The effect of anonymity in idea evaluation

Author

Roy Milder

Rotterdam School of Management,

Erasmus University

September 2019

Thesis advisors

Prof. Dr. Daan Stam,

Rotterdam School of Management,

Erasmus University

&

Dr. Mark Boons,

Vrije Universiteit Amsterdam

Abstract

We studied the effects of anonymity on the outcome of idea evaluations within the context of organizational ideation. Due to the rising popularity of topics like *open innovation* and *knowledge of the crowd* research on the subject of ideation in general and idea evaluation in specific is becoming more important. Response Bias is a known problem in Idea Evaluation. Evaluations are influenced by, for example, Social Desirability Effects and Rater Bias. Research shows that response bias can negatively affect the validity of results between 10% and 75%. Organizations that are using idea evaluations in there innovation or ideation strategy rely on the outcome of these evaluations for there future strategy. Therefor it is important for these organizations to be aware of the effect of response bias and to take steps for eliminating these biases. One of the known methods of eliminating response bias is anonymity and more specific, in the context of idea evaluation, the anonymity of both ideator and evaluator. In this study we have researched the influence of anonymity of both the ideator and evaluator on the Crowd Evaluation Score. We have conducted a field experiment in two organizations. We collected 863 idea evaluations which we used for analysis. Our study shows that there is an interaction effect between the anonymity of both ideator and evaluator. We have found that the Crowd Evaluation Score is considerably lower when an evaluation is completely anonymous.

Keywords: Ideation, Idea Evaluation, Idea management, Idea generation, Response Bias, Anonymity

Introduction

In 2018 we started with the development of our own ideation platform called 'Bamboo IMS'. Later we rebranded our platform to 'Hét Ideeënbureau'. Our goal was to develop a simple, effective and fun ideation platform that was build around scientifically funded processes and strategies. With our platform we help organizations develop an effective ideation program to support their long term innovation strategy.

In this research we take a close look on the impact of anonymity in the evaluation phase of ideation. One of the factors that one needs to take into consideration when using the results of these (idea) evaluations is *Response bias*. Response bias is the tendency of a person to answer questions on an evaluation untruthfully or misleadingly. Social Desirability Effects or Rater bias are two examples of response bias. One of the methods one can use to eliminate the effects of response bias is *anonymity*. Although different research has been conducted on the effects of anonymity of the ideator, there hasn't been much research on the effects of anonymity of the evaluator and the interaction between anonymity of both ideator and evaluator. Therefore, our study focuses on the question:

Does anonymity effect the evaluation of ideas in (organizational) crowd sourced ideation?

Following the principles of crowd-based idea collection, organizations more and more turn to crowdsourcing for review, selection, gaining internal and external support and evaluation of generated ideas. However, using the crowd for idea evaluation and selection has some disadvantages like response bias towards the ideator (Reitzig & Sorenson, 2013). Internal crowd evaluation could for example suffer from upward appraisals (Antonioni, 1994; Nederhof, 1985) and previous social interactions between ideator and evaluator. In general, studies show that respondents could show different response behavior when asked to response anonymous (Kerin & Peterson, 1976).

Using the crowd as source for idea generation has shown considerable success in recent history. Various organizations such as Dell (Bayus, 2012; Gangi di & Wasko, 2009), Starbucks (Hossain & Islam, 2015) and Proctor & Gamble (Dodgson, Gann, & Salter, 2006). Crowdsourced ideas tend to be more novel and beneficial than expert generated ideas (Poetz & Schreier, 2012). Crowdsourcing enables organizations to reach out to stakeholders, both internally and externally and involve them in the idea collecting and innovation process (Ågerfalk & Fitzgerald, 2008; Howe, 2006; Miner, 2005; Pisano & Verganti, 2008). Whereby most ideation platforms tend to aim at external users, harnessing employee ideas as input for idea generation has proven to be efficient and essential for organizations as well (Banbury & Mitchell, 1995). These ideas contribute to both radical innovations as well as more incremental process and practice innovations (Banbury & Mitchell, 1995).

Idea management (or *ideation)* as a method to gain access to new ideas has gained popularity among companies and non-profit organizations (Morgan & Wang, 2010; Piller & Walcher, 2006). As part of the concept of *Open Innovation* it is considered as 'the way to go' for organizations to survive in the current fast changing world (Gassmann, Enkel, & Chesbrough, 2010). Organizations need to innovate to stay ahead of competition. And every innovation has its origin in an idea (Rogers, 1995). Hence, access to a significant number of ideas is essential for innovation. Without ideas an organization is unable to innovate and without innovation it loses its raison d'être.

As ideation is becoming more important as part of open innovation, research on this topic is becoming more relevant. Recent studies are focusing on idea management in general or on idea generation specifically. Idea evaluation on its own can benefit from more research. It is commonly known that the results of surveys, questionnaires or evaluations are affected by response bias. Organizations need to be aware of this when using the outcomes for their future strategy. Organizations that are able to interpret the effects of response bias could gain (strategic) advantages. The studies that research idea evaluation tend to focus on the idea / ideator characteristics only and do not take the background of evaluators from the crowd in consideration. However to understand the effects of response bias on the outcome of evaluations, one needs to know how the identity of both ideator and evaluator affect the outcome of these evaluations. Therefor our study contributes to the academic development of idea evaluation in organizational ideation. We aim to contribute to the understanding of the effects of anonymity in crowd sourced idea evaluation and more specific organizational ideation. We're specifically interested in how the identify of ideator and evaluator affect the evaluation and if there is any interaction between them.

Our results give new insights for organizations that implement an ideation strategy which involves (internal) crowd sourcing. Our findings are also used in the development of our own ideation platform. For our research we have studied 2 health care organizations in the Netherlands. Approximately 4.857 employees have been asked to submit their ideas in an online idea management system. This resulted in 181 crowd generated ideas. Of these 181 ideas, 169 ideas have been evaluated. We have collected 863 idea evaluations by 425 unique evaluators.

Theoretical underpinnings

Idea management

As mentioned in our introduction, organizations need ideas as input for successful innovations (Rogers, 1995). An idea is commonly understood to be a concept or plan formed by mental effort (Newell, A; Shaw, J.C.; Simon, 1962). For organizations idea generation is a vital part of the long-term innovation strategy. Idea management is therefore seen as a vital (sub)process within the subject of innovation

strategy (Vandenbosch, Saatcioglu, & Fay, 2006). Idea management is about the process of continuously collecting of, collaborating on and evaluating of ideas. Over the years idea management has gained some sort of research attention, however the subject should be researched more as an independent area (Vandenbosch et al., 2006).

The concept of idea management can be defined as the process of recognizing the need for ideas, generating ideas and evaluating them. The first step, *recognizing*, affects the willingness of management to acknowledge the intertwining between on one hand (business) problems and ideas, *one needs an invention to solve a problem and problems arise from disaffection with inventions*, and on the other hand (business) opportunities and ideas, *one needs an idea to capitalize on an opportunity and opportunities are an occasion to test an idea* (Arrow, 1974). The second step, *generating*, is the where ideas are collected from (a group of) people. We will discuss this topic further on. The last step, *evaluating*, involves reviewing and selecting the best out of the generated ideas.

Open Innovation

Open Innovation is defined as 'the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and to expand the markets for external use of innovation respectively.' (Chesbrough, 2006). It has been proposed as a new paradigm for the management of innovation within organizations. (Gassmann, 2006; Hemphill, 2005). Open Innovation found its origin in the shortcomings of the traditional R&D model (van de Vrande, de Jong, Vanhaverbeke, & de Rochemont, 2009) which was based on the principle of discovering, developing and commercializing technology internally and was labeled as the closed innovation model (Hemphill, 2005). Closed Innovation has been the strategy that enabled large firms to outperform smaller rivals for years (Teece, 1986). Due to the rise of global developments such as labor mobility, easy access and sharing of knowledge, abundant access to capital and rapidly developing technology, organization are no longer able to innovate on their own and gain competitive advantage but they need to engage in alternative innovation practices (van de Vrande et al., 2009). When an organization engages in an open innovation strategy it uses both internal as external resources and it successfully challenges the Not Invented Here syndrome.

