The effect of varying degrees of anonymity on brainstorming

Author: Stefan Hellweg
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT

Brainstorming is a much-utilized tool to generate ideas, but the ideal way of organizing a brainstorming session has not been identified yet. Previous research on the effect of anonymity on brainstorming has been inconclusive. This study aims to investigate the relationship between varying degrees of anonymity and the effectiveness of a brainstorming session. *Varying degrees of anonymity* in this context refers to which ideas could be directly linked to the identity of the person that generated them. It was theorized that selective anonymity (a form of anonymity where only the top 10% of ideas could be linked to the creator) would prove to be the best way of conducting a brainstorming session and that the variables *evaluation apprehension* and *free-riding* would act as mediator variables. This was tested by creating an electronic nominal brainstorming session and randomly assigning participants into one of three groups (full anonymity, no anonymity, selective anonymity). The number of high-quality ideas generated per person was compared between the groups with a Kruskal-Wallis H test which showed no significant differences between the groups. An ANOVA test showed no significant effect from the degree of anonymity on evaluation apprehension or free-riding. The ordinal regression analysis showed that evaluation apprehension and free-riding were mostly insignificant predictors for the number of high-quality ideas generated. These results suggest that the degree of anonymity does not affect the effectiveness of an electronic nominal brainstorming session and that the variables evaluation apprehension and free-riding do not act as mediators in this context.

Graduation Committee members:

Dr. Tim Schweisfurth, Dr. Dorian E. Proksch

Keywords

Electronic nominal Brainstorming, Evaluation apprehension, Social-loafing, Anonymity, Idea generation, High-quality ideas

1. INTRODUCTION

Brainstorming in theory, is a group problem-solving technique that is based on writing down all the ideas of all group members without judgement and afterwards evaluating which idea is the best solution to the problem. Since brainstorming is a practice that any kind of group can use to generate new creative ideas, it is very common in organizations like corporations, educational institutions or political organizations. Naturally, everybody involved wants to optimize the technique to generate ideas as efficiently as possible, and it has therefore been the subject of research for decades. Researchers have already investigated the effectiveness of brainstorming sessions based on many different variables like group size or evaluation apprehension, but the results have not been conclusive.

There are multiple studies that suggest a positive effect of anonymity on brainstorming. An article by L. Mabel Camacho and Paul B. Paulus (1995) concludes that social anxiety has a negative effect on interactive brainstorming. Anonymity in brainstorming aims to reduce this effect by removing the social factor from the idea generation phase. Another study by Michael Diehl and Wolfgang Stroebe (1987) examined the different factors that lead to productivity loss when moving from an individual brainstorming setting to a group setting. The factor that was found to explain the most significant percentage of productivity loss was blockage. Blockage refers to the negative effect of group settings which is that only one person at a time can talk, and others must listen and wait for their turn to contribute to the conversation. Evaluation apprehension was also found to explain a significant amount of productivity loss. Although anonymity would be a solution to counterfeit the effects of evaluation apprehension, in theory, a study by Shepherd et al. (1995) argues that anonymity is to be seen as a "mixed blessing" because it supposedly solves the problem of evaluation apprehension but introduces the problem of social loafing. The study found that social comparison reduces social loafing but still deemed anonymous brainstorming the superior approach over identified brainstorming. In this context, social loafing can be defined as "the tendency of [group]members to do less than their potential" (Chidambaram & Tung, 2005) (more on this in 2.3). This inconsistency in the literature leads to the belief that neither approach may be the best solution to brainstorming.

The objective of this research is to evaluate the effectiveness of a third approach: selective anonymous brainstorming. For this way of brainstorming, all the ideas will be evaluated anonymously. Only the top 10% of ideas will be revealed, including the name of the idea creator. The effectiveness of selective anonymous brainstorming will be compared to the previously mentioned approaches of no anonymity and full anonymity. Selective anonymous brainstorming might be the best approach since it supposedly combines the best of both worlds. On the one hand, there is the benefit of removing the fear of evaluation apprehension by anonymizing the submitted ideas, but on the other hand, there is the benefit of reducing social loafing or free-riding by adding an element of social comparison. This research will show if revealing only the best ideas has the same effect on free-riding as other forms of social comparison.

Previous studies have utilized several methods to evaluate a brainstorming session's effectiveness. Variables such as the number of unique ideas generated or the quality of the generated ideas frequently appear (Mullen et al., 1991; Michinov, 2012). Since only the best ideas from a brainstorming session will be pursued, it is logical to evaluate the effectiveness of a brainstorming session by investigating only the best ideas. A quantitative approach when comparing brainstorming groups would be to look at the number of high-quality ideas generated.

When taking into consideration the current state of the research on this topic and the resulting research objectives, the following research question can be concluded:

How do varying degrees of anonymity affect the number of high-quality ideas generated during a brainstorming session?

2. LITERATURE

2.1 Brainstorming and its pitfalls

Advertising executive Alex Faickney Osborn first popularized the term brainstorming in his book Applied imagination in 1957. According to Osborn, brainstorming can be used to solve simple problems as a group. Furthermore, he argues that solving creative problems in a group can lead to synergy effects, where people can build on each other's ideas. Since then, many researchers have conducted experiments based on this concept introduced by Osborn to test the validity of his claims.

More recent studies contradict Osborn's findings. It has been shown that concepts such as evaluation apprehension, social anxiousness, production blocking and free-riding negate the synergy effects of brainstorming in a group setting (Camacho, L. M., & Paulus, P. B. 1995; Diehl, M., & Stroebe, W. 1987). These concepts are commonly referred to as productivity loss. Another study (Mullen et al., 1991) compares the effectiveness of brainstorming groups (interactive groups) as described by Osborn to the effectiveness of nominal (non-interacting) groups. It has been found that nominal groups outperform brainstorming groups in terms of quantity and quality of ideas. Bigger group sizes and the presence of researchers have been shown to increase the effects of productivity loss. With these findings on the pitfalls of brainstorming, researchers have since tried to develop improvements for the original technique.

2.2 Evaluation apprehension

The already mentioned concept of *evaluation apprehension* is a commonly named factor affecting the outcome of a brainstorming session. It refers to the social phenomenon of participants hesitating to contribute ideas due to the fear of judgement from others. A study by (Zhou et al., 2019) found that evaluation apprehension had a negative impact on the quantity and diversity of ideas generated within a collaborative setting.

