Does god play dice? Randomness vs. deterministic explanations of crowdsourcing success

Roiser, S. [a], Franke, N. [b], Lettl, C.[b], Türtscher, P.[c]

[a] FH Wiener Neustadt für Wirtschaft und Technik GmbH, Johannes Gutenberg-Straße 3, A-2700 Wiener Neustadt, AUSTRIA
[b] WU Wien, Welthandelsplatz 1, A-1020 Wien, AUSTRIA
[c] VU University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, THE NETHERLANDS

ABSTRACT:
Which factors are responsible for the success of crowdsourcing tournaments? Current theorizing on crowdsourcing appears to assume that there is a deterministic relationship between factors such as the organization of the tournament, characteristics of the participants attracted, and specific situational factors on the one hand and the quality of their contributions gained on the other. Based on theory that views creativity as a process of blind variation and retention, we introduce the alternative idea that in fact the quality of any participants' idea is largely random and thus the success of the tournament rests primarily on the number of participants attracted. In order to compare the explanatory power of randomness and 22 deterministic factors derived from literature we conducted a huge experiment in which 1,089 participants developed ideas for smartphone apps. Our finding is unambiguous: the single factor of randomness out-performs all deterministic explanations collectively by far. It appears that in crowdsourcing, God indeed plays dice.

1 INTRODUCTION

Recently, many companies have introduced crowdsourcing tournaments to outsource creative tasks such as idea generation, product or logo design to undefined solution bases outside the firm to gather high quality solutions with high levels of originality and innovativeness (Afuah & Tucci, 2012; Bayus, 2012; Bullinger, Neyer, Rass, & Moeslein, 2010; Terwiesch & Xu, 2008). While a number of such crowdsourcing tournaments are reported to be tremendously successful (Boudreau & Lakhani, 2009) others failed to produce the desired outcome. This prompted many scholars to investigate the factors determining crowdsourcing success, both conceptually (e.g. Afuah & Tucci, 2012; Terwiesch & Xu, 2008) as well as empirically (e.g. Boudreau, Lacetera, & Lakhani, 2011; Jeppesen & Lakhani, 2010). The underlying (implicit) assumption of the vast majority of studies is that the creative output generated by crowdsourcing tournaments is (1) determined by specific activities set by the crowdsourcing organizers (such as the design of the tournament or the incentives given), the characteristics of the participants attracted, and specific situational factors, (2) which are sufficiently general and stable to describe them empirically. Regarding characteristics of participants, for example, Frey et al. (2011: 398) "concentrate on the roles of motivation and knowledge as *determinants* of the quality of such individual contributions" (emphasis by the authors). Yet findings are quite heterogeneous. According to a pre-study conducted by the authors also expert practitioners such as platform providers of crowdsourcing tournaments or managers of firms that seek creative input via crowdsourcing tournaments diverge to a large extend with respect to the relative importance of success factors.

In this article, we suggest that the observed lack of a consensus in empirical studies and among expert practitioners may be due to the inherent limitations of the deterministic perspective. We investigate to what degree the output quality of crowdsourcing tournaments is in fact random. There are many indicators that this also holds in the business context (Langer, 1975; Mlodinow, 2009). Managers and other decision makers often misperceive patterns, order, and causality in data structures that are in fact random (Kahneman, 2011). CEOs, for example, intuitively construct causal stories such that their decisions make sense (Salancik & Meindl, 1984).

This might also explain why research has not taken up this perspective in the numerous studies on crowdsourcing success even though some few prior studies reveal indications of randomness. Bayus (2012) found that successful contributors in crowdsourcing tournaments are unlikely to repeat their early success, which allows the conclusion that maybe also the initial success was in fact a matter of luck rather than the result of specific reasons. Our research question thus is how the explanatory power of randomness and deterministic causal factors compare in explaining the success of crowdsourcing tournaments. It is important to know which of the two explanations matter more as they have opposing managerial implications. If on the one hand the success of crowdsourcing is determined by specific factors, then it is important to carefully design tournaments in a way that it corresponds to these factors. If on the other hand success is random, all that matters was to get an as large crowd as possible, corresponding to the law of large numbers.

2 METHOD

In order to measure the relative explanatory power of randomness and causal factors we conducted an experimental crowdsourcing tournament with a typical ideation task. We manipulated the organization of the tournament and measured the participants' expertise, their skills, their personality traits, and situational factors. Success of crowdsourcing in new product ideation is mostly conceived as originality of ideas, i.e. the degree how much they differ from existing paradigms and involve radically new functions, designs, and elements.