The Era of Open Innovation has just begun (Gassmann, Enkel, & Chesbrough, 2010) and as such different trends and patterns can be identified in the development of Open Innovation. One of these trends, as an opposite to the dominant stage-gate process of the 80's and 90's (Cooper, 1994), is a more probe-and-learn process (Gassmann et al., 2010; Lynn, Morone, & Paulson, 1996). As such, organizations are actively interacting in an early stage with customers, suppliers, R&D partners and their own employees. And although internal innovation processes are more and more professionalized

(Gassmann et al., 2010), there is still a shortcoming with the metric systems to successfully monitor and evaluate activities with their origin from open innovation.

Idea generation

Harnessing employee ideas as input for an idea management program has proven to be efficient and essential for organizations (Banbury & Mitchell, 1995). These ideas contribute to both radical innovations as well as more incremental process and practice innovations (Banbury & Mitchell, 1995). The role of organizational processes in idea generation has been examined in both social psychology literature as well as in innovation management literature (Girotra, Terwiesch, & Ulrich, 2008). Although views are changing whether the most effective way of generating ideas is by a team or hybrid structure (Girotra et al., 2008) in all circumstances idea generation involves some sort of employee involvement and can be seen as a group process. Open Innovation and Crowdsourcing fall within the same paradigm (Albors, Ramos, & Hervas, 2008): both distributed knowledge and both can provide an competitive advantage for companies when used to open a firm's R&D process. Following the rising popularity of open innovation, organizations reach out to external users as part of their idea generation strategy as well and are therefore more and more are using crowdsourcing as a vital source for collecting ideas.

Crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call (Howe, 2006). Crowdsourced ideas tend to be more novel and beneficial than expert generated ideas (Poetz & Schreier, 2012). Crowdsourcing enables organizations to reach out to stakeholders, both internally and externally and involve them in the idea collecting and innovation process (Ågerfalk & Fitzgerald, 2008; Howe, 2006; Miner, 2005; Pisano & Verganti, 2008).

Idea evaluation

In imitation of the success of crowdsourced idea generation and as part of the process of Idea management, organizations reach out to the crowd for evaluating the best idea out of all generated ideas. The concept of 'The wisdom of the crowds' has been described by James Surowiecki in his book from 2004 (Surowiecki, 2004) and although his book was critically reviewed, several studies validate the success of crowd sourcing and knowledge of the crowd in idea generation and innovation in general (Kijkuit & Van Den Ende, 2007; Zhu, Kock, Wentker, & Leker, 2018). The involvement of the crowd in idea evaluation leads to ideas of a higher quality (Blohm, Riedl, Leimeister, & Krcmar, 2011). Traditionally, ideas for innovation are selected by expert panels. However, crowd evaluations has been proven to efficiently complement these expert panels when selecting new ideas (Ehrlich, Galbraith,

Ehrlich, & Denoble, 2006). Building on the Socio-technical system theory (Geels, 2004; Trist, E; Bamforth, 1954) user-behavior is one of the perspectives to be taken in consideration when evaluating ideas within an idea management system.

Response behavior

Response bias

Building on the social network theory, social interactions are defined by weak and strong ties. Therefor the relationship between ideator and evaluator influences the evaluation of an idea (Bonds et al., 2012; Dell et al., 2012). For example, the hierarchical distance between ideator and evaluator effects the outcome of an evaluation (Zaggl, Schöttl, Schweisfurth, & Raasch, 2018). Another study shows the effect of reviewer identity in online review systems (Forman, Ghose, & Wiesenfeld, 2008). The phenomenon of 'Response bias' has been extensively researched over the years (Furnham, 1986). Response bias is the generic term for a whole range of responses to interviews, surveys or questionnaires which bias the response. For example, socially desirable or faking-good response, or the opposite, faking bad response. Research shows that response bias negatively effects the validity of test results between 10 and 75% (Nederhof, 1985). One of the most important factors of response bias is 'Social Desirability'.

Social Desirability Effects

Social desirability reflects the tendency to say things which place the speaker in a favorable light. One of the most used instruments for measuring social desirability is the Marlowe-Crowne Social Desirability Scale (CROWNE & MARLOWE, 1964). This effect is especially interesting in our study. How do these social desirability effects influence the outcome of idea evaluations? One study indicates an relation between response and respondents anonymity (Becker, 1976). However, effects of anonymity on response behavior lack decent research.

Rater Bias

When asked to rate peer, superior or subordinates' ideas the effects of Rater Bias must be taken into consideration. Especially *Halo Effects*, the tendency to suggest that due to the existence of a certain quality other qualities also exists, and *Differential Dimensions,* where ratings are influenced by superior, peer or subordinate relationships (Holzbach, 1978) could have an effect on the outcome of idea evaluations. When ideation takes place in an organizational context, due to the existing relationships between employees, the effects of Rater Bias on the outcome can be significant.

Anonymity

Generally speaking, one could say that relations between employees / colleagues tend to be of normal or positive nature. To keep up with appearances we tend to believe that in general, evaluators do evaluate ideas more positive if their identity is exposed to the ideator. Instead of rating solely based on the nature of the idea, the final score is influenced by the social connections between ideator/ evaluator.

Within organizations there are relatively strong social networks between employees. It is plausible to say that there is a relatively large chance that the evaluator or ideator, not necessarily in person, knows the other. So specifically in organizational ideation, there is a greater risk of response bias due to existing social connections.

Anonymity increases the focus on the content of ideas presented versus who is presenting them (Connolly, Jessup, & Valacich, 1990) and could eliminate response bias in idea evaluations. When speaking of anonymity in idea evaluation, our research focused on the effects of anonymity of both ideator and evaluator and, more specifically, on the interaction effects between these two.

Previous research found that anonymity of the evaluator in peer review contexts leads to lower ratings then reviews where the anonymity of the evaluator was identifiable (Lu & Bol, n.d.). Similarly, Howard et al. (Howard, Barrett, & Frick, 2010) found that participants who anonymously provide feedback through online communities are approximately five times more likely to give critical feedback than students who were identifiable.

In our research we made some assumptions on which we based our hypothesis. The first assumption we made, as we mentioned at beginning of this paragraph, is that the existing relations between employees / colleagues are of a normal or positive nature. The second assumption made is that these relations are all more or like the same. Based on the previous research and assumptions the following can be concluded. One can assume that evaluators, when their identity is exposed, tend to give a higher Crowd Evaluation Score due to the Social Desirability Effect. One can also assume that due to the effect of rater bias and the assumption that relations between employees are of a positive nature, evaluations were the ideator' identity is exposed, tend to have a higher CES mean.

This leads to the following two hypothesis.

H1. In idea evaluation, if the evaluator, in advance, knows that his/her identity is exposed to the ideator, the crowd evaluation score will be higher than when his/her identity stays anonymous.

H.2. The effects of H.1. are abolished when evaluations are anonymous for both ideator and evaluator.

Methods, design and sample

This study was conducted at two Dutch organizations in the health care sector. Organization A has 3.881 employees and Organization B has 988 employees. Organizational hierarchy is relatively flat consisting of four levels being (from bottom to top) operational, middle management, upper management and board of directors. Both organizations are decentralized within the Netherlands consisting of many small (sub) units. Each unit has a high level of autonomy. Both organizations have an open culture meaning that the distance between operational and upper management / board of directors is small.

Ideation process

For our research we used a web-based idea management system (IMS). During a 3-month period we have used a systematic approach for generating ideas, collecting feedback and evaluating. In the first phase we actively approached employees by e-mail and asked them to submit ideas. Participation was on a voluntary base. There was no reward for suggesting ideas. The first phase took approximately four – five weeks.

Phase 1

	Employees	Ideas	Idea/Employee ratio
Organization A	3.881	113	0,029
Organization B	988	53	0,053
Total		166	

In the second phase we asked all employees to contribute to the development of the suggested ideas by reading them and giving feedback. The second phase took 4-5 weeks.

Phase 2

	Employees	Ideas	Comments	Comments / Employees ratio	Likes	Likes / Employees
Organization A	3.881	113	101	0,026	525	0,134
Organization B	988	53	77	0,078	77	0,078

Phase 3

In the third phase we collected all evaluations used in our research. All employees were asked to evaluate the generated ideas. The third phase took approximately 7 days. Each evaluation consisted of two questions. Ideas were rated on a Likert scale from 1 (totally disagree) to 5 (totally agree).