2.3 Social loafing

Social loafing (from here on, referred to as *free-riding*) can be defined as an individual's lack of participation due to a lack of intrinsic and extrinsic motivation. The concepts suggest a connection between an individual's participation in a group project and the rewards related to the amount of effort put in. Participants might put in less effort if they feel like the efforts of others are already enough to lead to a sufficient group output or if they can not see a direct relationship between their effort put in and the resulting rewards. It has been found that the task effort of individuals in an electronic brainstorming session was higher when the participants were identified compared to anonymous (McLeod, 2011).

2.4 Facilitators

It has been found that the use of facilitators can lead to a better performance of interacting groups up to the point of these groups performing on a similar level as nominal groups (Oxley et al., 1996, Offner et al., 1996). The role of a facilitator in a brainstorming session is to actively encourage participants by supporting their ideas and generally stimulating their idea generation process. Although the effects of facilitators were not further investigated in this study, they will be important for the

practical implications (6.3) and the recommendations (6.5) resulting from this study.

2.5 Types of brainstorming

Further research has been conducted on the topic of electronic brainstorming in comparison to verbal brainstorming. Experiments show that electronic brainstorming leads to less production blocking and evaluation apprehension, especially when combined with larger group sizes (Gallupe et al., 1992). Contrary to these findings, it was claimed that the empirical evidence that electronic brainstorming was superior to verbal or nominal brainstorming was weak (Pinsonneault et al., 1999). The experiment that they conducted suggests that nominal brainstorming yields the best results. A third alternative to the classic verbal brainstorming method and electronic brainstorming is *brainwriting*. This technique involves participants silently writing down ideas on paper and passing them on to their neighbors. Using different colored pencils for different group members is supposed to increase accountability and seeing the ideas that other members have written is supposed to positively stimulate the idea-generating process. It was found that when using this technique, groups outperformed nominal groups that wrote down ideas individually (Paulus & Yang, 2000). A more recent study compared the two approaches of brainwriting and electronic brainstorming with exciting results: It was found that participants in the electronic brainstorming group generated more non-redundant and less redundant ideas than the participants of the brainwriting group (Michinov, 2012).

2.6 Anonymity

Because of the findings on electronic brainstorming, the assumption could be made that the anonymity that electronic brainstorming provides compared to verbal brainstorming was the main reason for the improvements in evaluation apprehension. Several researchers have since investigated the effects of anonymity on brainstorming with inconclusive results. While there has been reported to be no effect of anonymity on ideational performance (Valacich et al., 1992), others have found an increase in flexibility when anonymity was combined with a transformational leadership approach (Sosik et al., 1998). In the context of this study, flexibility was seen as one dimension of creativity. The inconclusive findings regarding anonymity in brainstorming suggest that the ideal way of conducting brainstorming has not been identified yet.

3. THEORY

3.1 Dilemma

Considering the research on the topic of brainstorming that has been elaborated on above, evaluation apprehension and free-riding seem to be essential variables for the effectiveness of a brainstorming session. The factor of anonymity appears to play an important role for influencing these variables. Still, there is a problem: As already elaborated in the literature review, previous studies have suggested that full anonymity would lead to less evaluation apprehension, but also to more free-riding among participants in a brainstorming group. The other extreme of no anonymity has been shown to have the opposite effect. Due to the identified idea generation process, free-riding could be reduced but evaluation again would be increased. It appears that both extremes solve one problem while introducing another one.

3.2 A possible solution

Selective anonymity may be an approach that combines the best of both; we expect to negate the adverse effects of evaluation apprehension by only showing the top 10% of ideas with the creators' names. Also, by not keeping all the ideas completely anonymous and therefore keeping some accountability and incentive for participants to contribute more, we expect free-riding to occur less. Applying this reasoning, the following hypothesis can be derived: *Doing brainstorming with selective anonymity will result in more high-quality ideas than utilizing no anonymity or full anonymity.*

3.3 Mediator variables

Essential variables within the theoretical framework are free-riding and evaluation apprehension, meaning that the effect of the degree of anonymity on them as well as their effect on the number of high-quality ideas generated need to be tested. Our theory is based on the assumption that the degree of anonymity affects the number of high-quality ideas generated by decreasing free-riding and evaluation apprehension. This means that we expect these variables to act as mediators within our theoretical framework. The second hypothesis that can be derived due to this is: *The degree of anonymity influences the degree to which free-riding and evaluation apprehension occur. Selective anonymity will decrease the effect of evaluation apprehension in comparison to no anonymity and the effect of free-riding compared to full anonymity.*

The last relationship that needs to be observed in order to validify our theoretical framework is between the mediator variables and the number of high-quality ideas generated. Based on the results from previous studies we expect to observe a negative relationship between free-riding and evaluation apprehension, and the number of high-quality ideas generated. The third hypothesis that can be derived from this is: *The lower the occurrence of free-riding and evaluation apprehension is, the higher the number of high-quality ideas generated will be.*

The diagram below is supposed to illustrate the relationships between the variables and to which direct relationship the hypotheses relate:

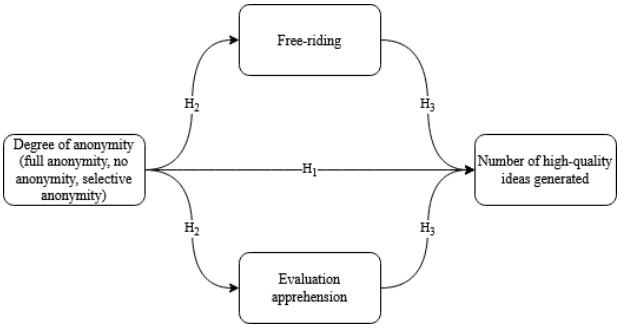


Figure 1 Theoretical framework

4. METHODOLOGY, DATA AND ANALYSIS

4.1 Research design

To answer the research question on the effect of different degrees of anonymity on the number of high-quality ideas generated, quantitative research has been conducted. The data used has already been gathered before this study but has not yet been evaluated in this way. One hundred six participants took part in a randomized control trial to generate business ideas for the student market. More specifically, the exact task given to the participants was: “You have been retained by a manufacturer of sports and fitness products to identify new product concepts for the student market. The manufacturer is interested in any product that might be sold to students in a sporting goods retailer. The manufacturer is particularly interested in products likely to be

appealing to students. These products might be solutions to unmet needs or improved solutions to existing needs. Please come up with ideas for new product concepts in the field of sports and fitness products for the student market." This research design has been chosen because the goal is to study a causal relationship between two variables, and by conducting an experiment, it is possible to deeply investigate the relationship between a dependent and an independent variable. In order to manipulate the independent variable, it was made sure that the participants were randomly distributed into one of the three groups with varying degrees of anonymity.