Our dependent variable is therefore the originality of the ideas generated. This setting allowed us to analyze data on two different levels. First, we could determine on the individual level in how far the numerous causal factors we measured actually explain the originality of the individual participants' ideas. Randomness is the "invisible guest" in this analysis – it corresponds to the remaining unexplained variance. On the aggregate level, we could model randomness more directly. In this second step, we used our data for a simulation of crowdsourcing tournaments. This corresponds to the perspective of a company that is rather interested in the outcome of the total crowdsourcing tournament than in the performance of each participant. For each crowdsourcing tournament simulated, we randomly drew participants from the overall sample. As we knew their ideas, we could measure what would have happened had we organized the tournament in this specific way and had been able to attract a crowd with these specific characteristics in these specific situational circumstances. The dependent variable was the originality of the best ideas obtained in this specific tournament. The independent variables were the specific crowds' overall expertise, their skills, their personality traits, and situational factors. Randomness was captured by the size of the crowd, which we varied from 10 to 100 in the 36,400 tournaments we simulated. This is so because the law of large numbers basically says that the chance of getting a high number of spots is a function of the number of dices rolled, given that die casts are actually random. Comparing the variances explained by the size of the crowd with the variance explained by all deterministic factors allows an answer to our guiding question in how far the success of crowdsourcing is determined by randomness.

2.1 Setting and participants of the crowdsourcing experiment

We conducted an ideation-based crowdsourcing tournament for smart phone apps in its natural web 2.0 environment. We had chosen the development of ideas for smart phone apps as the object of our study for two reasons. First, seeking for novel app ideas by means of crowdsourcing tournaments appears quite typical, which is visible by recent examples. Second, the ideation task can also be generalized to many other crowdsourcing tournaments as it has a broad, almost infinite solution space, and the quality of solutions is not arbitrary. Succeeding in such a tournament requires not only creativity and imagination but also skills (e.g. to describe the idea properly) and knowledge (particularly regarding already existing apps). We invited participants to an idea competition regarding “mobile communication of the future”, sponsored by the smart phone producer Apple Computer and the network provider Orange. As an incentive to participate, we announced prizes amounting to a total of € 50,000. We obscured the study intention in order to avoid problematic self-selection biases and the possibility to prepare an idea beforehand. People interested were informed that the contest would take place in two relatively narrow time slots (five hours in the evening time of two different days, a Thursday and a Saturday) we had defined in order to control the experiment.

We broadly promoted the study and thus attracted a relatively large gross sample of n= 2,599 participants out of which 55 % completed our questionnaire. We precluded double entries by controlling for IP addresses. However, we had to exclude cases due to data lost (223), a critical number of missing values (8), regular aborts (2), a task time of less than 100 seconds (40), inconsistent responses (28), multiple participation (24), minimum completing time of 20 seconds per form in all questionnaires (16). Thus, our net sample consists of 1,089 participants. Most participants were male (68.7 %) students (87.9 %) with a mean age of 25.7 years (s.d. = 7.1), thus the sample largely corresponds to typical crowdsourcing tournament participants (Füller 2010). They came from Austria (69.5 %), followed by Germany (25.7 %) and Switzerland (2.7 %).

2.2 Experimental design

We employed an online 2*2*2 between-subject experiment on a website we had programmed for this study. It consisted of seven phases. (1) When entering the website, participants logged in and received a short introductory text that described the procedure of the crowdsourcing tournament. Particularly, they were informed that they would get a specific task and had a maximum of 15 minutes to complete it. (2) Participants were then asked to self-assess their creativity (see operationalization). We measured this factor before the participants submitted their ideas in order to avoid a potential halo-effect. (3a) Then participants received the actual task of the tournament: they should develop an idea for a novel and innovative every-day-app for smart phones that should be interesting to as many users as possible. They were randomly assigned to one of the two conditions of the task framing (narrow or broad, see next section) and (3b) to one of the two conditions regarding the incentive (incentive or none, see next section). (4) Then idea generation started. There was a blank field for a clear headline and a blank field that allowed for a text with a maximum of 1,000 characters. Participants were permanently informed about the remaining time. (5) After exactly 1 minute it was randomly determined whether they would keep on working alone or in the interaction condition. In the latter case, the participant was assigned as interaction partner to another participant of the study. This dyad was then enabled to assist each other. (6) After a maximum of fifteen minutes participants had to submit their idea, which they did on average after 9.73 minutes (s.d.=3.79). (7) After this, they were directed to an online questionnaire with the key variables. From the net sample, 622 participants were in the interaction condition and 467 worked alone. 540 participants received the broad task framing as opposed to 549 who received the narrow task framing. The number of participants in the incentive condition accounts for 527, whereas those who had not the opportunity to win the award counted 562 participants. Tests showed that there are no significant differences regarding the distribution of independent variables between the eight treatment conditions which points to an effective randomization.