	1	2	3	4	5
This is a good idea					
My organization should further investigate this idea					

During the third phase we collected 863 evaluations as our sample.

	N	% of Total N
Organization A	584	67.7%
Organization B	279	32.3%
Total	863	100.0%

Before sending out the evaluations each evaluator was assigned an Evaluation Type which determined if the ideator and/or evaluator was (presented) anonymous. Each evaluator was randomly assigned to one evaluation type for all assigned evaluations thus assuring that the evaluators answers were not influenced by previous knowledge regarding anonymity.

Evaluation Type	N	% of Total N
Ideator anonymous / Evaluator anonymous	236	27.3%
Ideator anonymous / Evaluator known	214	24.8%
Ideator known / Evaluator anonymous	235	27.2%
Ideator known / Evaluator known	178	20.6%
Total	863	100.0%

Each evaluator received an invitation by e-mail to evaluate a single idea. Ideas were randomly assigned to an evaluator. After evaluating the first idea, evaluators had the opportunity to evaluate more ideas. This resulted in some evaluators evaluating more than 1 idea.

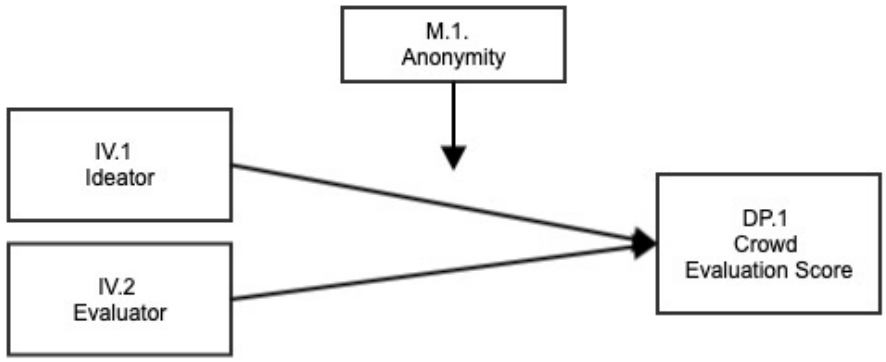
Table 1 Evaluations by same evaluator (N=evaluations)

	N	% of Total N
Less than 4 evaluations	487	56.4%
Between 4 and 8 evaluations	145	16.8%
9 or more evaluations	231	26.8%
Total	863	100.0%

To verify the correctness of the given answers by the evaluators we added a manipulation check at the end of each evaluation. We asked each evaluator to confirm the anonymity of both ideator and evaluator.

Manipulation check	N	% of Total N
Failed	149	17.2%
Passed	714	82,8%
Total	863	100.0%

Conceptual model



Variables

Dependent variable

Evaluation score (DV1): This score indicates the evaluation score for an idea.

Independent variables

Our independent variables are “ideator” (IV.1) being the person who submitted the idea and “evaluator” (IV.2), being the person who evaluated the idea.

Moderator

Anonymity (M.1) is our moderator.

Results

We differentiate our dataset into four (4) subsets. Dataset 1 (DATA1) being our raw dataset (N=863), dataset 2 (DATA2), being the dataset without the failed manipulation checks (N=799), dataset 3 (DATA3), being the dataset without the high occurrence evaluators (N=499). The last dataset (DATA4) consists of only cases with a passed manipulation check and less than 4 evaluations per evaluator. Details of each dataset are shown in Table 2. We have determined 3 different subsets besides the main set to get more insights in the effects of anonymity and to exclude important factors that could muddle the results.

Table 2. Overview of (sub) datasets

Dataset	Description	N	Mean	SD	
DATA1	Raw data	863	3,8239	1,02051	100%
DATA2	Evaluations with passed manipulation check	714	3,5674	1,03168	83%
DATA3	DATA2 excluding evaluators with more than 8 evaluations	499	3,6378	1,12982	58%
DATA4	DATA2 excluding evaluators with more than 3 evaluations	372	3,9849	0,95312	43%

Manipulation check

Each evaluator was asked to verify the given evaluation type after each evaluation. Was the evaluator aware that both the identity of the ideator and/or identity of his/herself was anonymous or not? 714 evaluations passed this check. However out of the 149 evaluation that did not pass the check, some did pass the check with another evaluation. Due to the fact each evaluator was assigned one evaluation type for all assigned evaluations, we can conclude that number of registered passed manipulation check is lower than the actual evaluations were the evaluator was correctly aware of the anonymity type. Therefor we could argue that the raw dataset is valid to use for our analysis.

Variables

We defined several variables in our dataset which we used for our analysis. The most important variables are the *Crowd Evaluation Score*, which is the mean score of both questions that were asked in each evaluation. *Ideator/Evaluator anonymity*, which indicates the anonymity type (known / anonymous) for both ideator and evaluator in an evaluation. *Evaluation Type,* which indicates the anonymity combination for an evaluation (

Table 4).

Table 3. Variables

Variable	Measurement
Crowd Evaluation Score	Mean score of each idea by crowd evaluations
Ideator anonymity	Ideator was anonymous
Evaluator anonymity	Evaluator was anonymous
Evaluation Type	Level of anonymity for an evaluation
Gender Ideator	If the ideator was male (dummy)
Gender Evaluator	If the ideator was male (dummy)
Manipulation check	If the manipulation check passed (dummy)
Evaluator occurrence	If the evaluator was responsible for 9 or more evaluations (dummy)

Table 4. Evaluation Type

	Ideator Known	Ideator Anonymous
Evaluator Known	Type = 4	Type = 2
Evaluator Anonymous	Type = 3	Type = 1

As we're interested in the effects of anonymity on the CES, our first step is to analyze the means of the different evaluation types. Means of all Evaluation Types in all four datasets are shown in Tables 5 - 8 and are visualized in Figure 1. Results show some interesting differences between the Evaluation Types. Evaluations of type 2 and 3 show the highest mean CES. Evaluations of type 4 show a slightly lower mean, however differences are minimal. Evaluation Type 1 (Anonymous/Anonymous) however, does show a considerably lower mean than the other three types.

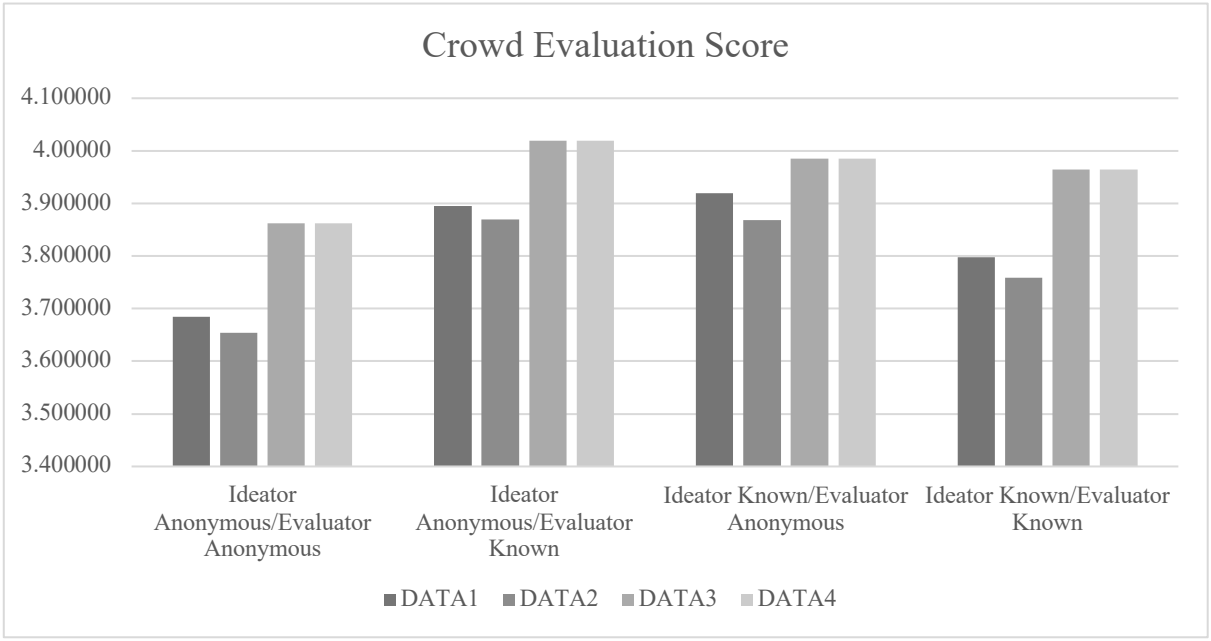


Figure 1 Crowd Evaluation Score by Evaluation Type

Table 5. Descriptives on Evaluation Type DATA1

Evaluation Type	Mean	N	SD
Ideator Anonymous/Evaluator Anonymous	3.6843	236	1.01002
Ideator Anonymous/Evaluator Known	3.8949	214	.96265
Ideator Known/Evaluator Anonymous	3.9191	235	1.03355
Ideator Known/Evaluator Known	3.7978	178	1.07019
Total	3.8239	863	1.02051