4.2 Data collection

The participants were students, and the trial was conducted via an online survey. Due to this, the session was facilitated electronically. Participants were allowed to write down up to ten ideas fitting the description. Afterwards they were asked to answer multiple questions (usually on a scale of 1-7) to measure their level of evaluation apprehension, free-riding, and other variables measured for the purpose of other studies. One example of one of those questions is: "I really took this task seriously" (1=" I fully disagree"; 7=" I fully agree").

4.21 Subjects

Participants of this study were students from the University of Twente. In total 106 individuals participated. The average educational background, age, sex and dropout rate of the participants is unknown at this point and will be further elaborated in the final version of the thesis. The students participated voluntarily, and their participation had no effect on their academic performance. They were asked to consent to the data collection and analysis practices and the study was reviewed and approved by the BMS ethics committee of the University of Twente.

4.22 Treatment

There are three different groups that were investigated over the course of this study varying in the degree of anonymity that was assigned to them. The degree of selective anonymity can be seen as the treatment and the groups of no anonymity and full anonymity as the control groups. For each group, the ideas were evaluated and ranked anonymously. The differences between the groups are what happens after the evaluation. For the first treatment of *no anonymity*, the participants were told that the description, overall rank, evaluation score, and name of the creator will be published for each idea. Participants in the second group - *full anonymity*- were told that none of this information would get published. The last group – *selective anonymity*- was told that this information will get published only for ideas that rank among the top 10% of all ideas.

4.23 Dependent variables

The dependent variable we will be investigating throughout this study varies based on the hypothesis that is being tested. For hypotheses H_1 and H_3 the number of high-quality ideas generated is the dependent variable. Evaluation apprehension and free-riding are the dependent variables for H_2. Since brainstorming aims to create great ideas to solve a specific problem, we can identify the number of great ideas generated as the effectiveness of a brainstorming group. The way that rating the quality of ideas was done was in alignment with previous research on evaluating creativity via a consensual assessment technique (Amabile, 1983). Requirements for the judges, according to this study, include experience with the domain in question, independent assessment, assessment on other dimensions than creativity alone, rating of ideas in comparison to each other rather than in comparison to a standard that the judge might have and lastly for all judges to view the products in different orders. For this study, the judges were seven different undergraduate students in the field of international business administration. The judges rated the ideas based on three factors: novelty (how novel is the idea), user value (how valuable is the idea for potential users), and purchase intent (how likely is the judge to buy the product) on a scale of 1 to 7.

4.3 Data analysis

In order to identify the number of great ideas generated per person, it was first necessary to recode the scores given by the judges into one overall score to rank the ideas. The first step of this process was to exclude all ideas from the data set where ratings from judges were missing. Sometimes ideas were not fitting for the brainstorming task assigned to the participants, so judges could not assess the score for all dimensions properly.

After cleaning up the data set, there were 101 participants left that had submitted a valid idea, and the subsequent decision that had to be made was how to recode the scores from the raters into one overall score. The most intuitive way would be to calculate the mean score given for each idea. However, after carefully assessing the data set, it became clear that this was possibly not the optimal solution. This can be best illustrated with an example: Many business ideas were simple products that already exist, such as "towel". The average scores given by the judges for novelty, user value and purchase intent were: 1, ~5,29 and ~4,71 resulting in an average score of ~3,67 for the idea "towel". In comparison: The idea of "Hydrating resistance vests" received average scores of ~4,29 novelty, ~3,43 user value and ~2,86 purchase intend, resulting in an average of ~3,52. Although the average of the first-mentioned idea is higher, an argument can be made that the second idea is overall more viable given the context. Since the novelty of the idea is stressed explicitly in the task given to the participants, a high score in the other two categories should not be able to make up for this deficit.

4.31 Ranking ideas

A more reliable way of assessing the overall quality of an idea is by taking the lowest average score across the three criteria as the overall score of the idea. By doing this, an idea can be seen as a chain, meaning that an idea is only as strong as its weakest "link" or criterion. This method ensures that only ideas that are sound in all three dimensions can make it into the category of a *high-quality* idea. After ranking the ideas, it still had to be decided which ideas qualify as a high-quality idea. In order to control for biases within the data set, the statistical data analysis was done for three different cutoff points. For the first analysis, the top 5% of ideas were considered a high-quality idea. The analysis was then done again with 10% and 15% as the cutoff points. Depending on the cutoff point, a total number of great ideas could be assigned to each participant. Figure 2 is supposed to provide some clarity on what exactly had to be done to the original data set to get to the data set that the different hypothesis could be tested with.

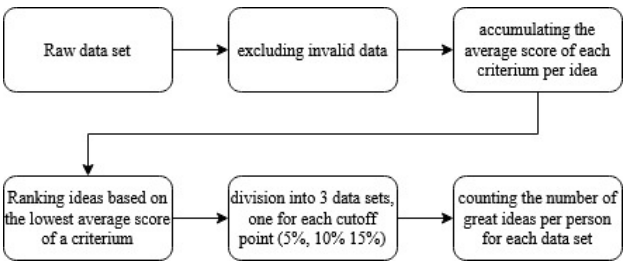


Figure 2 Data analysis

4.32 Testing H_1

The results for the average amount of great ideas generated per person were compared to the degree of anonymity which is the independent variable for this research design. A non-parametric test was used to analyze the data. In this case, the Kruskal-Wallis H test has been chosen. It was necessary to choose a non-parametric test to analyze this set of data because none of the parametric tests were viable due to their assumptions not being met. Parametric tests require specific distributions of the data (e.g. normal distribution) to be applicable. The distribution of the dependent variable *number of high-quality ideas* was tested and we found that the distribution was not specifiable. Non-parametric tests are not based on the assumption that the data follows a certain distribution, which makes a non-parametric test the logical choice. In order to validate this choice, first, the assumptions for this test must be checked. The detailed results for this are in appendix A.

