

Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2019 Proceedings

Crowds, Social Media and Digital Collaborations

Distant Search, but Local Implementation? Using the Crowd's Evaluation to Overcome Organizational Limitations in the Selection of Crowdsourced Ideas

Lisa Wimbauer

University of Passau, lisa-kristina.wimbauer@bmwgroup.com

Patrick Figge

University of Passau, patrick.figge@uni-passau.de

Carolin Haeussler

University of Passau, carolin.haeussler@uni-passau.de

Follow this and additional works at: <https://aisel.aisnet.org/icis2019>

Wimbauer, Lisa; Figge, Patrick; and Haeussler, Carolin, "Distant Search, but Local Implementation? Using the Crowd's Evaluation to Overcome Organizational Limitations in the Selection of Crowdsourced Ideas" (2019). *ICIS 2019 Proceedings*. 1.

https://aisel.aisnet.org/icis2019/crowds_social/crowds_social/1

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2019 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Distant Search, but Local Implementation? Using the Crowd's Evaluation to Overcome Organizational Limitations in the Selection of Crowdsourced Ideas

Completed Research Paper

Lisa Wimbauer

University of Passau
Innstr. 27, 94032 Passau, Germany
wimbau01@gw.uni-passau.de

Patrick Figge

University of Passau
Innstr. 27, 94032 Passau, Germany
patrick.figge@uni-passau.de

Carolin Haeussler

University of Passau
Innstr. 27, 94032 Passau, Germany
carolin.haeussler@uni-passau.de

Abstract

Organizations increasingly apply crowdsourcing to search for novel ideas. However, in selecting ideas, organizations face the challenge to select the best suggestions without prematurely rejecting the distant solutions they initially set out to find. We argue that local search leads to differences between crowd and organizational evaluators. Further, organizational evaluation is restricted by the use of formal evaluation criteria. We propose that crowds can alleviate this by detecting ideas otherwise overlooked by organizational evaluators. Our analysis is based on 869 crowdsourced ideas evaluated by the crowd and by organizational evaluators. Our results suggest that the favorite ideas of the crowd are more novel than the organization's favorites. We also find that the crowd can be used to detect ideas that were, despite their potential, initially overlooked but later implemented by the organization. We contribute to literature on local search and its repercussions in the evaluation of crowdsourced ideas for implementation.

Keywords: Crowdsourcing, idea evaluation, idea implementation, local search

Introduction

"If I had asked people what they wanted, they would have said faster horses." – Henry Ford¹

Organizations often doubt that outsiders are able to express and communicate their needs for innovative products. In particular, automotive producers question their customers' ability to generate innovative ideas that account for the complexity of automotive production and are in line with the industry's quality standards (Illi et al. 2010; Seidel and Langner 2015). This corresponds to the manufacturer-active

¹ Although often attributed to Henry Ford, the quote is unconfirmed. Nonetheless, the continued and widespread use of the quote exemplifies the prevalence of the expressed mindset.

paradigm in innovation research, according to which product development is performed by organizational experts with users in a passive, merely supporting role. This paradigm has been questioned by a view that proposes that users can successfully play a more active part in the innovation process (Baldwin and von Hippel 2011; von Hippel 1978).

Crowdsourcing is increasingly applied by organizations in order to search for new ideas outside of their organizational boundaries (Afuah and Tucci 2012). Although the ideas generated by crowds can often match internally generated ideas in terms of customer benefit and higher novelty (Poetz and Schreier 2012), organizations face the challenge to identify promising ideas within a large pool of suggestions. Especially for highly novel ideas, traditional selection processes seem problematic. When organizations have to assess more and more ideas, they increasingly pay less attention to distant suggestions (Piezunka and Dahlander 2015). This tendency is contrary to the organizations' reasons to use crowdsourcing in the first place. While crowds are deployed to enable more distant search, selection processes may still favor ideas which are based on knowledge that is more familiar to the organization. Organizations have started to use crowds for idea selection in the hope to address these challenges. Crowds are used to reduce the workload of organizational evaluators in handling large numbers of ideas. Furthermore, crowd evaluations may provide access to additional insights and improve the organization's decision-making on idea implementation. Using the crowd's knowledge, skills, and opinions not only for idea generation, but also for idea selection, is key in contexts that requires diverse knowledge (Lakhani et al. 2013).

The purpose of this paper is to generate insights into the differences between crowd and organizational evaluations, and the crowd evaluations' predictive value for subsequent implementation by the organization. Although one of the main reasons for organizations to use crowdsourcing is to profit from more distant search processes and source more novel ideas (Afuah and Tucci 2012), we know surprisingly little about how more distant ideas fare in the selection process, and whether crowds can be used to balance potential biases of organizational evaluators. Despite first indications that crowds can be useful in idea selection (Mollick and Nanda 2016), little is known about why and how crowd and organizational evaluators differ in their assessment. Moreover, prior studies often argue that crowds are prone to biases, whereas potential biases of organizations have received less attention in crowdsourcing contexts. For instance, certain aspects in the representation of decision situations, complex language, and appealing visual presentation of ideas can distort the crowd's evaluation (Mollick and Nanda 2016; Simmons et al. 2011). But how are organizations biased and restricted in their evaluation behavior when deciding on which ideas to implement?

We use data from a leading German automotive manufacturer's idea generation contests in 2010. Our sample contains 869 ideas, which were all evaluated by organizational experts and received votes from the crowd. We investigate the differences between crowd and organizational evaluators in their assessment of novel ideas. Based on the theoretical foundation of local search behavior, we argue that organizational evaluators systemically undervalue crowdsourced ideas which are more novel to the organization and that the evaluation criteria that organizations use to assess ideas further restrict their ability to select promising ideas from the submissions. We compare the ideas favored by the crowd and by the organization with projects implemented by the organization eight years after the contest. Therefore, we are able to examine whether and to what extent organizations could have profited from considering the ideas the crowd would have selected.

Indeed, our findings confirm that organizations favor less novel ideas than the crowd. Further, we find that the consideration of crowd evaluations can prevent organizations from dismissing ideas prematurely. Despite their potential, some ideas were first ignored by the organization, but have later been implemented. We show that among the ideas that an organization implements several years after the contest, a considerable proportion was among the crowd's (but not the organization's) favorites during the contest.

Theory

Organizational and Crowd Evaluation

We follow the definition of crowdsourcing as “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, p. 1). In our empirical context, the function that has been outsourced is the generation of ideas for new products and services. Scholars have started to investigate if crowds can also be used for the selection of ideas (Magnusson et al. 2016; Velamuri et al. 2017).

One reason for organizations to use crowds is to improve idea selection, i.e. to identify more ideas worth implementing and to make fewer mistakes in doing so. Similar to the motivation of organizations to use crowdsourcing for idea generation, integrating crowds in the selection process promises additional insights in addition to the experts’ assessment. Crowd members generally combine more diverse knowledge than organizational employees and often have access to different knowledge than the organization. Another reason to use crowds is to relieve the pressure from experts to evaluate the often high number of ideas (Criscuolo et al. 2017). To the extent that crowd evaluation yields similar conclusions, organizations might be able to replace expert judgment without making concessions with regards to selection quality. Due to the large number of individuals in the crowd and their often non-monetary incentives, organizations might be able to save resources by using crowds, for instance, by pre-selecting a smaller number of ideas with the help of the crowd, which are then evaluated by organizational experts (Velamuri et al. 2017).