Table 6. Descriptives on Evaluation Type DATA2

Evaluation Type	Mean	N	SD
Ideator Anonymous/Evaluator Anonymous	3.6544	217	.99438
Ideator Anonymous/Evaluator Known	3.8696	138	.88850
Ideator Known/Evaluator Anonymous	3.8680	197	1.08696
Ideator Known/Evaluator Known	3.7593	162	1.08379
Total	3.7787	714	1.02462

Table 7. Descriptives on Evaluation Type DATA3

Evaluation Type	Mean	N	SD
Ideator Anonymous/Evaluator Anonymous	3.7281	171	1.025
Ideator Anonymous/Evaluator Known	3.9250	100	.92216
Ideator Known/Evaluator Anonymous	3.9356	132	1.10586
Ideator Known/Evaluator Known	3.9740	96	.91009
Total	3.8697	499	1.00903

Table 8. Descriptives on Evaluation Type DATA4

Evaluation Type	Mean	N	SD
Ideator Anonymous/Evaluator Anonymous	3.8625	120	.90508
Ideator Anonymous/Evaluator Known	4.0190	79	.88232
Ideator Known/Evaluator Anonymous	3.9854	103	1.0466
Ideator Known/Evaluator Known	3.9643	70	.9755
Total	3.9489	372	.95312

Results of Evaluation Type 2 and 3 do not support our first hypothesis (H.1.). As the means show, there is no significant difference in score between evaluations with a known or anonymous evaluator. Before we reject H.1. we performed a correlation test on all four datasets (Tables 9 – 12). Table 9 shows no correlation between CES and Evaluator Anonymity, suggesting that anonymity does not have any effect on CES. We see the same results on DATA2, 3 and 4 (Tables 10 – 12).

However, we did notice a correlation effect between Evaluation Type and CES (Sig.=0,048, Sig.=0.046, Sig.=0.093, Sig.=0.046), although 3 out of four are not at our minimum level (Sig.=0.050). Nonetheless, these results could indicate an interaction effect between IV.1 and IV.2 on DV.1 which could support our second hypothesis.

To test for the existence of any interaction between IV.1. and IV.2 on DV.1. we used an ANOVA test. We tested all four datasets. All sets were tested for the normality requirement (Figure 6, Figure 7, Figure 8 & Figure 9). We did not test for *homogeneity* using Levene's test: all four groups are considered large enough (N>50) to be equal. The smallest and largest group do not differ more than factor 4 (Table 17, Table 18, Table 19 & Table 20).

Results of the ANOVA test are shown in Table 13, Table 14, Table 15, Table 16 and visualized in Figure 2, Figure 3, Figure 4 & Figure 5.

DATA1, DATA2 and DATA4 show an interaction effect between IV.1 and IV.2 on DV.1. However, differences of CES on DATA4 are not considered enough to be of any significance. Nonetheless, results on DATA1 (Sig.=0.018) and DATA2 (Sig.=0.037) are significant and based on these results we can conclude that there is an interaction effect between IV.1. and IV.2. Therefor we support hypothesis 2.

Table 9. Descriptives and Correlations DATA1

		Mean	SD	Min	Max	1	2	3	4	5	6	7	8
1	Crowd evaluation Score	3,82	1,02051	1,00	5,00	1							
2	Ideator anonymity	0,52	0,500	0	1	-0,040	1						
3	Evaluator anonymity	0,55	0,498	0	1	-0,024	-0,045	1					
4	Gender Ideator	0,14	0,343	0	1	-0,001	-0,061	0,001	1				
5	Gender Evaluator	0,16	0,368	0	1	0,008	0,035	-0,005	0,001	1			
6	Evaluation Type	2,41	1,097	1	4	0,048	-0,891**	-0,413**	0,055	-0,029	1		
7	Manipulation Check	0,83	0,378	0	1	-0,097**	-0,106**	0,150**	-0,016	0,025	0,029	1	
8	Evaluator Occurrence	0,27	0,443	0	1	-0,129**	-0,123**	-0,069*	0,036	0,006	0,143**	0,165**	1

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

Table 10. Descriptives and Correlations DATA2

		Mean	SD	Min	Max	1	2	3	4	5	6	7
1	Crowd evaluation Score	3,78	1,02462	1,00	5,00	1						
2	Ideator Anonymity	0,50	0,500	0	1	-0,040	1					
3	Evaluator Anonymity	0,58	0,494	0	1	-0,026	0,063	1				
4	Gender Ideator	0,13	0,340	0	1	0,004	-0,035	-0,026	1			
5	Gender Evaluator	0,17	0,372	0	1	0,006	0,055	-0,034	-0,019	1		
6	Evaluation Type	2,43	1,144	1	4	0,046	-0,902**	-0,487**	0,042	-0,034	1	
7	Evaluator Occurrence	0,30	0,459	0	1	-0,135**	-0,140**	-0,085*	0,039	0,020	0,159**	1

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

Table 11. Descriptives and Correlations DATA3

		Mean	SD	Min	Max	1	2	3	4	5	6
1	Crowd evaluation Score	3,87	1,00903	1,00	5,00	1					
2	Ideator anonymity	0,54	0,499	0	1	-0,075	1				
3	Evaluator anonymity	0,61	0,489	0	1	-0,063	0,053	1			
4	Gender Ideator	0,12	0,330	0	1	0,049	-0,033	-0,033	1		
5	Gender Evaluator	0,16	0,367	0	1	0,016	0,105*	0,027	-0,016	1	
6	Evaluation Type	2,31	1,134	1	4	0,093*	-0,903**	-0,478**	0,043	-0,104*	1

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

Table 12. Descriptives and Correlations DATA4

		Mean	SD	Min	Max	1	2	3	4	5	6
1	Crowd evaluation Score	3,95	0,953	1,00	5,00	1					
2	Ideator anonymity	0,53	0,499	0	1	-0,027	1				
3	Evaluator anonymity	0,60	0,491	0	1	-0,038	0,008	1			
4	Gender Ideator	0,12	0,330	0	1	0,024	-0,026	-0,010	1		
5	Gender Evaluator	0,16	0,371	0	1	0,020	0,020	0,021	-0,012	1	
6	Evaluation Type	2,33	1,116	1	4	0,041	-0,898**	-0,446**	0,028	-0,027	1

**. Correlation is significant at the 0.01 level.

*. Correlation is significant at the 0.05 level.