4.33 Testing H_2

In addition to the previously mentioned analysis, which focuses on the direct relationship between the degree of anonymity and the number of great ideas generated per person, a mediation analysis was conducted to investigate the effects of the variables evaluation apprehension and free-riding. The first statistical analysis needed for this was the ANOVA test. It was necessary to conduct the test twice – first with evaluation apprehension as the dependent variable and second with free-riding as the dependent variable. In both cases, *scenario* was the independent variable. The term scenario was used in this research to describe the group of anonymity a participant was put into. The detailed results for the assumption testing are included in appendix B.

4.34 Testing H_3

The last relationship investigated during this study was between the independent variables evaluation apprehension and free-riding and the dependent variable of the number of high-quality ideas generated. To investigate this relationship, an ordinal regression analysis was conducted. Detailed results for the assumption tests are included in appendix C. Due to the assumption of proportional odds being violated for the 15% cutoff point, this analysis is limited to the 5% and 10% cutoff point.

5. RESULTS

Table 1

The direct effect of the scenario on the number of high-quality ideas generated

cutoff point	Scenario	*N*	*mean*	*SD*	*mean rank*	*Kruskal-Wallis H*	*P*
5%	anonymous	34	0,15	0,36	48,13	2,12	0,347
	non anonymous	26	0,23	0,59	49,19		
	selective anonymous	41	0,34	0,66	54,52		
10%	anonymous	34	0,38	0,70	49,5	2,072	0,355
	non anonymous	26	0,31	0,62	46,73		
	selective anonymous	41	0,61	1,02	54,95		
15%	anonymous	34	0,59	0,86	51,49	2,787	0,248
	non anonymous	26	0,38	0,70	44,17		
	selective anonymous	41	0,78	1,13	54,93		

Table 2

Means, standard deviations and one-way analysis of variance of the mediator variables depending on the scenario

Measure	Scenario	*N*	*Mean*	*SD*	*F*	*P*
evaluation apprehension	anonymous	34	2,68	1,27	1,919	0,152
	non anonymous	26	3,25	1,25		
	selective anonymous	41	3,14	1,21		
free-riding	anonymous	34	3,47	1,15	1,273	0,285
	non anonymous	26	3,16	0,93		
	selective anonymous	41	3,56	0,94		

Table 3

The effect of evaluation apprehension and free-riding on the number of high-quality ideas generated

cutoff point	Parameter estimates						Model Fitting Information		
	Effect	Estimate	*SE*	*P*	*90% CI*				
					LL	*UL*	Model	*Chi-Square*	*P*
5%	evaluation apprehension	-0,1	0,212	0,638	-0,448	0,249	Intercept Only		
	free-riding	-0,499	0,274	0,068	-0,949	-0,049	Final	4,277	0,118
	Effect	Estimate	*SE*	*P*	*90% CI*				
					LL	*UL*	Model	*Chi-Square*	*P*
10%	evaluation apprehension	-0,084	0,174	0,631	-0,37	0,203	Intercept Only		
	free-riding	-0,247	0,217	0,255	-0,604	0,11	Final	1,678	0,432

Note CI = Confidence interval; LL = Lower limit; UL = Upper limit

5.1 Direct effects of scenario on the number of high-quality ideas generated per person

The main results of the Kruskal-Wallis test are listed in table 1. When analyzing the descriptive output, the first thing to mention is the mean score of great ideas per person for each scenario. For every cutoff point, the scenario selective anonymous scores the highest in mean and standard deviation. However, the mean rank shows the opposite, with the selective anonymous scenario scoring the lowest for every cutoff point. The null hypothesis that is being tested with a Kruskal-Wallis test is:

h0: $Mdn_a=Mdn_{na}=Mdn_{sa}$

"The three samples that have been drawn come from populations with identical medians."

We do not have enough evidence to reject the null hypothesis based on the Kruskal-Wallis H and the resulting P-values (0,347, 0,355, 0,248 with $\alpha=0,1$). This result stays the same for each cutoff point. That means that we do not have enough statistical evidence to prove a significant difference in the number of great ideas per person in a brainstorming session among the three groups.

5.2 Effects of scenario on the mediator variables

Table 2 lists the results of the ANOVA tests. The null hypothesis for these tests is:

h0: $\mu a=\mu na=\mu sa$

"There is no difference in means of evaluation apprehension / free-riding between the different groups of anonymity".

Based on the P-values of 0,152 and 0,285 (with $\alpha=0,1$), we do not have enough evidence to reject the null hypothesis for either dependent variable. This means that there is no statistically significant difference between the mean score for evaluation apprehension or free-riding based on the three different degrees of anonymity.

5.3 Effects of the mediator variables on the number of high-quality ideas generated per person

The most important results from the ordinal regression analysis are listed in table 3. As we can see by the P-values of 0,118 and 0,432, there is no significant improvement in the fit of the final model over the null model for each cutoff point (with $\alpha=0,1$). Both cutoff points have insignificant P-values (0,638 and 0,631 with $\alpha=0,1$) for the effect of evaluation apprehension on the number of high-quality ideas. This means that we do not have enough evidence to confirm an effect of evaluation apprehension on the number of high-quality ideas generated. The P-value for the effect of free-riding on the number of high-quality ideas generated is insignificant for the 10% cutoff point (0,255 with $\alpha=0,1$). However, the same P-value for the 5% cutoff point is significant (0,068 with $\alpha=0,1$). This means that the results for the relationship between free-riding and the number of high-quality ideas generated are inconclusive.

6. DISCUSSION

6.1 Interpretation

The above-presented results allow us to answer the research question of how varying degrees of anonymity affect the number of high-quality ideas generated during a brainstorming session. Regarding the Kruskal-Wallis test, we can conclude from this study that no direct relationship could be statistically proven between the degree of anonymity and the number of high-quality ideas generated per person. However, this does not mean that the degree of anonymity has no effect on the mediator variables evaluation apprehension and free-riding. With this test result alone, it could be possible that selective anonymity has a positive relationship with one of the variables and a negative relationship with the other one so that the effects may cancel each other out when looking at the direct effect on the number of high-quality ideas generated. Because of this, it was necessary to test for a relationship between the varying degrees of anonymity and the mediator variables evaluation apprehension and free-riding. The results of the ANOVA tests shown in table 2 suggest that there is no significant effect between the degree of anonymity and a person's score for evaluation apprehension or free-riding. This rules out the possibility of evaluation apprehension or free-riding acting as a mediator variable. Additionally, it has been found that there was no statistically significant relationship between evaluation apprehension and the number of high-quality ideas generated. For the variable free-riding, the results were inconclusive. Still, since the model fit was insignificant, it seems reasonable to assume that if free-riding is a predictor of the number of high-quality ideas generated, then it is only a minor predictor in this study.

