MIS Quarterly

RESEARCH NOTE

HOW INFORMATION CONTRIBUTED AFTER AN IDEA SHAPES NEW HIGH-QUALITY IDEAS IN ONLINE IDEATION CONTESTS[1]

Aron Lindberg
Stevens Institute of Technology, 1 Castle Point Terrace,
Hoboken, NJ 07030 U.S.A. {aron.lindberg@stevens.edu}

Ann Majchrzak
Marshall School of Business, University of Southern California, 3670 Trousdale Parkway,
Los Angeles, CA 900898 U.S.A. {majchrza@usc.edu}

Arvind Malhotra
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, 300 Kenan Center Drive,
Chapel Hill, NC 27599 U.S.A. {arvind_malhotra@kenan-flagler.unc.edu}

Findings on how prior high-quality ideas affect the quality of subsequent ideas in online ideation contests have been mixed. Some studies find that high-quality ideas lead to subsequent high-quality ideas, while others find the opposite. Based on computationally intensive exploratory research, utilizing theory on blending of mental spaces, we suggest that the effects of prior ideas on the generation of subsequent ideas depend on the alignment of (1) crowd participants' subjective quality assessments of prior ideas and (2) subsequent problem-related contributions made by the crowd. When a prior idea is assessed as high-quality, this motivates the crowd to emulate that idea. When this motivation is aligned with subsequent contributions that expand the mental space of the prior idea, a new high-quality idea can be created. In contrast, when a prior idea is assessed as low-quality, it motivates the crowd to redirect away from that idea. When this motivation is aligned with subsequent contributions that shift the mental space of the prior idea, a new high-quality idea can be created. The mixed findings in the literature can then be explained by a failure to consider non-idea information contributions made by the crowd.

Keywords: Online ideation contests, crowds, idea generation, group creativity, mental spaces, blending, innovation, open innovation

Introduction

As part of open innovation (Chesbrough, 2006), firms are using online ideation contests to draw on the "wisdom of crowds" (Bayus, 2013; Terwiesch & Xu, 2008). Scholars studying online ideation contests are paying increasing attention to how crowd participants influence each other's ideation to create high-quality ideas (Bayus, 2013; Füller et al., 2011; Majchrzak & Malhotra, 2016). High-quality ideas can be defined as "fresh ideas for changing products, services, and processes to better achieve the organization's goals" (Amabile et al., 2005, p. 367). These ideas are not only novel but implementable and key to competitive advantage.

The literature on crowdsourcing of ideas has yielded mixed results: some find that crowds posting high-quality ideas create a positive feedback cycle for more high-quality ideas (Kohn et al., 2011), while others find the opposite, that prior high-quality ideas beget subsequent low-quality ideas (Bayus, 2013; Liu et al., 2014; Wang et al., 2018). Moreover, crowds

[1]Brian Butler was the accepting senior editor for this paper. Eivor Oborn served as the associate editor.

do not always post high-quality ideas (Füller et al., 2011). In such a situation, analogous to online brainstorming (Dennis & Williams, 2003; Seeber et al., 2017), low-quality ideas may create mimetic pressure, producing more low-quality ideas. Hence, it is not clear how the quality of the ideas offered by crowds affects the quality of subsequent ideas.

Additionally, in many online ideation contests, the crowd not only contributes ideas but may also contribute additional information about the problem posed to them (Majchrzak & Malhotra, 2020). Before making a contribution, a crowd participant reads the information that has already been contributed by others and recorded in digital traces. In the course of doing so, the crowd participant has an internal conversation and reflects on the information in preparation for a new contribution (Baralou & Tsoukas, 2015). What is not known is whether the temporality and content of information contributed between ideas impact the relationship between the quality of a prior idea and the subsequent idea provided by the crowd. Therefore, we ask the following research question:

How does the information contributed after a prior idea affect the quality of the subsequent idea?