Table 13. Between-Subjects Effects DATA1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7.929a	3	2.643	2.552	.054
Intercept	12456.166	1	12456.166	12025.012	.000
Anonymous Ideator	1.010	1	1.010	.975	.324
Anonymous Evaluator	.423	1	.423	.408	.523
Anonymous Ideator * Anonymous Evaluator	5.866	1	5.866	5.663	.018
Error	889.799	859	1.036		
Total	13516.500	863			
Corrected Total	897.728	862			
a R Squared = .009 (Adjusted R Squared = -.005)					

Table 14. Between-Subjects Effects DATA2

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6.126[a]	3	2.042	1.953	.120
Intercept	9936.071	1	9936.071	9502.309	.000
Anonymous Ideator	.462	1	.462	.442	.506
Anonymous Evaluator	.490	1	.490	.469	.494
Anonymous Ideator * Anonymous Evaluator	4.542	1	4.542	4.344	.037
Error	742.410	710	1.046		
Total	10943.500	714			
Corrected Total	748.536	713			
a R Squared = .004 (Adjusted R Squared = -.004)					

Table 15. Between-Subjects Effects DATA3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.353[a]	3	1.784	1.760	.154
Intercept	7.157.000	1	7.157.000	7.061.699	.000
Anonymous Ideator	1.944	1	1.944	1.918	.167
Anonymous Evaluator	1.636	1	1.636	1.614	.205
Anonymous Ideator * Anonymous Evaluator	.743	1	.743	.733	.392
Error	501.680	495	1.013		
Total	7.979.500	499			
Corrected Total	507.033	498			
a R Squared = .004 (Adjusted R Squared = -.004)					

Table 16. Between Subjects Effects DATA4

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1.438[a]	3	.479	.526	.665
Intercept	5.571.225	1	5.571.225	6.109.243	.000
Anonymous Ideator	.103	1	.103	.113	.736
Anonymous Evaluator	.407	1	.407	.446	.504
Anonymous Ideator * Anonymous Evaluator	.701	1	.701	.769	.381
Error	335.592	368	.912		
Total	6.138.000	372			
Corrected Total	337.030	371			
a R Squared = .004 (Adjusted R Squared = -.004)					