All three hypotheses that have been concluded in the "Theory"-section concerning the effects of varying degrees of anonymity on the number of high-quality ideas produced and the variables evaluation apprehension and free-riding have not been proven correct. However, it is important to notice that these results are only applicable for this exact brainstorming method (electronic nominal brainstorming). Likely, the results for a similar study conducted for in-person brainwriting would be different. This is one possible explanation for the difference between the anticipated and actual results of the study.

6.2 Theoretical implications

6.21 Anonymity

Considering the inconclusive results that other researchers have found concerning the effects of anonymity on brainstorming, this study is naturally validating the results from some while invalidating the results from others. The fact that there was no direct relationship observable between the degree of anonymity and the number of high-quality ideas generated is in line with a study by Valacich et al., 1992 which also found no effect of anonymity on ideational performance.

6.22 Evaluation apprehension

The assumption of researchers, however that anonymity would lower evaluation apprehension was not confirmed in this study. A possible explanation for this can be found in the research design. It has already been suggested that electronic brainstorming lowers evaluation apprehension compared to verbal brainstorming. Although the group of no anonymity was told that all their ideas, including their name, would be available to other participants and researchers, this concept of no anonymity is likely still different from being in the same room as other people. The fear of being judged by others might be higher in person than online. Because of this, it is possible that the fact that the survey took place online already lowered the evaluation apprehension of all groups, so the differences between groups were lowered. This theory is further supported by the fact that evaluation apprehension was an insignificant predictor of the number of high-quality ideas generated in this study. Rosenthal et al. (2009) mention in their book (pp. 231) that, among others, the experimental setting and the need for approval play a role in affecting evaluation apprehension and responses in studies. Due to the very low interaction with others during the experiment, the low observed evaluation apprehension is in line with previous research.

6.23 Free-riding

The theoretical implications concerning the effects of free-riding are, given the inconclusive results of the study in this regard, limited. Although previous literature suggests that free-riding should lower the task effort of individuals (McLeod, 2011) and thus lower the effectiveness of a brainstorming session, the results of this study support this only partly. With only one out of two cutoff points showing a significant effect of free-riding on the number of great ideas, it can be assumed that the degree of free-riding is at least not as important as it was ought to be. This contradicts the findings of Shepherd et al. (1995) who found an increase in productivity of 63% when introducing social comparison to counter the effects of free-riding. The results of this study suggest that the increase in productivity that Shepherd et al. (1995) found was not entirely due to the improvements in terms of free-riding. This implies that social comparison can have other positive effects than just reducing the amount of free-riding.

6.3 Practical implications

Due to the mostly statistically insignificant findings in this study, the practical implications that arise are limited. When managers are planning an electronic nominal brainstorming session, they should be aware that anonymity is not a factor affecting the number of high-quality ideas produced during the session. Another key takeaway is related to the process of choosing the ideal form of brainstorming. If managers feel like the participants are particularly likely to be negatively influenced by evaluation apprehension, for example, due to their character traits, electronic nominal brainstorming can be seen as a good option since evaluation apprehension was observed to not influence the performance of participants significantly in this study. Concerning free-riding, the practical implications are not entirely clear. The results of this study suggest free-riding to be of more importance for the number of high-quality ideas generated per person. This means that managers should rather focus on lowering the possibility of free-riding rather than evaluation apprehension when faced with a choice between the two options. One way of doing this would be to specifically instruct a facilitator to pay attention to free-riders within the group rather than focusing on lowering the evaluation apprehension of participants.

6.4 Limitations

First, it needs to be acknowledged that the sample size for the experiment was relatively small and that a bigger sample size might have led to more conclusive results. Furthermore, the sample was drawn exclusively from students currently studying in the Netherlands, meaning that the results may not be generalizable for students studying in countries with majorly different cultural dimensions. It also means that one needs to be careful when generalizing the results from this study and taking practical advice for brainstorming sessions where vastly different people are involved (student populations tend to be limited in terms of age and educational background for example, but a company might not be).

The statistical test chosen to conduct the analysis is the Kruskal-Wallis test. This non-parametric test is weaker than parametric tests and was only selected because the assumptions for parametric tests were not met. It seems possible that with a more significant sample, the distribution of the number of high-quality ideas would come closer to a Poisson distribution. This would enable us to use a parametric test to analyze the data and draw a more conclusive result (Poisson regression).

The methodological choices of this study were constrained by the experiment that had already taken place before this. It is possible that electronic nominal brainstorming was not the best choice to investigate the relationship between the degree of anonymity and the effectiveness of brainstorming since this form of brainstorming involves very little human interaction. Different degrees of anonymity might have had a more significant impact on other brainstorming setups. Nonetheless, the results of this study are relevant for the type of brainstorming that was investigated and can be seen as the most valid results in this context until a study with bigger sample size is conducted. Additionally, it was essential to test the relationship between anonymity and electronic nominal brainstorming since it is a type of brainstorming that is being used by multiple organizations today, meaning that optimizing this type of brainstorming is desirable.

6.5 Recommendations

Given the current state of the research concerning the effects of anonymity on brainstorming, there are a lot of possibilities for future research. In a more general sense, it needs to be further investigated which type of brainstorming works the best under which circumstances (verbal brainstorming, brainwriting, electronic nominal brainstorming, …) since the common literature tends to contradict each other in this regard.

The relationship between the degree of anonymity and its effects on brainstorming leaves much room for further investigation. Results from this study should be tested by conducting a study with a bigger sample size for the reasons mentioned above. Furthermore, it would be interesting to test the effects of varying degrees of anonymity on other kinds of brainstorming. Past research suggests that electronic nominal brainstorming might be less affected by the varying degrees of anonymity than other kinds of brainstorming. If this is true, then finding significant differences between groups should be easier when testing, for example brainwriting. Another factor that should be investigated further is the impact of facilitators when varying degrees of anonymity play a role. Previous research on the effects of facilitators was limited to identified interactive brainstorming. The effects of facilitators in anonymous or selective anonymous settings have not yet been investigated and represent a gap in the current knowledge on brainstorming.

7. CONCLUSION

The purpose of this study was to investigate the relationship between the degree of anonymity and the number of high-quality ideas generated during a brainstorming session. Based on the statistical analysis, it can be concluded that the degree of anonymity is not a significant factor. It must be mentioned that this is only true for the type of brainstorming conducted for this study's purpose, which is electronic nominal brainstorming.