Only few prior studies compare the evaluation of crowds and experts. For instance, Magnusson et al. (2016) analyze the conformance between user and expert evaluations for different levels of expertise and rating methods. Similarly, Mollick and Nanda (2015) compare the funding decisions of crowd and experts for theater projects. Both studies show that crowd evaluations and expert assessments show considerable levels of agreement. However, prior research also shows that, unlike experts, crowds can be swayed by characteristics of ideas that might not be very informative for idea quality. For instance, they tend to favor ideas with a certain presentation style (e.g., ideas with pictures) and a certain writing style (e.g., low complexity).

Organizational evaluators are usually experts in the knowledge domains that are thought to be pertinent to the contest and their assessments rest on their technical, procedural, and solution knowledge (Poetz and Schreier 2012). However, organizational evaluators base their assessment not only on their expert domain knowledge, but their evaluation may also be influenced by the organizational context. For instance, organizational evaluators who have to assess a large number of ideas face an additional workload, which reduces the attention they devote towards highly novel ideas (Piezunka and Dahlander 2015). In addition, their decision may not only be affected by their expert knowledge and their limited capacity. Compared to crowds, organizational evaluators do not only evaluate the information in the idea description but also react to cues in the organizational environment. Of the few studies that compare crowd and organizational evaluation, the influence of organizational context on idea evaluation is so far not investigated. For instance, Magnusson et al. (2016) analyze users’ and experts’ evaluations in a field experiment, and Mollick and Nanda (2016) compare crowd and expert funding decisions in the context of crowdfunding for art projects, where such considerations do not play a role.

Local and Restricted Search

In an organizational context, experts may consider aspects unknown to the crowd, such as how well an idea reflects current strategic thrusts of the organization or how well it fits to ongoing projects and capabilities. Yet, prior studies have not yet investigated how organizational context leads to differences in the assessment by experts and the crowd.

To do so, we build on theory of local search (Katila and Ahuja 2002; Piezunka and Dahlander 2015; Stuart and Podolny 1996) to develop hypotheses on the differences between the organization’s and the crowd’s evaluation of ideas. Organizations search for information inside but also outside their boundaries to solve problems or to generate ideas (Afuah and Tucci 2012; Jeppesen and Lakhani 2010). Search behavior can be characterized according to the degree to which new knowledge is explored by the organization, i.e. if an organization’s search is local, close to its own knowledge base, or more distant (Katila and Ahuja 2002; Levinthal and March 1981). The search and processing of information precedes idea generation as well as idea evaluation. If organizational members search and use more local information, the ideas they generate are less novel. Similarly, during idea evaluation, organizational evaluators activate or acquire relevant information according to which they make their decisions. The kind of information available for idea evaluation will also influence which ideas are selected (Piezunka and Dahlander 2015).

The term ‘local search’ describes the tendency of organizations to build on information and experiences that are readily available to them, rather than to search for new, distant knowledge (Lopez-Vega et al. 2016; Stuart and Podolny 1996). First, local search can occur as organizational members fall into ‘competency traps’ (Rosenkopf and Nerkar 2001). Organizations decide to specialize in specific research and development areas in order to secure a competitive advantage. Thus, organizational members gain more experience in these domains over time. This limits their awareness for alternatives. Especially if the environment is changing, this may lead to suboptimal decisions based on outdated information (Reichwald and Piller 2009). Second, organizational members may search locally as they pursue an ‘economic mindset’. The search for and transfer of distant knowledge is more difficult and takes longer, and is therefore associated with higher costs. To avoid additional costs of more distant search, organizational members focus on their own knowledge and search more locally, for instance by consulting in-house experts (Reichwald and Piller 2009). Third, the ‘functional fixedness’ of organizational members can be another reason for local search behavior. Organizational members tend to prefer to use known, pre-existing solutions, even if the organization is confronted with a new situation. Decisions are made relying on previous experiences and pre-existing knowledge, even if no longer adequate in a particular situation (Gambeta et al. 2019; Reichwald and Piller 2009; Sorensen and Stuart 2000).

Local search is a general phenomenon that can affect organizational evaluators as well as crowd members. The crowd, however, is made up of a large number of individuals with diverse knowledge. Therefore, on the aggregate level, the local search behavior of its individual members has the potential to cover a broad knowledge area. What constitutes local search for one individual can constitute distant search for another. In addition, the information that organizational evaluators use to assess an idea can also be restricted by the evaluation criteria that are applied in the selection process. Organizations themselves determine the evaluation criteria according to which crowdsourced ideas are assessed (Mueller et al. 2018). These criteria may further limit the search for information and the kind of considerations that organizational evaluators engage in. Formal evaluation criteria determine the questions that evaluators try to answer. Rather than searching for information freely, their search is constrained to specific directions. For instance, if the market potential of an idea is to be evaluated, information search might be directed towards comparing sales data on current, comparable products – while disregarding other search directions. These constraints on their evaluators’ information search and evaluation behavior are self-imposed by organizations. While they may be effective in reaching certain goals, they might also have unintended side effects.

Hypotheses

In the following, we introduce two sets of hypotheses. First, we argue that three facets of an idea’s ‘novelty’ create evaluation differences between the crowd and the organization. As organizational evaluators are restricted by local search, more novel ideas are more likely to be favored by the crowd. Second, we will argue that the more formal evaluation criteria of organizational evaluators constitute a further constraint to their evaluation behavior. We argue that the crowd, unconstrained by these formal criteria, also selects ideas that the organization initially overlooks, but later decides to implement.

Evaluation Differences due to Novelty

As organizational evaluators are restricted by local search, more novel ideas are more likely to be favored by the crowd. As idea novelty is conceptualized and operationalized differently within each study, a comparison and generalization of outcomes across studies is challenging. For instance, Mollick and Nanda (2016) measure novelty by an expert assessment whereas Haas et al. (2015) use a language-based method. A literature review conducted by Dean et al. (2006) summarizes measurements of idea novelty used in over 90 studies. Their review reveals three recurrent dimension of novelty, which are most widely used: originality, rarity, and paradigm modification.

Originality

Original ideas “have the characteristic of being ingenious or imaginative” (Dean et al. 2006, p. 659). An original idea is an out-of-the-box thought that is “unusual, unique, and fresh” (Magnusson et al. 2016, p. 8). In our context, these are the ideas that deviate clearly from an organization’s product pipeline and

innovation strategy. Organizations follow strategic thrusts and use them as reference point for idea assessment (Behrens et al. 2014). Thus, an idea is original for an organization if it departs from its strategic thrusts.

Organizational evaluators will consider these strategic thrusts in their evaluation. When an organization decides to pursue a strategy and builds competencies accordingly, ideas which are not in line with this strategy will be evaluated less favorably (Criscuolo et al. 2017; Froehlich et al. 2016). This is not self-evident and may actually conflict with the organization's goals, as crowdsourcing is used to tap into knowledge resources outside of the company's usual search scope. We argue that there are two reasons why organizational evaluators judge original ideas less favorably. First, for ideas that comply with the organization's strategy, evaluators are more likely to either possess the knowledge required to assess them, or have access to that knowledge via colleagues. The knowledge needed to assess more original ideas is more difficult to obtain for them, and poorly understood ideas will not be among the organizational evaluators' favorites. One could argue that evaluators might invest more time to acquaint themselves with more original ideas, but this is unlikely to happen. Prior research shows that ideas which do not match with an evaluator's expertise receive less attention (Haas et al. 2015). Second, organizational evaluators are employees voting on behalf of the organization. They try not to consider personal preferences but rather focus on the ideas' perceived value for the organization. Ideas that are aligned with the organization's current strategy are more likely to be perceived as valuable (Di Gangi and Wasko 2009; Hienerth and Riar 2013).