3 FINDINGS

Randomness rules in our crowdsourcing tournament. The originality of the contributions is explained only to a very limited degree by the 22 deterministic factors we derived from the general literature on creative problem-solving and the more specific literature on crowdsourcing tournaments and hence had incorporated in our measurement. Our simulation shows that randomness outperforms all deterministic factors collectively by 532 %. We differentiate findings on the individual level regarding the determinants of idea originality and on the aggregate level and discuss avenues for further research and limitations.

3.1 Determinants of idea originality on the individual level

We tested the causal explanators as described in the research framework with OLS regression analyses. Overall, results allow the conclusion that randomness indeed plays a major role in determining the originality of an idea submitted. The total model (Model 1) shows that although we include 22 independent variables and thus basically all causal factors discussed in the literature, 93.6 per cent of the variance of the dependent variable is left unexplained (R^2=.083, p<.001). Even if we take measurement error into account this suggests that the invisible guest of randomness has major prominence in our crowdsourcing tournament.

Among the variables discussed in the literature, ten of the variables turned out to be significant predictors as suggested by the literature (four of them only marginal). The sign of the coefficients is in the expected direction – with one exception: if we announced an incentive for delivering a good idea, this had a negative impact on the originality of the idea submitted (b=-.039, p<.1). A possible explanation are crowding out effects that have been reported in crowdsourcing before (Bayus, 2012; Frey et al., 2011).

The strongest influence has the group of situational variables with an R^2 of .047 (p<.001). Variance explained by the organization of the tournament (Model 2), participants' expertise (Model 3), skills (Model 4), and personality traits (Model 5) is surprisingly low

	DV=originality of submitted idea					
	Model 1: All variables	**Model 2: Organization**	**Model 3: Participants' expertise**	**Model 4: Participants' skills**	**Model 5: Participants' traits & roles**	**Model 6: Situation**
Organization						
Incentives	-.039†	-.033				
Interaction	.053*	.117***				
Task framing	.000	-.008				
Participants' expertise						
Domain-specific expertise	.048†		.070*			
Analogous domain expertise	-.024		.014			
Ideation task experience	.007		.003			
Lead userness	-.030		-.010			
Participants' skills						
Business skills	-.042			-.028		
Technical skills	.025			.052†		
Creative skills	.041			.067*		
Competence profile	-.012			.009		
Education level	.069*			.095**		
Participants' traits and roles						
Creativity	.039				.072*	
Information hub	-.037				-.033	
Boundary spanner	.033				.031	
Outsiderness	.053†				.019	
Age	.038				.066*	
Gender	-.049†				-.051*	
Situation						
External support	.063*					.066*
Motivation	.057*					.048†
Time spent	.121***					.129***
Timing of the tournament	-.141***					-.141***
Adj. R^2	**.064******	**.012******	**.001**	**.015******	**.011*****	**.044******
R^2	.083***	.015***	.005	.020***	.017**	.047***
F	4.367	5.353	1.249	4.316	3.033	13.453
N	1,089	1,089	1,089	1,089	1,089	1,089

† = p < 0.1, * = p < 0.05, ** = p < 0.01, *** = p < 0.001 (one-sided)
Standardized coefficients are shown.

Table 1. individual level analysis

3.2 Determinants of idea originality on the aggregate level

Again we use OLS regressions for analyses. Overall, the aggregation and particularly the inclusion of the crowds' size resulted in a high level of variance explained (Model 1, R^2=.725)2. Obviously, crowd size explained by far most variance, 5.32 times as much as the other 22 independent variables collectively (Model 2 and 3). The second strongest effect comes from the incentive – again as a crowding out effect (see Table 2).