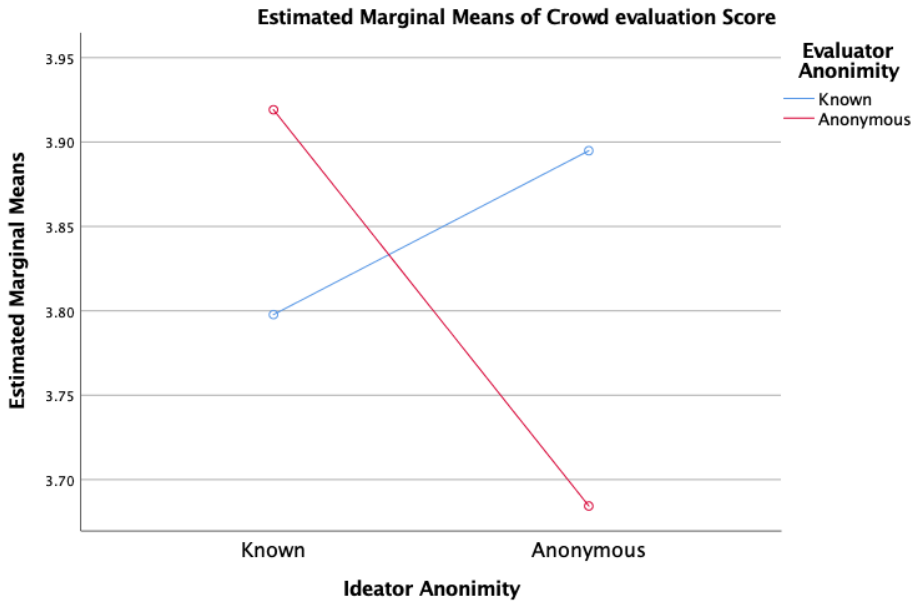


Figure 3 Interaction effect DATA1

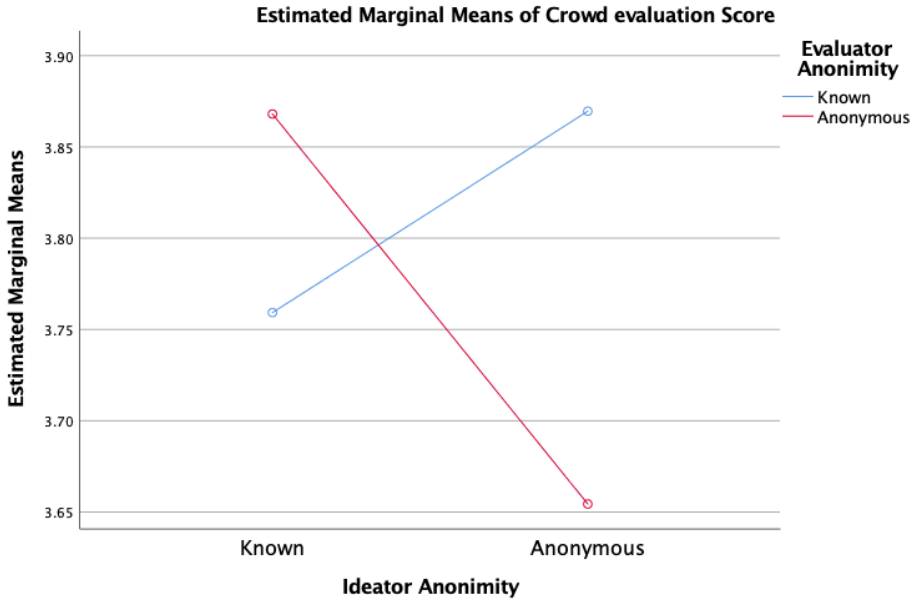


Figure 2 Interaction effect DATA2

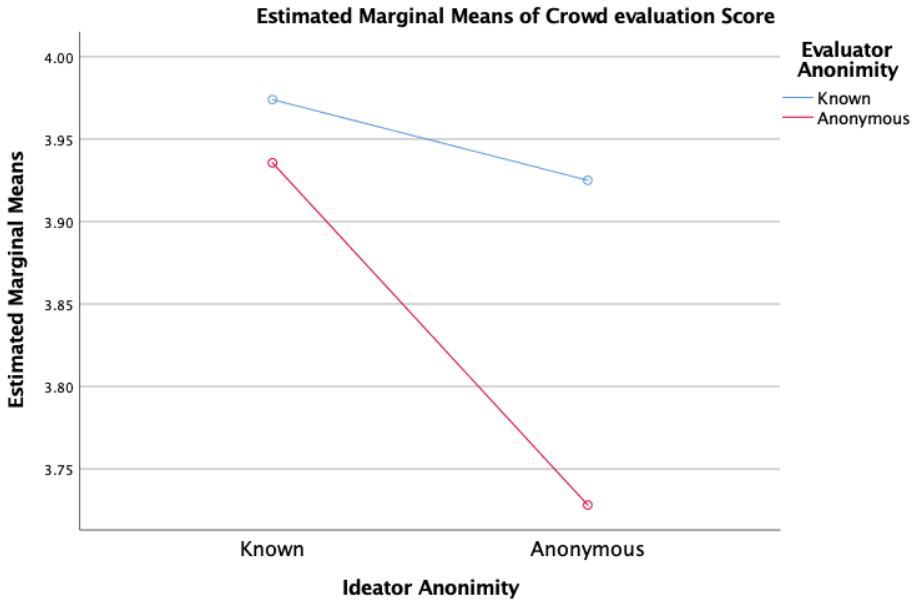


Figure 4 Interaction effect DATA3

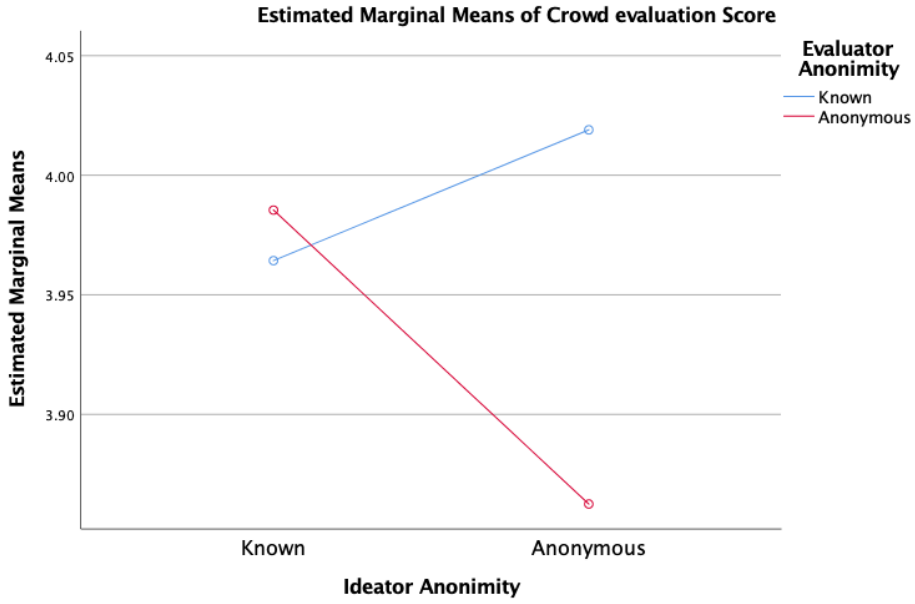


Figure 5 Interaction effect DATA4

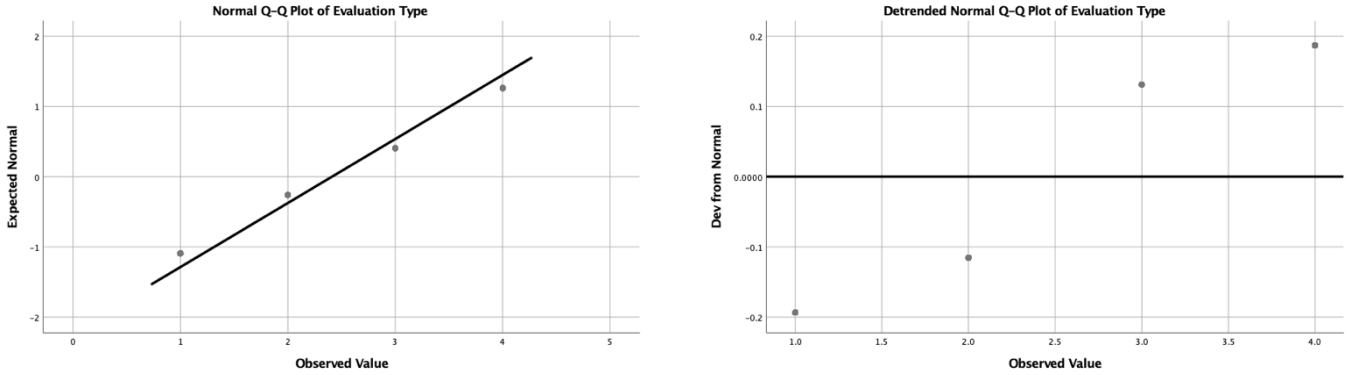


Figure 6 Normal Probability plot DATA1

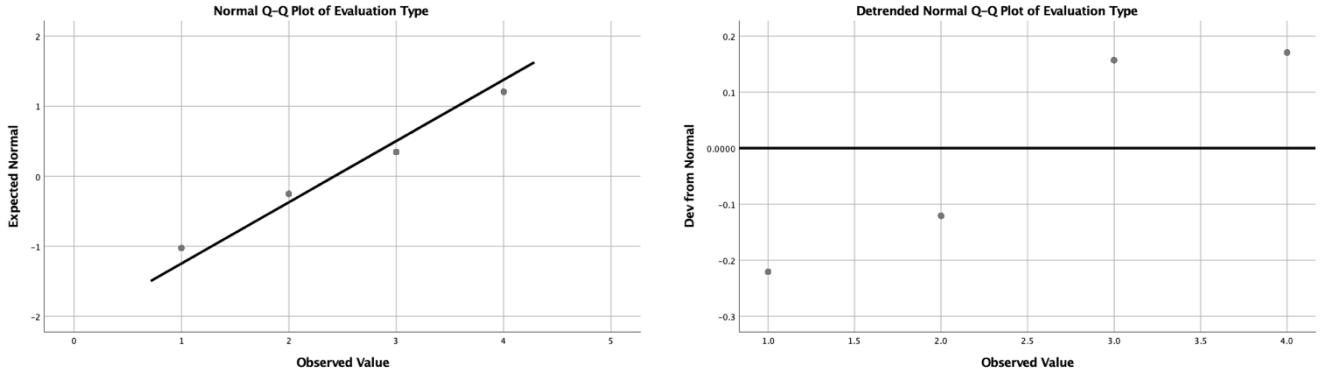


Figure 7 Normal Probability plot DATA2

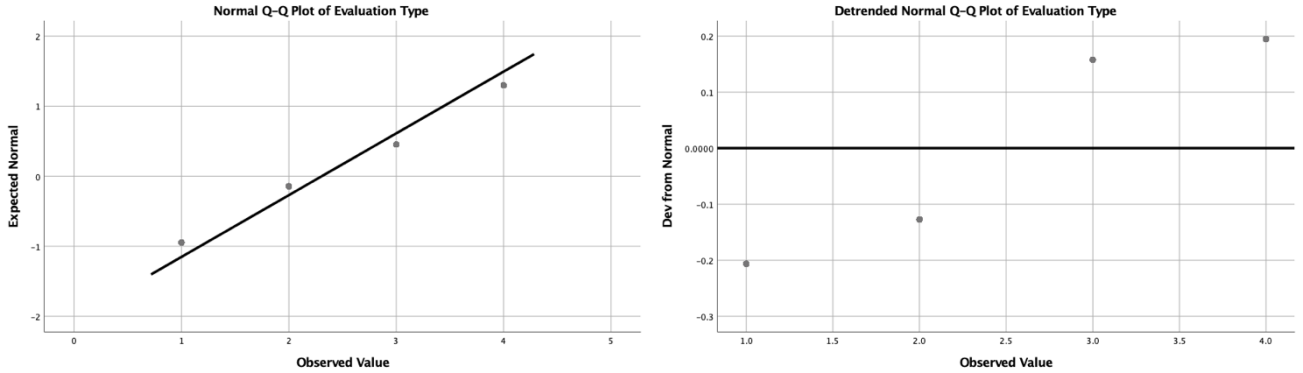


Figure 8 Normal Probability plot DATA3

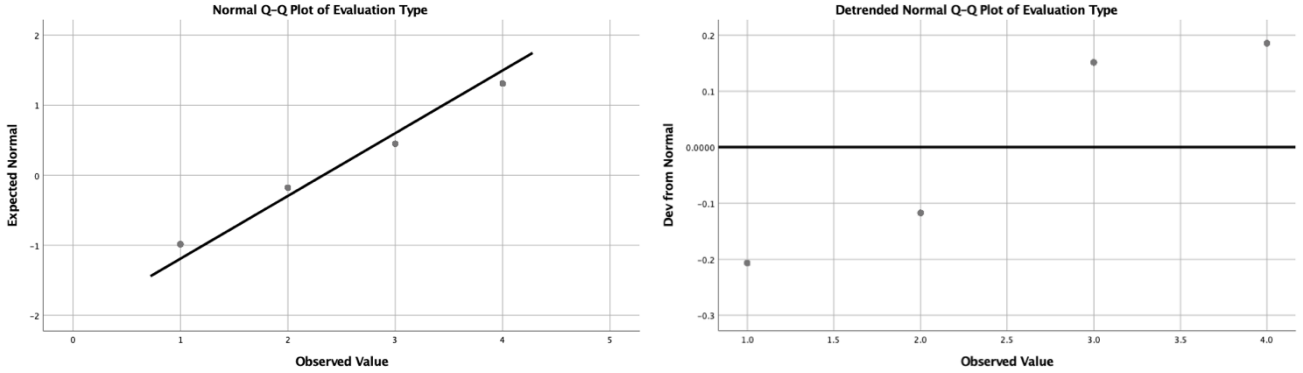


Figure 9 Normal Probability plot DATA4

Table 17. Descriptives and Variance Analysis DATA1

	Frequency	Percent	Valid Percent	Cumulative Percent
Ideator Anonymous/Evaluator Anonymous	236	27.3	27.3	27.3
Ideator Anonymous/Evaluator Known	214	24.8	24.8	52.1
Ideator Known/Evaluator Anonymous	235	27.2	27.2	79.4
Ideator Known/Evaluator Known	178	20.6	20.6	100.0
Total	863	100.0	100.0	

Table 18. Descriptives and Variance Analysis DATA2

	Frequency	Percent	Valid Percent	Cumulative Percent
Ideator Anonymous/Evaluator Anonymous	217	30.4	30.4	30.4
Ideator Anonymous/Evaluator Known	138	19.3	19.3	49.7
Ideator Known/Evaluator Anonymous	197	27.6	27.6	77.3
Ideator Known/Evaluator Known	162	22.7	22.7	100.0
Total	714	100.0	100.0	