Crowds, in contrast, are rather independent from the organization in their assessment. They are not compensated for selecting ideas and are not held responsible for their evaluations (Mollick and Nanda 2016). As their career prospects do not depend on their ability to predict the idea's success and they are not committed to the organization and its strategy in the same way, crowd evaluators favor idea independent of their accordance to the organization's strategy. Crowds may even appreciate more original ideas. For instance, new product development studies have found that crowds are more willing to pay for original concepts (Dahl and Moreau 2002). Crowd evaluators do not share the risks of organizational evaluators. For instance, organizational evaluators risk their reputation or career prospects, if they stand behind original ideas that may turn out to be nonstarters. Moreover, similar to the mechanism that organizations try to use in idea generation during crowdsourcing contests, the knowledge necessary to adequately understand and appreciate an original idea is likely to be present somewhere among the crowd members. Crowd members with the right expertise self-identify and assess the idea. Distant search for the required knowledge might not need to occur, as some of the diverse crowd members already possess the required knowledge.

Hypothesis 1a: Original ideas are more likely favored by the crowd rather than by the organization.

Rarity

The rarity of an idea expresses how exceptional and unusual the idea is compared to the total pool of ideas (Dean et al. 2006). In our context, these are the ideas that are different from the other ideas generated in the contest. Rare ideas are standalone ideas submitted by only one crowd member, as opposed to ideas that share a closer, more similar content description with other ideas in the contest (Haas et al. 2015; Piezunka and Dahlander 2015).

The emergence of similar ideas in a contest indicates that these topics address a widely held need, suggestive of high market acceptance of the innovation (Kornish and Ulrich 2011). If several crowd members submit an idea with similar content, clusters or "opportunity spaces" (Kornish and Ulrich 2011, p. 108) emerge. Organizational evaluators expect that these ideas are interesting for a higher number of potential customers. Organizational evaluators therefore anticipate higher demand for ideas within a cluster of ideas with similar content (Hofstetter et al. 2018; Kornish and Ulrich 2011). To make a similar assessment for unusual, exceptional ideas, organizational evaluators would have to engage in separate search processes that only support them in the evaluation of a single idea. It is more likely that they will only engage in extensive search for clusters of ideas, which seem more promising and render the search process more efficient as retrieved knowledge supports the evaluation of several ideas.

Crowd members assess ideas according to their individual preferences and needs. They often prefer niche products which are interesting only for a minority of customers, but address their particular needs best (Ozer 2009; Poetz and Schreier 2012). These individual tastes of the crowd members determine which ideas they engage in, and positively evaluate (Mueller-Trede et al. 2018; Peeters 2018). A study of online forums found that problem statements that stand out and are different from standard problems attract more attention by the forum's users (Haas et al. 2015). Therefore, crowd members will favor more rare, standalone ideas compared to organizational evaluators.

Hypothesis 1b: Rare ideas are more likely favored by the crowd rather than by the organization.

Paradigm Modification

Paradigm modifying ideas change existing patterns by introducing new elements or by altering the relationships between existing elements (Dean et al. 2006). In our context, ideas that change how major systemic components of a car interact, or ideas that introduce substantial new parts or modules are modifying the paradigm of how things have been done before.

Organizational evaluators possess solution knowledge, i.e. knowledge on how to address identified needs by creating products and services (von Hippel 1994). Thus, organizational evaluators are aware of the feasibility of a technology within the current organizational knowledge base and product architecture (Magnusson et al. 2016). Moreau et al. (2001) show that solution knowledge can make it more difficult for experts to assess the risks and potential of novel ideas. When organizational evaluators evaluate novel ideas based on their solution knowledge, this may lead to cognitive fixation in their evaluations. Paradigm modifying ideas challenge the current status of technology and product architecture in an organization. The experts' knowledge about industry standards, internal processes, as well as their individual expertise, will be of limited use in judging those ideas. Being fixated on their solution knowledge, we expect that organizational evaluators discount paradigm modifying ideas instead of questioning their current knowledge base and engaging in distant search.

On the contrary, crowd evaluators have limited solution knowledge and therefore limited insights into how an idea relates to current organizational paradigms (von Hippel 1994). However, crowd evaluators possess need knowledge, "an understanding of how the service creates value for the user" (Magnusson 2009, p. 580). As crowd's evaluations are based on this need knowledge, they favor ideas customized to their need (Ozer 2009; Poetz and Schreier 2012) independent of how many elements of an existing product have to be changed to realize the idea.

Hypothesis 1c: Paradigm modifying ideas are more likely favored by the crowd rather than by the organization.

Organizational Evaluation Criteria and Idea Implementation

The process that organizational evaluators use to assess ideas is often more structured and formal compared to how the crowds evaluates. Organizations intentionally guide their evaluators with specific criteria. However, this may have unwanted consequences. Popular evaluation criteria that organizations apply for the assessment of ideas are company fit, market potential, and creativity. In the following, we argue how these evaluation criteria constrain organizational evaluators' search and evaluation behavior. We argue that crowds, who do not share this constraint, are able to identify some of the ideas that organization's overlook: ideas, which are initially neglected by organizational evaluators, but implemented by the same organization several years later.

Company Fit

The evaluation criterion 'company fit' guides organizational evaluators to focus on the perceived alignment of the idea with the organizational knowledge base (Di Gangi and Wasko 2009). Organizational evaluators using this evaluation criterion are more inclined to search information on the compatibility of the idea with the current competencies and product portfolio (Kim and Wilemon 2002). Ideas with low fit are more difficult to implement and are consequently judged more critically (Baer 2012). As a result, organizational evaluators may initially overlook the potential of ideas they associate with lower company fit to the extent that they do not select them for implementation right away. Crowd evaluators, however, who are not told to use specific formal evaluation criteria, are not influenced by these considerations.

They recognize the potential of these ideas as they are not constrained in their information search and assessment by concerns about how well the idea aligns with the organizational knowledge base.

Hypothesis 2a: Implemented ideas initially evaluated to fit less to the organization are more likely to be favored by the crowd rather than by the organization.

Market Potential

Organizational evaluators are made to take into account if the crowdsourced idea addresses a wide potential customer group if they are to assess the evaluation criterion ‘market potential’ (Franke et al. 2006). In the organizational search for information this will divert attention to considerations such as the idea’s potential profitability or customers’ willingness to buy at a very early stage (Hofstetter et al. 2018; Mueller et al. 2018). Organizational evaluators may neglect ideas during the contest because they attribute low market potential to these ideas. Crowd evaluators are not constrained by these considerations and may favor such ideas although their market potential is not immediately apparent.

Hypothesis 2b: Implemented ideas initially evaluated to have lower market potential are more likely to be favored by the crowd rather than by the organization.

