower in the second round, this could indicate that there was an exhaustion effect in the study. The study might have been too long, exhausting the participants and making them less enthusiastic about generating ideas. As most brainstorming studies are between 12 and 20 minutes long, this study adhered to that format and made each brainstorming session last 12 minutes. However, most studies did not implement a within-subject design and therefore only had one round of brainstorming. This study had two rounds, making the total brainstorming time 24 minutes, which might have been too long. A solution for this could be to half the time for each round, making the total brainstorming time 12 minutes.

When analyzing the results per adjective, it became clear that the first few adjectives in the gamified version had the milestone score of 4 more often than in the normal version. This could indicate that people stuck with an adjective longer in the gamified version to score the point. So in this sense, the gamification worked to let people consider stimuli for a longer time. However, in the normal version people had a higher score for the later adjectives, indicating that it was easier to write down all the first ideas they had. This could indicate that a combination would be interesting, where people write down all their initial ideas per stimulus, but can then spend time on developing more ideas per adjective. Alternatively, this implementation of gamification could be tried on a smaller scale, by having less time or just fewer stimuli.

The results show that this form of gamification does not make a significant difference in outcomes of individual brainstorming, except for a decrease in self-rated performance in the gamified condition. Other forms of motivational affordances as listed by Hamari, Koivisto, & Sarsa (2014) might work better or differently to overcome negative aspects associated with brainstorming. A problem listed by the participants was that the points did not matter to them, as they could not do anything with them. This might be solved in ideation teams, where the second step of the brainstorming process would be the group stage where
everybody evaluates ideas. If the team members could use these points in the second stage, points would have a purpose and people might be more motivated to gain these points.

6. Conclusion

This study shows that a gamified tool does not overcome negative aspects associated with brainstorming individually. No difference was found in enjoyment, motivation and the number of ideas generated between using a gamified tool or a normal tool. The gamified tool additionally made people feel like they performed less well than a non-gamified tool. Milestones and progress towards these milestones make people write down more ideas per category, but do not make people more motivated to write these ideas or make them write down more ideas in total.

To overcome these negative aspects, another approach might be taken to applying gamification to ideation. Maybe other motivational affordances or another application of the same motivational affordances might give a better or different result. The context of ideation also matters a great deal, so making these gamified tests in the context of a real need for ideation could also give a better or different result. An alternative is to try to overcome these negative aspects by combining ideation with a different field than gamification.

Since some participants in this study said they did prefer the gamified version, it might just come down to preference. So gaining points by achieving milestones might be a good alternative for some people that like this better than a regular brainstorm session. For the design community, it means that applying gamification to products relating to ideation might not always be beneficial for all or any of the users. And that when applying gamification, other approaches to gamification should be taken.

6.1 Future research

Based on this research it might be beneficial to look at which motivational affordances work best with brainstorming, and which result each of them has concerning the negative aspects related to brainstorming alone. This research should perhaps be of a qualitative nature, to see why people prefer certain affordances over others. Additionally, any further research would benefit from separating gamers from nongamers when looking at gamification, or at least measure the participant’s attitude towards games.

A further study into overcoming negative aspects in brainstorming alone might consider including the entire brainstorming process, so including the group stage and the evaluation of the ideas. People will have the feeling their work matters more than if they would quit right after the brainstorming alone phase. This contextualised study should also look at possibilities of using the scores or other outcomes from the gamified sessions and taking them to the group stage, making the outcomes mean something. Additionally, whenever possible the study should use people that actually have a need to brainstorm about ideas, perhaps in a company atmosphere.

Lastly, this study does not give a convincing argument for gamification to overcome the negative aspects of brainstorming alone. Future studies should look at other branches of research to make brainstorming alone more enjoyable and more motivating while making the feeling of performance higher as well.
References


Appendix 1: Questionnaire introduction text example for gamified version

You will hold two brainstorm sessions on the website linked below, each of 12 minutes (timed by the application). After each session there will be a few questions on this form to fill out. The application will also generate a code, please copy and paste that into the designated field in this form. Note that each session has a different link to the application.

Please fill in these three questions and take your time reading the instructions below. [Question about age, gender and whether they brainstormed before]

For this session please adhere to the following brainstorming rules:
1. Quantity over quality; In this stage you want as many ideas as possible.
2. Try to not criticize your ideas, all of your ideas are useful. Even the wild ones!
3. Try to build on your older ideas, for example by combining two, and find new possibilities.

The application (see the picture below for a preview) lets you write and group ideas. Each idea can be written on a card (#2 on the picture below). To help you generate creative ideas the application gives you adjectives represented above each card (#1 on the picture below).

Stuck on getting new ideas and need a new adjective? Press the button (#3 in picture) to get a new adjective and card color. However, there is only a limited number of adjectives you can use, so try to use each adjective to generate ideas!

It also uses a drag and drop interface, so after filling in an idea on a card, you can drag and drop it onto the board under it to add it to your list of ideas. You can then drag and drop to order and sort ideas as you please.

Please write one idea per card.

The game tracks your ideas with stars, for each 4 ideas you get per adjective, you earn a star. Try to get as many as possible!

The 12 minutes will start as soon as your first idea is placed on the board.

The brainstorm topic for this session is:
What can universities do to introduce student life and all its aspects to new students

Illustration 10: Intruction image gamified version
Appendix 2: Questionnaire questions

**Interest / Enjoyment (intrinsic motivational values)**
I enjoyed doing this activity very much
This activity was fun to do.
I thought this was a boring activity. (R)
I would describe this activity as very interesting.

**Value / Usefulness / Goal achieving (extrinsic motivational values)**
I think this activity helped me to achieve my goals
I think that doing this activity is useful for generating ideas.
I would want to use this tool again if I have the need to generate ideas.
I think using this tool helped me generate ideas more ideas than I would without this tool.

**Perceived productivity / performance**
I think I did well on this task
I think I produced a high number of ideas
I think the ideas I produced are of high quality
If I was in a group with three others, I think I would have generated more ideas.
If I was in a group with three others, I think I would have generated better quality ideas.