Table 19. Descriptives and Variance Analysis DATA3

	Frequency	Percent	Valid Percent	Cumulative Percent
Ideator Anonymous/Evaluator Anonymous	171	34.3	34.3	34.3
Ideator Anonymous/Evaluator Known	100	20.0	20.0	54.3
Ideator Known/Evaluator Anonymous	132	26.5	26.5	80.8
Ideator Known/Evaluator Known	96	19.2	19.2	100.0
Total	499	100.0	100.0	

Table 20. Descriptives and Variance Analysis DATA4

	Frequency	Percent	Valid Percent	Cumulative Percent
Ideator Anonymous/Evaluator Anonymous	120	32.3	32.3	32.3
Ideator Anonymous/Evaluator Known	79	21.2	21.2	53.3
Ideator Known/Evaluator Anonymous	103	27.7	27.7	81.2
Ideator Known/Evaluator Known	70	18.8	18.8	100.0
Total	372	100.0	100.0	

Discussion

In this study we examined whether anonymity of ideator and/or evaluator is of any influence on the result of idea evaluations as part of an ideation program. Research shows that response bias is an import influencer on the outcome of surveys and evaluations. Therefore, it is important to know how response bias influences the outcome of idea evaluations in ideation. We have researched the effects of anonymity in four different evaluation settings.

We found that evaluation types 2, 3 and 4 result in the same mean Crowd Evaluation Score. Contrary to H.1., we do not see any difference in the outcome of an evaluation when only ideator or evaluator is anonymous or when both are identifiable. However, evaluation type 1 (both anonymous) does show a considerable lower mean CES (Figure 1). Figure 3, Figure 2 and Figure 5 show that there is an interaction effect between ideator and evaluator in terms of anonymity. Interestingly this only occurs when both are anonymous.

Our findings suggest that the effect of response bias also occurs in idea evaluation. This is based on the higher mean CES by evaluation where one of two identities, or both is/are known. However, different forms of response bias (social desirability, rater bias) do not seem to have a multiplier effect on the outcome (no interaction effect when both are known). The results of evaluation type 1 shows an interaction effect between ideator and evaluator when both are anonymous. As mentioned, the mean CES is significantly lower when both are anonymous, suggesting that there is no bias effect on the outcome. Our findings support previous research (Becker, 1976; Furnham, 1986; Nederhof, 1985) that social desirability bias has an effect on the outcome of idea evaluations. Results do suggest that idea evaluations are also affected by rater bias. We found evidence that idea evaluations that are completely anonymous result in a different (being considerably lower) mean CES. These findings are important because they impact the ideation strategy of an organization.

Theoretical Implications

Knowledge of the crowd is more and more considered as a valuable alternative to expert knowledge in idea management (Kijkuit & Van Den Ende, 2007; Zhu et al., 2018). As a result, the crowd is being used in idea evaluations. Especially in organizational ideation it is important to gain knowledge regarding social factors between ideator and evaluator that are of influence on the evaluation process. If organizations can identify the ideas with the most potential in an early stage, they can gain considerable advantage in their innovation strategy. However, it is important to understand the factors that can influence the evaluation process. The relation or social network connection between ideator and evaluator is of some influence on the evaluation outcome. An evaluation setting where ideator, evaluator of both are identifiable seem to result in higher mean CES. Therefore, response bias must be taken into consideration when organizations rely on idea evaluations for their future strategy. Becker (1976) suggested that anonymity in idea evaluation influences the outcome. Our researched shows that anonymity indeed affects idea evaluation. We found evidence suggesting the non-existence of response bias on the outcome when ideator and/or evaluator are both anonymous. This leads to a significantly lower mean CES. Future research could focus on this effect and if it results in a better forecasting model. Future research could also focus on the different types of bias and the type of bias has the most effect on the outcome.

Managerial Implications

Crowd sourced feedback proves to be useful for improving ideas on ideation platforms. Online interaction between ideator and 'the crowd' has the potential to improve ideas and thus helping organizations to be more successful with their ideation strategy. Crowd sourced evaluation is an effective way of selecting the best idea. However, when idea evaluation is influenced by response bias, social desirability and the relation between ideator and evaluator, organizations that operate an ideation platform need to take anonymity into consideration when using the (internal) crowd as evaluation source. This would create several challenges related to the design and/or operation of these platforms. The results of our study provide more insights into the effect of anonymity on idea evaluation. It shows that anonymity influences evaluation outcome and should be considered a factor when using crowd evaluations as an indicator for potential successful ideas. Our findings could also impact other forms of evaluation. For example, one could research how this affects the review of academic papers. Does response bias also affect the peer review process and are ratings lower when both reviewer and author are anonymous?

Limitations and future research

Our research consisted of a field research for which we made some assumptions and where we had to deal with some shortages. Due to the nature of our dataset we did not have access to certain meta data regarding ideator and evaluator. Hierarchical status, experience and/or age are some of the characteristics which could provide more insights (although our original intention was to have this data included in our research). However, our results do give a more generic conclusion on the effect of anonymity in idea evaluation. We have found that evaluations that are completely anonymous are rated lower than evaluations where the identity of ideator and/or evaluator is/are known. Our results do not provide an answer to which is best. Are the results of Evaluation Types 2, 3 and 4 an accurate indication for a potentially successful idea or do evaluations of Type 1 (Anonymous/Anonymous) provide a better prediction for success? Combined with a more detailed study of our data combined with characteristics of the ideator could lead to a model on which potentially successful ideas could be identified in the evaluation phase. We also made some assumptions (page 8) which could be argued. Another limitation of our research lies in the nature of the organization that were included in our study. We only studied one type of organization that consists of a large homogeneous type of employee. Future research should focus on different type of organizations.

References

Ågerfalk, P. J., & Fitzgerald, B. (2008). Outsourcing to an unknown workforce: Exploring opensourcing as a global sourcing strategy. *MIS Quarterly: Management Information Systems*, *32*(2), 385–409. Retrieved from https://www-jstor-org.eur.idm.oclc.org/stable/pdf/25148845.pdf?casa_token=STBn91N0pg8AAAAA:9JEsVvc_ZIzGDe7UQuhwrzxxiZpVojBdMGQcUWItDF3cXFXHYvT6fHf4ZuyKaFfZrUdEwX5-1KwcItV546EJWRiW-4GCMtWgVliZCV6xs7mbLwSoI-AjKw

Albors, J., Ramos, J. C., & Hervas, J. L. (2008). New learning network paradigms: Communities of objectives, crowdsourcing, wikis and open source. *International Journal of Information Management*, *28*(3), 194–202. https://doi.org/10.1016/j.ijinfomgt.2007.09.006

Antonioni, D. (1994). the Effects of Feedback Accountability on Upward Appraisal Ratings. *Personnel Psychology*, *47*(2), 349–356. https://doi.org/10.1111/j.1744-6570.1994.tb01728.x

Arrow, K. J. (1974). *The limits of organization*.

Banbury, C. M., & Mitchell, W. (1995). The effect of introducing important incremental innovations on market share and business survival. *Strategic Management Journal*, *16*(S1), 161–182. https://doi.org/10.1002/smj.4250160922

Bayus, B. L. (2012). Crowdsourcing New Product Ideas Over Time: An Analysis of Dell's Ideastorm Community. *Ssrn*, *59*(1). https://doi.org/10.2139/ssrn.1979557

Becker, W. M. (1976). Biasing Effect of Respondents' Identification on Responses to a Social Desirability Scale: A Warning to Researchers. *Psychological Reports*, *39*(3), 756–758. https://doi.org/10.2466/pr0.1976.39.3.756

Blohm, I., Riedl, C., Leimeister, J. M., & Krcmar, H. (2011, October 1). Idea Evaluation Mechanisms for Collective Intelligence in Open Innovation Communities: Do Traders Outperform Raters? Retrieved from https://papers-ssrn-com.eur.idm.oclc.org/sol3/papers.cfm?abstract_id=1936590

Bonds, I., Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., … Kraut, R. E. (2012). *Building Member Attachment in Online Communities: Applying Theories of Group*. *Source: MIS Quarterly* (Vol. 36).