Creativity

It is often unclear how ideas which are assessed as highly creative can be implemented. At first glance, they might seem unrealistic. Creative ideas may not conform to the prevailing organizational standards and the solution knowledge of evaluators. While organizational evaluators are primed to assess to what extent an idea defies their expectations, crowd evaluators may simply perceive them as more interesting and favor them vice versa less creative ideas.

Hypothesis 2c: Implemented ideas initially evaluated to be more creative are more likely to be favored by the crowd rather than by the organization.

Method

Data

We collected data from two external crowdsourcing contests conducted by a leading German automotive manufacturer in 2010. The crowd consists of external participants who have been recruited via an open call on social media platforms by the company. Participants were asked to submit innovative ideas for new products and services. One contest was on “the future of mobility”, the other was about “the interior of the car”. Besides submitting ideas, crowd members also evaluated the ideas on the platform. Like in prior crowdsourcing contests, the crowd used a binary voting mechanism (e.g., Hofstetter et al. 2018) and organizational evaluators assessed more formal evaluation criteria on a rating scale (e.g., Poetz and Schreier 2012).

The crowd members who voted on the ideas are on average 30 years old (S.D. = 9.38, Min. = 16, Max. = 74), 83% own a car, 86% are male, and 87% have completed higher education such as a university degree. Although the contests were conducted by an automotive manufacturer, only 10% of the crowd evaluators are working in the automotive industry. The crowd and the organization have both evaluated all 869 ideas that were submitted. Contrary to most other studies, this allows us to compare the evaluation of all ideas submitted to the contest, not only the limited number of pre-filtered ideas for which an organizational evaluation usually exists. In the two contests, a total of 377 crowd members voted using a binary ‘like/dislike’ button. Binary voting is a common method in crowd evaluation as it has the benefits of being intuitive and easy to understand (Hofstetter et al. 2018; Velamuri et al. 2017). Ten organizational experts rated the ideas on the criteria creativity, market potential, and company fit, and provide their overall assessment on a five-point Likert Scale. This allows us to investigate potential evaluation differences resulting from a restriction to the evaluation behavior of organizational evaluators by the additional evaluation criteria they are asked to pay attention to (cf. H2a-H2c).

Measurement

Dependent Variables

In order to test Hypotheses 1a-c, we measure differences in idea evaluation. We follow the approach from Mollick and Nanda (2016, p. 1543) by assigning ideas to one of four categories: the crowd's favorites (which are favored by the crowd but not the organization), the organization's favorites (which are favored by the organization but not the crowd), and ideas which are either favored by none or both of the two groups of evaluators. The 25% of ideas with the highest evaluation of the respective evaluator group are categorized as "favorites". This approach is preferable to a continuous measure, since we are not equally interested in evaluation differences for ideas which can be rather clearly identified as sub-par. Similarly, organizations usually do not try to rank all ideas but are interested in identifying only the best ideas generated in their contests. Table 1 shows the number of ideas assigned to each of the four categories.

	Bottom 75% Organizational Evaluation	Top 25% Organizational Evaluation	Total
Top 25% Crowd Evaluation	Crowd's Favorites 197	Everyone's Favorites 89	286
Bottom 75% Crowd Evaluation	No one's Favorites 432	Organization's Favorites 151	583
Total	629	240	869

Table 1. Descriptive Overview on the Measure for Idea Evaluation Differences

To test Hypotheses 2a-c, we require a measure that captures whether an idea was implemented by the organization after the contest. We asked six innovation experts employed by the automotive manufacturer to assess if the ideas generated in the contests in 2010 correspond to ideas in their current R&D project pipeline eight years later, or to features already introduced to the market, but not already under development in 2010. In their assessment, experts were allowed to consult specialized internal databases and all other information they had access to. For a limited number of ideas (50, i.e. about 6%) we were able to calculate the interrater reliability of expert assessments, which showed a rather high level of agreement (0.92). As the assessment process was very time consuming, it was not feasible to request more than one expert assessment for the other ideas. Of the 869 ideas, 186 were identified as implemented. We call our dependent variable to test Hypotheses 2a-c 'hidden treasure'. This binary variable has the value 1 for all implemented ideas which were (at the time of the contest) favored by the crowd but not by the organization, i.e. where a 'hidden treasure' to the organization at the time of the contest.

Explanatory Variables

In our study, we include a proxy for each of the three novelty dimensions identified by Dean et al. (2006). The first variable, Originality, is measured by comparing an idea to the organization's strategic thrusts, which are defined in the organizational strategy (the broad areas of autonomous, connected, electrified, and shared products and services). Two independent coders blind to our study assessed if any of the four areas were related to the idea description and discussed the few cases where they initially were in disagreement. We generate the sum of areas matched with an idea and inverted the variable, so that high values denote fewer matches. Empirically, coders matched an idea to either one or two areas (therefore, the minimum value of the variable is $4 - 2 = 2$). The maximum value for originality is 4, where none of the four areas matched with the idea description.

The second dimension of novelty, Rarity, indicates how distant an idea is from the other ideas generated in the contest. Higher values denote ideas with a high average distance to the other ideas, whereas lower values denote ideas that were closer to other submitted ideas. We follow the approach from Piezunka and

Dahlander (2015) to calculate the distance between idea descriptions. First, we process the text by converting all words into lowercase and by removing all punctuation, white spaces, and non-alphabetic characters. Second, we remove all stop words, such as “and” or “to”, which have little informational value for the assessment of an idea’s content. Third, we stem the text by changing each word to its root form (e.g., “driver,” “driving”, or “drive”, all become “drive”). Forth, we transform the texts into vectors of words and compare their cosine similarity. The higher the resulting value, the more similar an idea’s text is to all other idea descriptions, i.e. the less rare the idea is. In order to improve the interpretation of the variable, we subtracted the values from the mean: positive values indicates that an idea is more rare than the average idea, negative value indicate lower than average rarity.

The last dimension of novelty, Paradigm Modification, was assessed by two independent coders blind to our study. The binary variable was coded as 1, if an idea changes more than one major element of a product and/or changes the relationship between elements in a major way. The interrater reliability was 0.87, a third independent coder made the final decision in cases where the two coders did not agree.

The three organizational evaluation criteria were assessed by organizational experts during the contest. Experts were asked to evaluate the ideas on the criteria Company Fit, Market Potential, and Creativity. All criteria were rated on a five-point Likert scale.

Control Variables

In our analyses, we include several control variables for how the ideas were presented, which have often shown to have an influence on how ideas are evaluated. The variable Visual Presentation indicates whether the idea was visualized with a picture or not. Pictures have been shown to attract the attention of the crowd (Li et al. 2016) and positively influence their evaluation (Mollick and Nanda 2016). Further, we include the Length of Description, measured by the number of words used to describe the idea (e.g., Piezunka and Dahlander 2015). We also control for the Number of Comments by participants about an idea, as a proxy for the attention that an idea received (Piezunka and Dahlander 2015). Importantly, we also control for the time an idea was available on the platform as ideas posted later have less time to accumulate “likes”. Days on Platform is measured as the number of days an idea was online. The variable is calculated by subtracting the date of the first activity per idea from the end date of the contest. Further, we control for idea content. The binary variable Physical Product is 1 if the idea is a product and 0 otherwise (e.g., service idea or new function for navigation system). Contest Scope indicates the two different contests. The variable is binary with values of ‘1’ for the broader contest topic (i.e., future of mobility) and ‘0’ for the narrower topic (i.e., interior of the car). As the idea quality and their implementation are dependent on the crowdsourcing problem, we take the contest scope into account (Nickerson and Zenger 2004).