Although the results found do not match the expectations from the theoretical framework, there are explanations for this deviance that are coherent with previous research. The theoretical framework was derived from common literature on the topic of brainstorming. Since many studies in this field are based on other kinds of brainstorming than investigated in this study, it seems plausible that a theoretical framework developed for the field of general brainstorming does not perfectly align with a very specific kind of brainstorming. These findings align with the previous theory on electronic brainstorming that suggested that electronic brainstorming led to less evaluation apprehension than conventional brainstorming methods. The lacking influence of the degree of anonymity on evaluation apprehension found in this study suggests that evaluation

apprehension might have been generally low, to begin with, due to the chosen form of brainstorming.

Because of this, the next logical step for researchers should be to investigate the same relationships as this study but for other forms of brainstorming. It seems reasonable to assume that forms of brainstorming where evaluation apprehension is a bigger factor might see significant differences between the various degrees of anonymity.

The most relevant takeaway for practical purposes would be to not think about anonymity as a factor when planning an electronic nominal brainstorming session.

8. REFERENCES

Amabile, T. M. (1983). A consensual technique for creativity assessment. The Social Psychology of Creativity, 37–63. https://doi.org/10.1007/978-1-4612-5533-8_3

Camacho, L. M., & Paulus, P. B. (1995). The role of social anxiousness in group brainstorming. Journal of Personality and Social Psychology, 68(6), 1071–1080. https://doi.org/10.1037/0022-3514.68.6.1071

Chidambaram, L., & Tung, L. L. (2005). Is out of sight, out of mind? an empirical study of social loafing in technology-supported groups. Information Systems Research, 16(2), 149–168. https://doi.org/10.1287/isre.1050.0051

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. Journal of Personality and Social Psychology, 53(3), 497–509. https://doi.org/10.1037/0022-3514.53.3.497

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Nunamaker, J. F. (1992). Electronic brainstorming and group size. Academy of Management Journal, 35(2), 350–369. https://doi.org/10.5465/256377

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best idea. Management Science, 56(4), 591–605. https://doi.org/10.1287/mnsc.1090.1144

McLeod, P. L. (2011). Effects of anonymity and social comparison of rewards on computer-mediated group brainstorming. Small Group Research, 42(4), 475–503. https://doi.org/10.1177/1046496410397381

Michinov, N. (2012). Is electronic brainstorming or brainwriting the best way to improve creative performance in groups? an overlooked comparison of two idea-generation techniques. Journal of Applied Social Psychology, 42. https://doi.org/10.1111/j.1559-1816.2012.01024.x

Mullen, B., Johnson, C., & Salas, E. (1991). Productivity loss in brainstorming groups: A meta-analytic integration. Basic and Applied Social Psychology, 12(1), 3–23. https://doi.org/10.1207/s15324834basp1201_1

Offner, A. K., Kramer, T. J., & Winter, J. P. (1996). The effects of facilitation, recording, and pauses on group brainstorming. Small Group Research, 27(2), 283–298. https://doi.org/10.1177/1046496496272005

Osborn, A. F. (1957). Applied imagination: Principles and procedures of creative problem-solving. C. Scribner's sons.

Oxley, N. L., Dzindolet, M. T., & Paulus, P. B. (1996). The effects of facilitators on the performance of brainstorming groups. Journal of Social Behavior and Personality, 11(4), 633–646.

Paulus, P. B., & Yang, H.-C. (2000). IDEA generation in groups: A basis for creativity in organizations. Organizational Behavior and Human Decision Processes, 82(1), 76–87. https://doi.org/10.1006/obhd.2000.2888

Pinsonneault, A., Barki, H., Gallupe, R. B., & Hoppen, N. (1999). Electronic brainstorming: The illusion of productivity. Information Systems Research, 10(2), 110–133. https://doi.org/10.1287/isre.10.2.110

Rosenthal, R., & Rosnow, R. L. (2009). Variables influencing the evaluation apprehension process. In Artifacts in behavioral research Robert Rosenthal and Ralph L. Rosnow's classic books: A re-issue of Artifact in behavioral research, experimenter effects in behavioral research and the volunteer subject. essay, Oxford University Press.

Shepherd, M. M., Briggs, R. O., Reinig, B. A., Yen, J., & Nunamaker, J. F. (1995). Invoking social comparison to improve electronic brainstorming: Beyond anonymity. Journal of Management Information Systems, 12(3), 155–170. https://doi.org/10.1080/07421222.1995.11518095

Sosik, J. J., Kahai, S. S., & Avolio, B. J. (1998). Transformational leadership and dimensions of creativity: Motivating idea generation in computer-mediated groups. Creativity Research Journal, 11(2), 111–121. https://doi.org/10.1207/s15326934crj1102_3

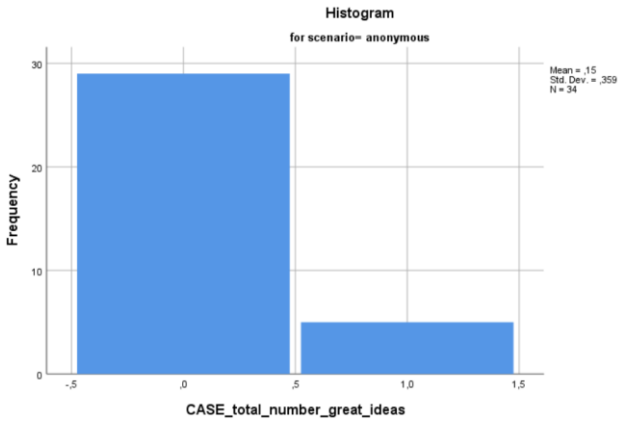
Valacich, J. S., Dennis, A. R., & Nunamaker, J. F. (1992). Group size and anonymity effects on computer-mediated idea generation. Small Group Research, 23(1), 49–73. https://doi.org/10.1177/1046496492231004