Chesbrough, H. W. (2006). *Open innovation: Researching a new paradigm*. Oxford University Press on Demand.

Connolly, T., Jessup, L. M., & Valacich, J. S. (1990). Effects of Anonymity and Evaluative Tone on Idea Generation in Computer-Mediated Groups. *Management Science*, *36*(6), 689–703. https://doi.org/10.1287/mnsc.36.6.689

Cooper, R. G. (1994). Third-Generation New Product Processes. *Journal of Product Innovation Management*, *11*(1), 3–14. https://doi.org/10.1111/1540-5885.1110003

CROWNE, D. P., & MARLOWE, D. (1964). The approval motive, studies in evaluative dependence.

Dell, N., Vaidyanathan, V., Medhi, I., Cutrell, E., Thies, W., & of Washington, U. (2012). *"Yours is Better!" Participant Response Bias in HCI.*

Dodgson, M., Gann, D., & Salter, A. (2006). The role of technology in the shift towards open innovation: The case of Procter & Gamble. *R and D Management*, *36*(3), 333–346. https://doi.org/10.1111/j.1467-9310.2006.00429.x

Effects of Message Valence, Communicator Credibility, and Source Anonymity ...: EBSCOhost. (n.d.). Retrieved October 23, 2019, from https://web-b-ebscohost-com.eur.idm.oclc.org/ehost/pdfviewer/pdfviewer?vid=1&sid=2ae1fa02-897e-4105-ba66-64d28ee9f653%40pdc-v-sessmgr06

Ehrlich, S., Galbraith, C. S., Ehrlich, S. B., & Denoble, A. F. (2006). Predicting Technology Success: Identifying Key Predictors and Assessing Expert Evaluation for Advanced Technologies. *Article in The Journal of Technology Transfer*. https://doi.org/10.1007/s10961-006-0022-8

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*. https://doi.org/10.1287/isre.1080.0193

Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, *7*(3), 385–400. https://doi.org/10.1016/0191-8869(86)90014-0

Gangi di, P. M., & Wasko, M. (2009). Steal my idea! Organizational adoption of user innovations from a user innovation community A case study of Dell IdeaStorm Gangi Wasko 2009 (caracteristicas de la innovacion).pdf. *Decision Support Systems*, *48*, 303–312.

Gassmann, O. (2006, June). Editorial: Opening up the innovation process: Towards an agenda. *R and D Management*. https://doi.org/10.1111/j.1467-9310.2006.00437.x

Gassmann, O., Enkel, E., & Chesbrough, H. (2010, June). The future of open innovation. *R and D*

Management. https://doi.org/10.1111/j.1467-9310.2010.00605.x

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy*, *33*(6–7), 897–920. https://doi.org/10.1016/j.respol.2004.01.015

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2008). Idea Generation and the Quality of the Best Idea. *Ssrn*, *56*(4), 742. https://doi.org/10.2139/ssrn.1082392

Hemphill, T. A. (2005). Open Initiative: The New Imperative for Creating and Profiting from Technology. *Academy of Management Perspectives*, *19*(1), 164–165. https://doi.org/10.5465/ame.2005.15842012

Holzbach, R. L. (1978). Rater bias in performance ratings: Superior, self-, and peer ratings. *Journal of Applied Psychology*, *63*(5), 579–588. https://doi.org/10.1037/0021-9010.63.5.579

Hossain, M., & Islam, K. M. Z. (2015). Generating Ideas on Online Platforms: A Case Study of "My Starbucks Idea." *Arab Economic and Business Journal*, *10*(2), 102–111. https://doi.org/10.1016/J.AEBJ.2015.09.001

Howard, C. D., Barrett, A. F., & Frick, T. W. (2010). ANONYMITY TO PROMOTE PEER FEEDBACK: PRE-SERVICE TEACHERS' COMMENTS IN ASYNCHRONOUS COMPUTER-MEDIATED COMMUNICATION. *J. EDUCATIONAL COMPUTING RESEARCH*, *43*(1), 89–112. https://doi.org/10.2190/EC.43.1.f

Howe, J. (2006). Wired 14.06: The Rise of Crowdsourcing. https://doi.org/10.1086/599595

Kerin, R. A., & Peterson, R. A. (1976). Personalization, respondent anonymity, and response distortion in mail surveys. *Journal of Applied Psychology*, *62*(1), 86–89. https://doi.org/10.1037/0021-9010.62.1.86

Kijkuit, B., & Van Den Ende, J. (2007). The organizational life of an idea: Integrating social network, creativity and decision-making perspectives. *Journal of Management Studies*, *44*(6), 863–882. https://doi.org/10.1111/j.1467-6486.2007.00695.x

Lu, R., & Bol, L. (n.d.). *A Comparison of Anonymous Versus Identifiable E-Peer Review On College Student Writing Performance and the Extent of Critical Feedback Repository Citation*. Retrieved from www.ncolr.org/jiol

Lynn, G. S., Morone, J. G., & Paulson, A. S. (1996). Marketing and Discontinuous Innovation: The

Probe and Learn Process. *California Management Review*, *38*(3), 8–37. https://doi.org/10.2307/41165841

Miner, T. (2005). Book Review: The Wisdom of Crowds: Why the Many are Smarter than the Few, and how Collective Wisdom Shapes Business, Economies, Societies, and Nations. *Journal of Experiential Education*, *27*(3), 351–352. https://doi.org/10.1177/105382590502700325

Morgan, J., & Wang, R. (2010). Tournament for ideas. *Tournament for Ideas*, *52*(2), 77–97.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*(3), 263–280. https://doi.org/10.1002/ejsp.2420150303

Newell, A; Shaw, J.C.; Simon, H. A. (1962). *The process of creative thinking*. Retrieved from http://shelf1.library.cmu.edu/IMLS/BACKUP/MindModels.pre_Oct1/creativethinking.pdf

Piller, F. T., & Walcher, D. (2006). Toolkits for idea competitions: A novel method to integrate users in new product development. *R and D Management*, *36*(3), 307–318. https://doi.org/10.1111/j.1467-9310.2006.00432.x

Pisano, G. P., & Verganti, R. (2008). Which Kind of Collaboration Is Right for You? - Harvard Business Review. *Harvard Business Review*, 1–7. https://doi.org/Article

Poetz, M. K., & Schreier, M. (2012). The value of crowdsourcing: Can users really compete with professionals in generating new product ideas? *Journal of Product Innovation Management*, *29*(2), 245–256. https://doi.org/10.1111/j.1540-5885.2011.00893.x

Reitzig, M., & Sorenson, O. (2013). Biases in the selection stage of bottom-up strategy formulation. *Strategic Management Journal*, *34*(7), 782–799. https://doi.org/10.1002/smj.2047

Rogers, E. M. (1995). *Diffusion of Innovations* (4th Editio). New York: Free Press.

Surowiecki, J. (2004). *The Wisdom of Crowds*.

Teece, D. J. (1986). *Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy*.

Trist, E; Bamforth, K. (1954). Some social and psychological consequences of the long wall method of coal-getting. *Human Relations*.

van de Vrande, V., de Jong, J. P. J., Vanhaverbeke, W., & de Rochemont, M. (2009). Open innovation in SMEs: Trends, motives and management challenges. *Technovation*, *29*(6–7), 423–437.

https://doi.org/10.1016/j.technovation.2008.10.001

Vandenbosch, B., Saatcioglu, A., & Fay, S. (2006). Idea management: A systemic view. *Journal of Management Studies*, *43*(2), 259–288. https://doi.org/10.1111/j.1467-6486.2006.00590.x

Zaggl, M. A., Schöttl, C., Schweisfurth, T. G., & Raasch, C. (2018). Hierarchical distance and idea evaluation in enterprise crowdfunding. *International Conference on Information Systems 2018, ICIS 2018*, (Berg 2016), 1–17.

Zhu, H., Kock, A., Wentker, M., & Leker, J. (2018). How Does Online Interaction Affect Idea Quality? The Effect of Feedback in Firm-Internal Idea Competitions. *Journal of Product Innovation Management*, *36*(1), 24–40. https://doi.org/10.1111/jpim.12442