We also control for the language used in the idea submissions. For all language-based variables, we used the software Linguistic Inquiry and Word Count (LIWC), which is based on a validated and widely-applied word-count and dictionary-based method. For the definition and the selection of our language controls, we refer to established measures in prior research (e.g., Mollick and Nanda 2016; Piezunka and Dahlander 2015). Formality of Language is measured by the number of words longer than six letters. Prior studies found that organizational evaluators allocate more attention to formal ideas than crowd evaluators (e.g., Mollick and Nanda 2016; Onarheim and Christensen 2012). In addition, we control for the Complexity of Language measured by the average number of words per sentence (cf. Tausczik and Pennebaker 2010). We control for Emotional Tone (e.g., Mollick and Nanda 2016; Piezunka and Dahlander 2015) and the use of plural pronouns as they indicate identification with the community and the participants’ identification with their ideas, which may influence the evaluation behavior of both audiences (cf. Rouse 2018).

Table 2 shows descriptive statistics and Table 3 the correlation matrix for these variables.

Variable Name	Mean	S.D.	Min.	Max.
Originality	3.75	0.47	2	4
Rarity	0.00	0.01	-0.03	0.02
Paradigm Modification (b)	0.12		0	1
Company Fit	2.08	1.25	1	5
Market Potential	1.97	1.30	1	5
Creativity	2.74	1.12	1	5
Visual Presentation (b)	0.73		0	1
Length of Description	145.20	99.04	3	921
Number of Comments	5.53	3.75	0	30
Days on Platform	60.74	28.67	0.42	99.35
Physical Product (b)	0.85		0	1
Contest Scope (b)	0.31		0	1
Language Formality	21.75	7.33	0.00	77.42
Language Complexity	26.03	20.90	1	180
Emotional Tone	2.00	2.60	-8.08	14.29
First-Person Pronouns	0.70	1.40	0.00	9.80
Plural Pronouns	0.30	0.83	0.00	7.37

Table 2. Descriptive Statistics

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Originality																
2 Rarity	.19															
3 Paradigm Mod. (b)	-.26	-.14														
4 Company Fit	-.02	-.02	-.12													
5 Market Potential	-.15	-.13	.12	-.15												
6 Creativity	-.08	-.04	-.01	.30	.06											
7 Visual Present. (b)	.03	-.04	-.05	.06	.13	.16										
8 Length of Description	-.11	-.31	.10	.08	.10	.15	.10									
9 Numb. of Comments	-.07	-.07	.02	-.01	-.01	.02	.10	-.03								
10 Days on Platform	.29	.12	-.13	.07	-.30	-.21	-.12	-.05	.17							
11 Physical Product (b)	.25	.09	.05	-.10	-.05	-.05	.02	-.03	.03	.40						
12 Contest Scope (b)	-.39	-.17	.20	-.14	.25	.14	.06	.01	.04	-.35	-.44					
13 Language Formality	-.12	.07	-.01	.12	.07	.05	.01	.01	-.09	-.13	-.12	.11				
14 Language Complexity	-.04	-.10	.09	.06	.04	.01	-.01	.14	-.12	.01	-.06	.04	-.04			
15 Emotional Tone	.08	.01	-.04	.03	-.03	-.07	.10	-.05	.03	.14	.09	-.15	.04	-.09		
16 First-Pers. Pronouns	.08	.01	-.06	.01	-.11	-.08	.05	-.02	.05	.17	.06	-.15	-.20	-.09	.13	
17 Plural Pronouns	-.01	-.10	.01	.01	-.02	.05	.01	.05	-.02	.01	.05	.01	-.13	.12	-.02	.05

Table 3. Correlation Matrix

Results

In order to test the hypotheses on evaluation differences (Hypotheses 1a-c), we use a multinomial logit regression analysis (cf. Table 4) in which we compare the crowd's and the organization's favorite 25% of ideas with each other (cf. Table 5). As we are interested in how the organization's favorites differ from crowd's favorites, we use the organization's favorites as reference group. For these analyses we use all 869 ideas that have been submitted during the two contests.

We find support for Hypothesis 1a, as the coefficient for Originality is positive and significant (Table 4, column 1: $b=0.83$, $p<0.05$). This indicates that with increasing originality of an idea, it becomes more likely that the idea will be among the crowd's favorites rather than the organization's favorites. We also find support for Hypothesis 1c: The coefficient for Paradigm Modification is positive and significant (Table 4, column 1: $b=1.11$, $p<0.05$). For ideas with higher values of paradigm modification, it is more likely that they are among the crowd's rather than the organization's favorites. Thus, compared to organizational evaluators, the crowd favors more novel ideas in terms of Originality and Paradigm Modification. We find no support for Hypothesis 1b. Although the coefficient for Rarity is positive as predicted, it is not significant (Table 4, column 1: $b=12.58$, n.s.). Contrary to our hypothesis, the crowd's and the organization's favorites do not differ significantly in terms of how rare an idea is compared to the other ideas generated in a contest.²

Table 5 shows the number of ideas that experts judged as implemented. Out of all 869 ideas, 186 ideas are implemented (a further discussion of this table will follow). We test Hypotheses 2a-c in a logit model (cf. Table 6, column 1) based on only those 186 ideas identified as implemented. The dependent variable is 1 whenever an idea was among the crowd's, but not among the organization's favorites during the contest. These ideas are 'hidden treasures' at the time of the contest: only the crowd, not the organization, recognizes their 'true value' as represented by the organization's later decision to implement these ideas.

Our findings show that the organizational evaluation criterion Company Fit (Table 6, column 1: $b=-2.57$, $p<0.01$) as well as the criterion Market Potential (Table 6, column 1: $b=-0.96$, $p<0.01$) have significant negative coefficients. Both decrease the probability that an idea is a hidden treasure. Thus, we find support for our hypotheses H2a and H2b. As organizational evaluators assess ideas to have lower company fit or lower market potential, these ideas are more likely a crowd, but not an organizational favorite – but are implemented by the organization in the years after the contest. In addition, the coefficient for the organizational evaluation criterion Creativity is positive and significant (Table 6, column 1: $b=1.87$, $p<0.05$). This provides support for our hypothesis H2c: implemented ideas that have been evaluated as more creative are more likely a crowd, but not an organizational favorite, at the time of the contest.

Our analyses reveal several additional insights. In a logit analysis with all 869 ideas in our sample, we observe that all three evaluation criteria increase the probability of an idea to be implemented (Table 6, column 2). Although not hypothesized effects, these relationships may be informative for the interpretation of our findings. At first sight, it might seem counterintuitive that ideas initially judged to have low market potential and company fit are also implemented by the organization. This post-hoc analysis (Table 6, column 2) shows that for the full set of observations, this is – on average – not the case. Having higher scores for market potential and company fit increases an idea's likelihood to be among the ideas that are implemented by the organization. What we argue and find is that due to those evaluation criteria, the organization overlooks some ideas that would have been selected by the crowd: among all implemented ideas ($n=186$), the 57 ideas that are 'hidden treasures' (i.e., are among the crowd's but not the organization's favorites) score lower on company fit and market potential.