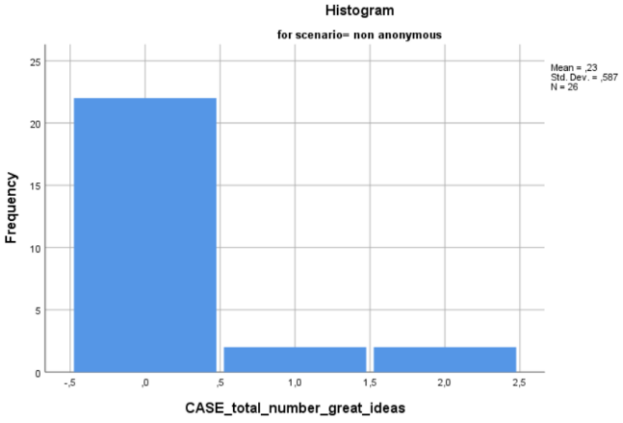
Zhou, X., Zhai, H.-K., Delidabieke, B., Zeng, H., Cui, Y.-X., & Cao, X. (2019). Exposure to ideas, evaluation apprehension, and incubation intervals in collaborative idea generation. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.01459

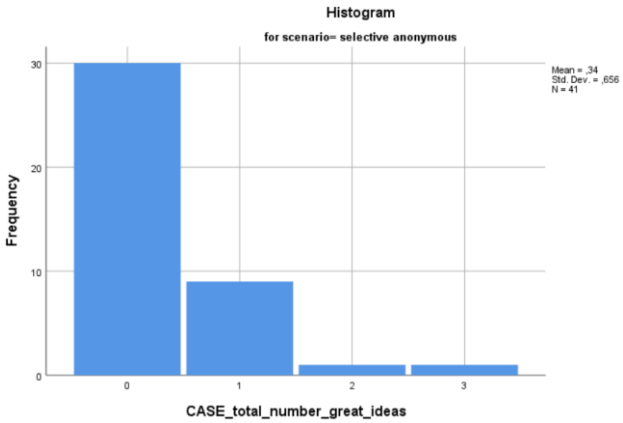
APPENDIX A

The first assumption is that the dependent variable needs to be measured at an ordinal or continuous level. Second, the independent variable must include 2 or more categorical independent groups. Third, the observations must have been made independently from each other. All of these three assumptions can be seen as fulfilled for this research design without having to check the assumptions with a statistical program. The fourth assumption, which states that the distribution for the different groups need to have the same shape, can be checked via visual investigation of the histograms generated from the data sets. The histograms show that the distribution has a similar shape for each group for each cutoff point.