	DV= mean originality of the top ten ideas		
	Model 1: All variables	**Model 2: All deterministic factors**	**Model 3: Crowd size**
Crowd size	.782		.782
Organization			
Incentives	-.214	-.218	
Interaction	.091	.102	
Task framing	.012	.014	
Crowds' expertise			
Domain-specific expertise	.019	.025	
Analogous domain expertise	.042	.038	
Ideation task experience	-.029	-.018	
Lead userness	.005	.004	
Crowds' skills			
Business skills	-.012	-.017	
Technical skills	.089	.093	
Creative skills	.006	.004	
Competence profile	.081	.088	
Education level	.017	.015	
Crowds' traits and roles			
Creativity	.036	.038	
Information hub	-.015	-.020	
Boundary spanner	.008	.007	
Outsiderness	-.035	-.048	
Age	.037	.024	
Gender	.017	.008	
Situation			
External support	.012	.013	
Motivation	.085	.092	
Time spent	.003	-.012	
Timing of the tournament	-.016	-.022	
Adj.. R^2	**.725**	**.115**	**.612**
R^2	.725	.115	.612
F	4180.553	215,168	57448.48
N	36,400	36,400	36,400

Standardized coefficients are shown.
Note that we do not indicate significance levels (due to the permutation design with an artificial sample size of 36,400 all coefficients are significant).

Table 2. tournament level analysis

4 DISCUSSION

We contribute to the quickly evolving literature that investigates the factors explaining the success of crowdsourcing tournaments (Boudreau et al., 2011; Jeppesen & Lakhani, 2010; Leimeister et al., 2009; Poetz & Schreier, 2012; Terwiesch & Xu, 2008). The factor we add to this line of research is systematically from extant factors as it involves a different weltanschauung, namely a non-deterministic perspective. In a way this resembles the discussion in quantum mechanics in how far the world is deterministic or governed by pure chance (Bell, 2004). Randomness is also systematically different from extant factors from another perspective: its effect size is much greater. The obvious conclusion for managers who consider starting a crowdsourcing tournament for their new product ideation processes is that they are well advised to recruit as many participants as possible. The degree in how far this is achieved is far more important than the exact organization and the composition of the crowd attracted. Certainly, there will be minimum qualifications for participants and also we must not forget that an unprofessional, unattractive, or unfair design of the tournament will inevitably result in recruitment problems. However, the clear focus must be to increase the number of participants.

This important conclusion leads to a number of follow up questions that constitute opportunities for further research. The first question is in how far our findings can be generalized. The task we employed was typical for crowdsourcing tournaments in the area of new product ideation. It does not require specific technical knowledge as in expertise-based crowdsourcing projects such as scientific problem solving tournaments (e.g. Boudreau et al., 2011; Jeppesen & Lakhani, 2010) where randomness might play a less prominent role. Furthermore, the self-selection of participants in ideation-based crowdsourcing tournaments tends to lead to a certain assimilation of participants' interest and knowledge about the product category (in our case: smartphones) which may suppress the importance of individual differences. Crowdsourcing tournaments in the area of new product ideation as applied in this study also differ quite significantly from lead user studies as such studies aim for the development of physical prototypes and not just pure ideation.

BIBLIOGRAPHY

[1] Afuah, A.; Tucci, C. L. (2012): Crowdsourcing as a Solution to Distant Search. In: Academy of Management Review, 37(3), S. 355-375.

[2] Bayus, B. (2012): Crowdsourcing New Product Ideas Over Time: An Analysos of Dell's Ideastorm Community. In: Management Science, Articles in Advance, S. 1-19.

[3] Boudreau, K. J. & Lakhani, K. R. (2009): How to Manage Outside Innovation. In: MIT Sloan Management Review, 50(4), S. 69-76.

[4] Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. In: Management Science, 57(5), S. 843-863.

[5] Bullinger, A. C., Neyer, A. K., Rass, M., & Moeslein, K. M. (2010): Community-Based Innovation Contests: Where Competition Meets Cooperation. In: Creativity and Innovation Management, 19(3), S. 290-303.

[6] Kahneman, D. (2011): Thinking, Fast and Slow. London: Penguin Group.

[7] Frey, K., Luethje, C., & Haag, S. (2011): Whom Should Firms Attract to Open Innovation Platforms? The Role of Knowledge Diversity and Motivation. In: Long Range Planning, 44(5-6), S. 397-420.

[8] Jeppesen, L. B. & Lakhani, K. R. (2010). Marginality and Problem-Solving Effectiveness in Broadcast Search. In: Organization Science, 21(5), S. 1016-1033.

[9]

[10] Langer, E. J. (1975): The Illusion of Control. In: Journal of Personality and Social Psychology, 32(2), S. 311-328.

[11] Mlodinow, L. (2009): The Drunkard's Walk: How Randomness Rules Our Lives. New York: Random House, Inc.

[12] Salancik, G. R. & Meindl, J. R. (1984): Corporate Attributions as Strategic Illusions of Management Control. In: Administrative Science Quarterly, 29(2), S. 238-254.

[13] Terwiesch, C. & Xu, Y. (2008): Innovation contests, open innovation, and multiagent problem solving. In: Management Science, 54(9), S. 1529-1543.