An interesting observation in our descriptive statistics concerns the question of how accurate the crowd and the organization ultimately were in predicting the ideas that were implemented. Table 5 classifies all implemented ideas ($n=186$) into the four groups that we used to construct our dependent variables. Out of the 286 ideas that the crowd liked best (cf. Table 5), 109 ideas or 38% were implemented. Out of the best ideas according to organizational evaluators, 47% were implemented. We find this remarkable, given that the decision to implement ideas was made by the organization (in the years after the contest). We believe

² These results are robust to excluding the three organizational evaluation criteria from the regression.

this is a rather conservative indication of the crowds' benefits in idea evaluation. We cannot compare the two evaluator groups with an objective reference value: both groups are measured against how well they predict the organization's implementation decision. There might be more ideas among the crowd's favorites that are worthy of implementation, but which are still misjudged by the organization. Therefore, this descriptive analysis, taken together with the results of our hypotheses tests, corroborates the crowd's usefulness in idea evaluation. Indeed, the highest share of implemented ideas is at the intersection of crowd and organizational favorites (58%). Implementation is most accurately predicted when both evaluator groups agree on the ideas merit.

		Crowd's Favorites (n=197)	Organization's Favorites (n=151)	Everyone's Favorites (n=89)	No one's Favorites (n=432)
Novelty	Originality	0.83** (0.35)		0.67* (0.36)	0.84*** (0.30)
	Rarity	12.58 (17.57)		-7.72 (18.24)	24.03 (15.41)
	Paradigm Modification	1.11** (0.55)		0.52 (0.59)	1.10** (0.50)
Evaluation Criteria	Creativity	-0.97*** (0.18)		-0.16 (0.19)	-1.24*** (0.16)
	Market Potential	-1.13*** (0.15)		-0.10 (0.15)	-1.19*** (0.14)
	Company Fit	-1.87*** (0.18)		-0.15 (0.18)	-1.82*** (0.17)
Controls for Presentation	Visual Presentation	0.17 (0.37)		1.00** (0.46)	-0.41 (0.33)
	Length of Description	-0.003* (0.001)		-0.001 (0.001)	-0.001 (0.001)
	Number of Comments	0.28*** (0.05)		0.26*** (0.05)	0.05 (0.05)
	Days on Platform	0.07*** (0.01)		0.06*** (0.01)	-0.001 (0.01)
Controls for Content	Physical Product	0.63 (0.46)		0.15 (0.45)	0.52 (0.38)
	Contest Scope	3.88*** (0.80)		3.42*** (0.83)	0.18 (0.64)
Controls for Language	Language Formality	-0.06** (0.02)		-0.02 (0.02)	-0.04** (0.02)
	Language Complexity	-0.01* (0.007)		-0.008 (0.008)	-0.001 (0.006)
	Emotional Tone	0.10 (0.06)		-0.03 (0.07)	0.07 (0.06)
	First-Person Pronouns	-0.06 (0.11)		0.01 (0.12)	-0.004 (0.10)
	Plural Pronouns	0.54*** (0.19)		0.40** (0.19)	0.12 (0.19)
	Constant	4.62*** (1.70)		-5.53*** (1.87)	13.48*** (1.49)
	Observations	869	869	869	869

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; Organizational favorites are the reference group.

Table 4. Multinomial Logit Model on Crowd Favorites and Organizational Favorites

	Bottom 75% Organizational Evaluation	Top 25% Organizational Evaluation	Total
Top 25% Crowd Evaluation	57 (29%)	52 (58%)	109 (38%)
Bottom 75% Crowd Evaluation	17 (4%)	60 (40%)	77 (13%)
Total	74 (12%)	112 (47%)	186 (21%)

Note: Percentages are the share of implemented ideas in this category (cf. Table 1, e.g. 60 ideas of 151 organization's favorites are implemented: $60/151 = 40\%$, cf. Table 1).

Table 5. Classification of Implemented Ideas

		Hidden Treasures N = 57	Implemented Ideas N = 186
Novelty	Originality	2.44* (1.27)	0.17 (0.22)
	Rarity	52.62 (39.05)	-13.85 (11.02)
	Paradigm Modification	3.04* (1.56)	-0.14 (0.32)
Evaluation Criteria	Creativity	1.87** (0.80)	0.58*** (0.09)
	Market Potential	-0.96*** (0.37)	0.31*** (0.08)
	Company Fit	-2.57*** (0.53)	0.55*** (0.09)
Controls for Presentation	Visual Presentation	-0.40 (0.96)	0.06 (0.22)
	Length of Description	-0.004 (0.005)	-0.001 (0.001)
	Number of Comments	0.19** (0.08)	0.08*** (0.03)
	Days on Platform	0.07** (0.03)	0.02** (0.008)
Controls for Content	Physical Product	0.62 (1.39)	-0.16 (0.28)
	Contest Scope	3.41* (1.97)	1.10** (0.46)
Controls for Language	Language Formality	-0.08 (0.071)	-0.007 (0.01)
	Language Complexity	-0.05 (0.04)	-0.01** (0.006)
	Emotional Tone	0.03 (0.18)	-0.05 (0.04)
	First-Person Pronouns	-0.09 (0.19)	-0.01 (0.07)
	Plural Pronouns	0.79 (0.51)	0.13 (0.10)
Constant		-1.84 (3.97)	-6.16*** (0.89)
Observations		186	869

Note: Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6. Logit Model on Hidden Treasures and Implemented Ideas

Discussion

Theoretical Implications

Our paper contributes to research on crowdsourcing, specifically on how organizations can enlist the help of crowds for another step in the innovation process: the decision about which crowdsourced ideas to implement. Prior research has focused on how crowds may be biased, mainly due to the ‘packaging’ of ideas, and has started to establish if and in what situations a crowd can adequately take over part of the idea evaluation, i.e. yields similar results to expert evaluators. Instead, we focus on organizational limitations in idea evaluation and whether crowds may be able to complement organizational evaluators. Further, we focus on a key characteristic of the content of crowdsourced ideas, rather than their presentation: an idea’s novelty. The search for more novel ideas is one of the primary reasons why a crowdsourcing contest is started in the first place. We extend prior crowdsourcing research by differentiating between three dimensions of novelty: idea originality, idea rarity, and paradigm modifying ideas. We find support for our hypotheses that compared to organization’s favorites, crowds favor ideas which are more original and more often require paradigmatic changes when implemented. We find no support, however, for differences in the evaluation of rare ideas. In sum, our findings show that crowd evaluations can be an important source of additional information for organizations that set out to search for new ideas.

We also contribute to research on local search, in particular how local search affects the organization’s decisions on idea implementation. A decisive reason why we are able to make these contributions is the data we are able to compile. Due to resource constraints it is uncommon that organizational evaluators assess all submitted ideas. More often, only ideas pre-selected via a positive crowd evaluation are judged by organizational experts. Our dataset includes assessments of all submitted ideas by both evaluator groups. Therefore, we are able to avoid the selection bias and the distortions that come with organizational evaluators assessing only the ideas which are most promising according to the crowd’s assessment. Further, research on crowd evaluations has highlighted time and again that there is a lack of studies that also investigate the implementation of ideas, rather than stopping at their assessment after the contest (e.g., Magnusson et al. 2016). We analyze 869 ideas from two contests of a leading German automotive manufacturer and have evaluations for all submitted ideas by both, crowd and organizational evaluators. Further, innovation experts of the organization assessed if the crowdsourced ideas match with completed or ongoing R&D projects, which were started after the contest.