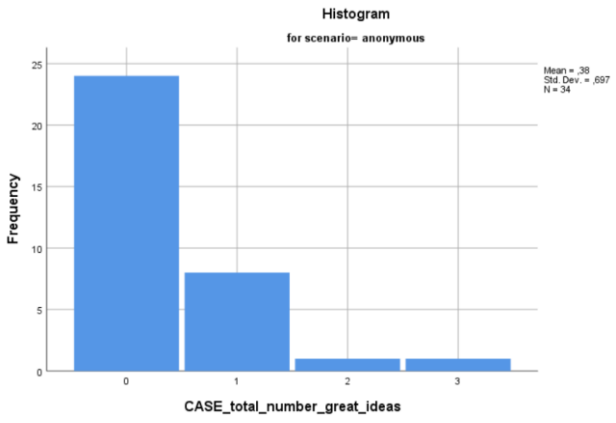
5% cutoff point

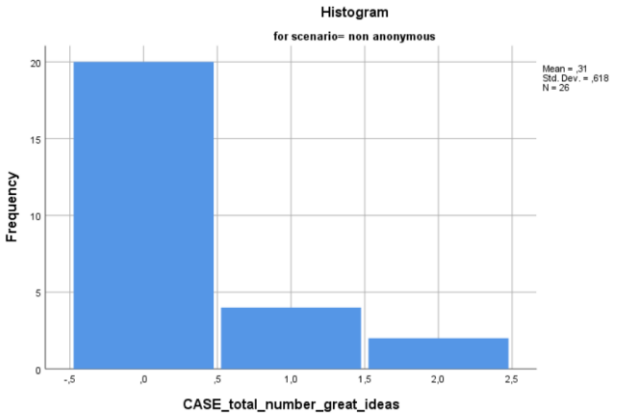


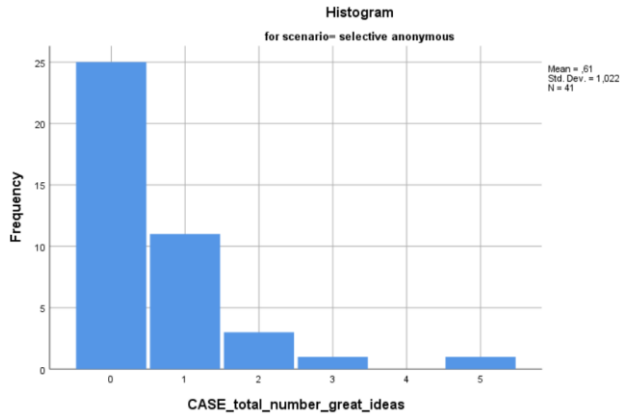




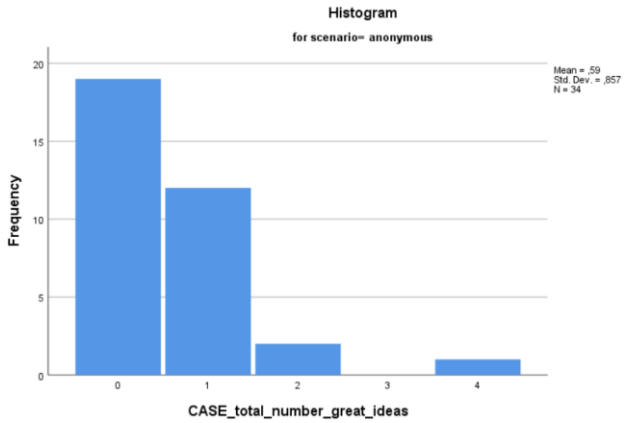
10% cutoff point

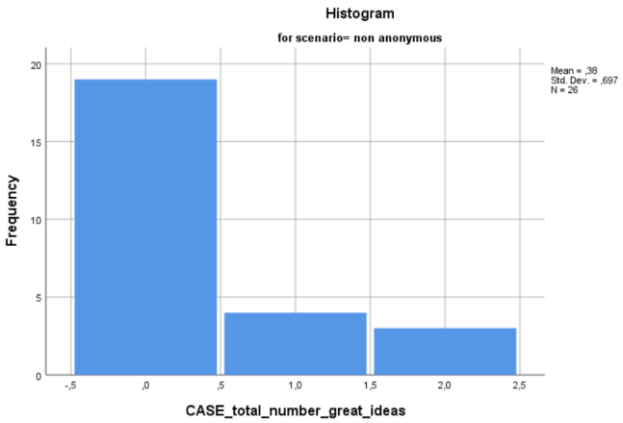


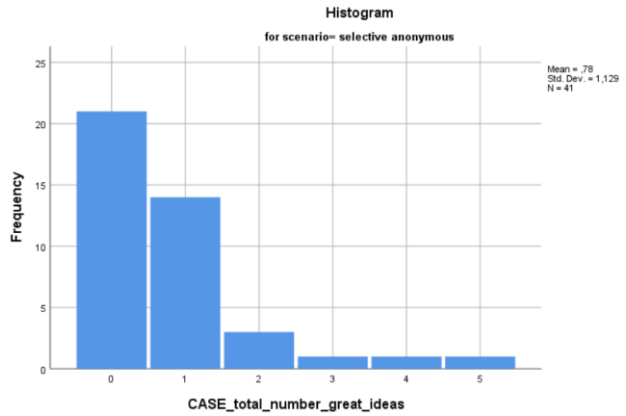




15% cutoff point







APPENDIX B

The first three assumptions for an ANOVA test are: The dependent variable should be continuous, the independent variable should include two or more independent groups, there should be independence of observations. These assumptions can be easily validated with the prior description of the research design. The other assumptions of no significant outliers, a normally distributed dependent variable for each group and the homogeneity of variances, have been tested via software.

Tests of Normality[a]

	scenario	Kolmogorov-Smirnov[b] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
evaluation_apprehension	anonymous	,114	34	,200*	,947	34	,102
inverse_freeriding	anonymous	,106	34	,200*	,964	34	,311

*. This is a lower bound of the true significance.

a. scenario = anonymous

b. Lilliefors Significance Correction

Tests of Normality[a]

	scenario	Kolmogorov-Smirnov[b] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
evaluation_apprehension	non anonymous	,117	26	,200*	,952	26	,254
inverse_freeriding	non anonymous	,147	26	,155	,961	26	,415

*. This is a lower bound of the true significance.

a. scenario = non anonymous

b. Lilliefors Significance Correction

Tests of Normality[a]

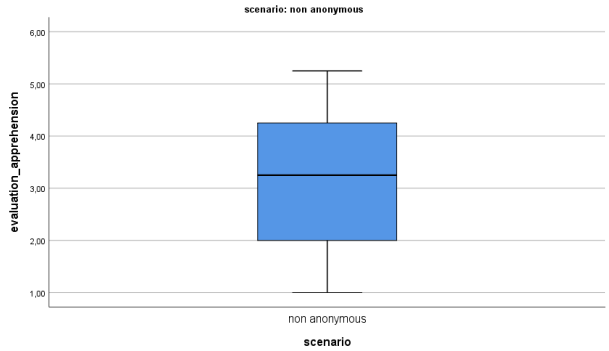
	scenario	Kolmogorov-Smirnov[b] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
evaluation_apprehension	selective anonymous	,119	41	,156	,968	41	,302
inverse_freeriding	selective anonymous	,110	41	,200*	,961	41	,171

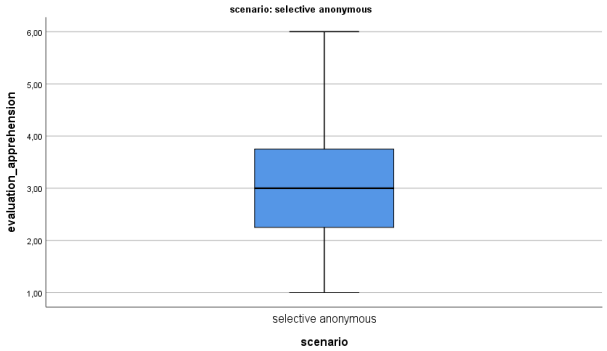
*. This is a lower bound of the true significance.

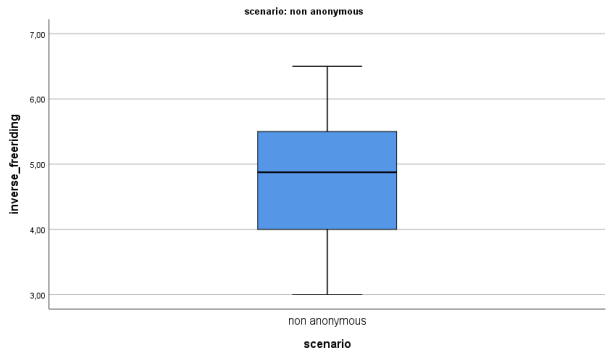
a. scenario = selective anonymous

b. Lilliefors Significance Correction

The Shapiro-Wilk test infers that there is a normal distribution for each individual group.







Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
evaluation_apprehension	Based on Mean	,139	2	98	,870
	Based on Median	,149	2	98	,862
	Based on Median and with adjusted df	,149	2	97,657	,862
	Based on trimmed mean	,144	2	98	,866
inverse_freeriding	Based on Mean	,141	2	98	,869
	Based on Median	,146	2	98	,864
	Based on Median and with adjusted df	,146	2	81,001	,864
	Based on trimmed mean	,155	2	98	,857

The boxplots show that there are no significant outliers for each group.

The Levene's test concludes that equal variances can be assumed.

Appendix C

The first assumption of the dependent variable being ordinal is met. The second assumption of the independent variables being continuous, ordinal or categorical is met. The detailed results for the assumption of proportional odds are listed in appendix C-1.05, C-1.10 and C-1.15.

C-1.05

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	113,020			
General	105,015[b]	8,005[c]	4	,091

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The log-likelihood value cannot be further increased after maximum number of step-halving.

c. The Chi-Square statistic is computed based on the log-likelihood value of the last iteration of the general model. Validity of the test is uncertain.

C-1.10

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	168,635			
General	166,398[b]	2,237[c]	6	,897

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The log-likelihood value cannot be further increased after maximum number of step-halving.

c. The Chi-Square statistic is computed based on the log-likelihood value of the last iteration of the general model. Validity of the test is uncertain.

C-1.15

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	198,286			
General	198,556[b]	.[c]	8	.

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The log-likelihood value cannot be further increased after maximum number of step-halving.

c. The log-likelihood value of the general model is smaller than that of the null model. This is because convergence cannot be attained or ascertained in estimating the general model. Therefore, the test of parallel lines cannot be performed.