Therefore, our study is able to go beyond the comparison of evaluation differences due to idea novelty, and can go on to investigate for which ideas the organizational or the crowd assessment correctly predicted later implementation. We compare the evaluations after the contest to the ideas that the organization did de facto decide to implement during the eight years after the contest. For some ideas, the crowd’s evaluation could have helped to correctly identify the potential of ideas that were initially overlooked by the organization. We find that the evaluation criteria that organization’s use are constraints to their search and evaluation behavior. These constraints guide organizational evaluators’ search in specific directions to the extent that they overlook ideas with potential for implementation. In contrast, unconstrained by formal evaluation criteria, the crowd is able to identify some of those ‘hidden treasures’: they select ideas that are not among the organization’s favorites, but are nevertheless implemented later on.

Managerial Implications

Our study provides several toeholds for organizations to improve the process of idea evaluation after crowdsourcing contests. First, by considering the crowd’s favorite ideas, organizations can reduce false negatives (i.e. underestimating and rejecting good ideas). We find that a substantial share of ideas that were later implemented is at first undervalued by organizational evaluators, but is among the favorites of the crowd. However, organizations should not simply leave the evaluation to the crowd entirely. The crowd identifies additional ideas that were later implemented by the organization, but they do not select the same ideas as organizations. Organizational evaluators in our sample showed that their assessment carries weight. The bias in their assessment can possibly be reduced by carefully considering whether they are given formal evaluation criteria and, if so, which ones. It may also be worth considering, if their search

behavior during the evaluation process can be opened up to attenuate the effects of local search. Finally, a multi-stage process might prove valuable. It might be a good practice to screen the crowd's favorites a second time, independent from the first assessment these ideas received by organizational evaluators.

Although we find that crowds sort out and select ideas with a comparable accuracy as organizational evaluators, our results indicate that crowd evaluations should rather complement than replace organizational idea selection. A sequential process where the crowd pre-filters ideas for experts to consider in the second stage is likely to miss valuable ideas that organizational evaluators, but not the crowd recognize. After all, organizational evaluators were most accurate in their assessment (cf. Table 5) and our hypotheses and inferential analyses were concerned with what ideas organizational evaluators might be biased against (i.e., more novel ideas) and what ideas they might overlook (i.e. later change their opinion on). We recommend organizations to choose a selection process based on their main target. If organizations aim to improve the accuracy in idea selection, we recommend that organizational evaluators continue assessing all ideas. Before making their final decision, taking a closer look at the favorites of the crowd might uncover additional ideas that were initially overlooked. However, if organizations are unable or not willing to commit resources to do so, the crowd might perform reasonably well in selecting a smaller set of ideas that is then assessed by organizational evaluators. With this approach, organizations have to consider that they will miss some of the ideas that only organizational evaluators would have detected.

Limitations and Future Research

In the contests we investigate, only the organizational evaluators use formal criteria, the crowd assesses the ideas using a binary “like/dislike” evaluation. Even though this evaluation design is most common in practice and prior research, this limits some of the comparisons we can draw. If crowds would be given formal criteria to assess, their evaluation may also become constrained by them. In addition, while our investigation of idea implementation goes an important step beyond mere evaluation results, we believe that going yet a step further would be exciting. It will be a promising avenue for future research to measure the commercial value of ideas in terms of sales data or profit contribution. In particular, our study found that organizations initially overlook ideas with low ratings for ‘market potential’: do these ideas, if implemented, indeed show lower sales? Similarly, it would be interesting to follow up on ideas initially overlooked due to low ratings of ‘company fit’: are those ideas particularly challenging to implement? Was the initial assessment overestimating implementation difficulties, or should the ideas not have been implemented as implementation efforts ultimately proved unsuccessful or too costly?

Further, our study relies on the assessment of experts to determine whether an idea is implemented in the years after the contest. It is not possible to track ideas over this period of time. Therefore, our study cannot inform on the mechanisms that caused idea implementation. We believe it would provide insightful if we could take a closer look into the ‘black-box’ of what happens to the ideas after a contest. Are teams formed to specify next steps and a plan for implementation of a submitted idea? Are those ideas entering internal product development workshops and merged with each other, or otherwise changed considerably before product development teams work on their implementation? Are they sometimes used as an instrument to lobby for similar ideas that have not yet been able to rally enough support to be implemented? In particular, we would be interested in the reasons why organizations decide to implement ideas which they first undervalued during the crowd contest: when and why did they change their opinion?

In conclusion, we believe that this study contributes to our understanding of the role that crowds can play in the evaluation process. We provide insights into how the crowd's evaluation supports organizations in overcoming limitations in the selection of crowdsourced ideas. Furthermore, our study is one of the first to compare evaluations with actual implementation. Therefore, we are able to pinpoint how local search tendencies in organizational evaluation can be overcome, as some ideas are selected by the crowd but not the organization at the time of the contest, although they get implemented by the organization later on. Crowds may provide valuable assistance to recognize the potential of those ideas early on, right after the contest.

References

- Afuah, A., and Tucci, C. L. 2012. "Crowdsourcing As A Solution To Distant Search," *Academy of Management Review* (37:3), pp. 355–375.
- Baer, M. 2012. "Putting Creativity To Work: The Implementation Of Creative Ideas In Organizations," *Academy of Management Journal* (55:5), pp. 1102–1119.
- Baldwin, C., and von Hippel, E. 2011. "Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation," *Organization Science* (22:6), pp. 1399–1417.
- Behrens, J., Ernst, H., and Shepherd, D. A. 2014. "The Decision To Exploit An R&D Project: Divergent Thinking Across Middle And Senior Managers," *Journal of Product Innovation Management* (31:1), pp. 144–158.
- Criscuolo, P., Dahlander, L., Grohsjean, T., and Salter, A. 2017. "Evaluating Novelty: The Role Of Panels In The Selection Of R&D Projects," *Academy of Management Journal* (60:2), pp. 433–460.
- Dahl, D. W., and Moreau, P. 2002. "The Influence And Value Of Analogical Thinking During New Product Ideation," *Journal of Marketing Research* (39:1), pp. 47–60.
- Dean, D., Hender, J., Rodgers, T., and Santanen, E. 2006. "Identifying Quality, Novel, And Creative Ideas: Constructs And Scales For Idea Evaluation," *Journal of the Association for Information Systems* (7:10), pp. 646–699.
- Di Gangi, P. M., and Wasko, M. 2009. "Steal My Idea! Organizational Adoption of User Innovations From A User Innovation Community: A Case Study of Dell IdeaStorm," *Decision Support Systems* (48:1), pp. 303–312.
- Franke, N., von Hippel, E., and Schreier, M. 2006. "Finding Commercially Attractive User Innovations: A Test Of Lead-user Theory," *Journal of Product Innovation Management* (23:4), pp. 301–315.
- Froehlich, J. K., Hoegl, M., and Gibbert, M. 2016. "Idea Selection In Suggestion Systems: A Thematic Similarity Perspective," *R&D Management* (46:5), pp. 887–899.
- Gambeta, E., Koka, B. R., and Hoskisson, R. E. 2019. "Being Too Good For Your Own Good: A Stakeholder Perspective On The Differential Effect Of Firm-employee Relationships On Innovation Search," *Strategic Management Journal* (40:1), pp. 108–126.
- Haas, M. R., Criscuolo, P., and George, G. 2015. "Which Problems to Solve? Online Knowledge Sharing and Attention Allocation in Organizations," *Academy of Management Journal* (58:3), pp. 680–711.
- Hienert, C., and Riar, F. 2013. "The Wisdom Of The Crowd Vs. Expert Evaluation: A Conceptualization Of Evaluation Validity.," Paper presented at DRUID, Spain, Barcelona.
- Hofstetter, R., Aryobsei, S., and Herrmann, A. 2018. "Should You Really Produce What Consumers Like Online? Empirical Evidence For Reciprocal Voting In Open Innovation contests," *Journal of Product Innovation Management* (35:2), pp. 209–229.
- Howe, J. 2006. "The Rise of Crowdsourcing," *Wired magazine* (14:6), pp. 1–4.
- Illi, S., Albers, A., and Miller, S. 2010. "Open innovation in the automotive industry," *R&D Management* (40:3).
- Jeppesen, L. B., and Lakhani, K. R. 2010. "Marginality And Problem-solving Effectiveness In Broadcast Search," *Organization Science* (21:5), pp. 1016–1033.
- Katila, R., and Ahuja, G. 2002. "Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction," *Academy of Management Journal* (45:6), pp. 1183–1194.
- Kim, J., and Wilemon, D. 2002. "Focusing the Fuzzy Front-End in New Product Development," *R and D Management* (32:4), pp. 269–279.
- Kornish, L. J., and Ulrich, K. T. 2011. "Opportunity Spaces In Innovation: Empirical Analysis Of Large Samples Of Ideas," *Management Science* (57:1), pp. 107–128.
- Lakhani, K. R., Lifshitz-Assaf, H., and Tushman, M. L. 2013. "Open innovation and organizational boundaries: Task decomposition, knowledge distribution and the locus of innovation," in *Handbook of economic organization: Integrating economic and organization theory*, A. Grandori (ed.), Northampton, MA: Edward Elgar, pp. 355–382.
- Levinthal, D., and March, J. G. 1981. "A model of adaptive organizational search," *Journal of Economic Behavior & Organization* (2:4), pp. 307–333.
- Li, M., Kankanhalli, A., and Kim, S. H. 2016. "Which Ideas Are More Likely To Be Implemented In Online User Innovation Communities? An Empirical Analysis," *Decision Support Systems* (84), pp. 28–40.
- Lopez-Vega, H., Tell, F., and Vanhaverbeke, W. 2016. "Where And How To Search? Search Paths In Open Innovation," *Research Policy* (45:1), pp. 125–136.

- Magnusson, P. R. 2009. "Exploring The Contributions Of Involving Ordinary Users In Ideation Of Technology-based Services," *Journal of Product Innovation Management* (26:5), pp. 578–593.
- Magnusson, P. R., Waestlund, E., and Netz, J. 2016. "Exploring Users' Appropriateness as a Proxy for Experts when Screening New Product/Service Ideas," *Journal of Product Innovation Management* (33:1), pp. 4–18.
- Mollick, E., and Nanda, R. 2016. "Wisdom Or Madness? Comparing Crowds With Expert Evaluation In Funding The Arts," *Management Science* (62:6), pp. 1533–1553.
- Moreau, C. P., Lehmann, D. R., and Markman, A. B. 2001. "Entrenched Knowledge Structures And Consumer Response To New Products," *Journal of Marketing Research* (38:1), pp. 14–29.
- Mueller, J., Melwani, S., Loewenstein, J., and Deal, J. J. 2018. "Reframing The Decision-makers' Dilemma: Towards A Social Context Model Of Creative Idea Recognition," *Academy of Management Journal* (61:1), pp. 94–110.
- Mueller-Trede, J., Choshen-Hillel, S., Barneron, M., and Yaniv, I. 2018. "The Wisdom Of Crowds In Matters Of Taste," *Management Science* (64:4), pp. 1779–1803.
- Nickerson, J. A., and Zenger, T. R. 2004. "A Knowledge-based Theory Of The Firm—the Problem-solving Perspective," *Organization Science* (15:6), pp. 617–632.
- Onarheim, B., and Christensen, B. T. 2012. "Distributed Idea Screening In Stage–gate Development Processes," *Journal of Engineering Design* (23:9), pp. 660–673.
- Ozer, M. 2009. "The Roles Of Product Lead-users And Product Experts In New Product Evaluation," *Research Policy* (38:8), pp. 1340–1349.
- Peeters, T. 2018. "Testing The Wisdom Of Crowds In The Field: Transfermarkt Valuations And International Soccer Results," *International Journal of Forecasting* (34:1), pp. 17–29.
- Piezunka, H., and Dahlander, L. 2015. "Distant Search, Narrow Attention: How Crowding Alters Organizations' Filtering Of Suggestions In Crowdsourcing," *Academy of Management Journal* (58:3), pp. 856–880.
- Poetz, M. K., and 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), pp. 245–256.
- Reichwald, R., and Piller, F. T. 2009. *Interaktive Wertschoepfung: Open Innovation, Individualisierung und neue Formen der Arbeitsteilung*, Wiesbaden: Gabler Verlag / GWV Fachverlage GmbH Wiesbaden.
- Rosenkopf, L., and Nerkar, A. 2001. "Beyond Local Search: Boundary-spanning, Exploration, And Impact In The Optical Disk Industry," *Strategic Management Journal* (22:4), pp. 287–306.
- Rouse, E. 2018. "Where you end and I begin: Understanding intimate co- creation," *Academy of Management Review* (Forthcoming).
- Seidel, V. P., and Langner, B. 2015. "Using an online community for vehicle design: project variety and motivations to participate," *Industrial and Corporate Change* (24:3), pp. 635–653.
- Simmons, J. P., Nelson, L. D., Galak, J., and Frederick, S. 2011. "Intuitive Biases In Choice Versus Estimation: Implications For The Wisdom Of Crowds," *Journal of Consumer Research* (38:1), pp. 1–15.
- Sorensen, J. B., and Stuart, T. E. 2000. "Aging, Obsolescence, And Organizational Innovation," *Administrative Science Quarterly* (45:1), p. 81.
- Stuart, T. E., and Podolny, J. M. 1996. "Local Search And The Evolution Of Technological Capabilities," *Strategic Management Journal* (17:S1), pp. 21–38.
- Tausczik, Y. R., and Pennebaker, J. W. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology* (29:1), pp. 24–54.
- Velamuri, V. K., Schneckenberg, D., Haller, J. B. A., and Moeslein, K. M. 2017. "Open Evaluation Of New Product Concepts At The Front End Of Innovation: Objectives And Contingency Factors," *R&D Management* (47:4), pp. 501–521.
- von Hippel, E. 1978. "A Customer-active Paradigm For Industrial Product Idea Generation," *Research Policy* (7:3), pp. 240–266.
- von Hippel, E. 1994. "'Sticky Information' And The Locus Of Problem Solving: Implications For Innovation," *Management Science* (40:4), pp. 429–439.