We address this question conceptually by drawing on the creative cognition literature (Finke et al., 1996), specifically Fauconnier and Turner's (1998, 2003) theory on blending of mental spaces, which we adapt to the context of online ideation contests. In particular, we suggest that ideas represent distinct mental spaces comprised of concepts and relationships between concepts. Therefore, when an idea has been posted, it is an explicit expression of a mental space. This idea is then followed by additional contributions made by crowd participants. Many of these contributions include additional information related to the problem posed to the crowd, which may modify the mental space of the prior idea. These contributions may then encourage a crowd participant who reads the prior idea and subsequent contributions to expend cognitive effort in blending their mental space with the modified mental space of the prior idea. Such blending can result in a subsequent idea representing a new, "blended" mental space.

We conducted a computationally intensive data-driven theory development study (Berente et al., 2019) that utilized log data from online ideation contests and variable-length Markov chains for sequence analysis (VLMC, see Gabadinho & Ritschard, 2016; Mächler & Bühlmann, 2004). We find that a subsequent high-quality idea is affected by whether assessments of the quality of a prior idea are aligned with subsequent information contributions about the problem statement that either *shifts* or *expands* the prior idea's mental space. Such shifting or expansion then encourages a subsequent crowd participant reading the prior idea and subsequent contributions to expend additional cognitive effort to blend a mental space of their own with the modified mental space associated with the prior idea. This theorizing suggests that prior research may not have found consistent effects of the quality of a prior idea on crowd ideation because it failed to take into account the type of information about the problem that the crowd offers subsequent to the prior idea. We also outline how our methodological approach may stimulate novel empirical research utilizing sequence analysis.

Conceptual Development

Mental spaces (Fauconnier & Turner, 1998, 2003) are constituted by concepts and their relationships and can be altered through discourse "for the purpose of local understanding" (Hart, 2007, p. 117). As such, mental spaces are situational and context-specific; they concern a particular ongoing conversation or situation experienced by an individual, exist only temporarily, and are continuously modified (Oakley & Hougaard, 2008, p. 4).

The concept "space" has often been conceived of in social terms (Kellogg, 2009; Latané & Liu, 1996; Rao & Dutta, 2012) and has been helpful in explaining how human practices are relationally situated in particular contexts and social constructions. For example, Latané and Liu (1996) indicate that social space is an "intersubjective matrix of psychological distances" that helps to explain social impact and clustering of attitudes, values, and identities (p. 26). Further, space has been used to conceptualize search and problem solving processes (Newell & Simon, 1972). Such spaces contain all possible combinations of various parameters. Mental spaces (Fauconnier & Turner, 1998, 2003), however, contain specific sets of concepts and their relationships (i.e., they contain a limited number of particular concepts, not all possible combinations) and are focused on the micro-foundations of cognition, as opposed to explaining social relationships and practice.

Multiple mental spaces can be "blended" (Fauconnier & Turner, 1998, 2003) together to create a new mental space, referred to as a "blended space," which contains a novel set of concepts and relationships. According to Turner (2014), blended mental spaces represent ideas. Blending occurs through selectively projecting some concepts from the "input" mental spaces to the blend, and then relating these concepts in a way that is different from how concepts were related in the input spaces (Fauconnier & Turner, 1998, 2003).

Blending in Online Ideation Contests

Blending theory allows us to view crowd ideation from a cognitive perspective. As blending theory conceptualizes ideas as mental spaces, this theory enables us to explain how crowd participants may engage in ideation by blending mental spaces pertaining to prior ideas with mental spaces of their own to create new ideas. For example, in an ideation contest for a toy company, a mental space may contain multiple concepts germane to dolls, such as "plastic body," "look," "coloring," "doll clothes," and "gender appeal." Some or all of these concepts from the mental space of the idea could then be blended with other mental spaces to create subsequent ideas. For example, the mental space pertaining to "dolls" could be blended with a mental space pertaining to "personalized manufacturing technologies" leading to a new idea of dolls personalized to a child's specifications.

Blending theory does not specifically explain how prior ideas of varying quality can be modified through discourse to yield further high-quality ideas. Therefore, we combine blending theory with the broader literature on creative cognition (e.g., Finke et al., 1996) to explain (1) how ideas may vary in the degree to which they capture the requirements of the initiating problem statement (what we refer to as quality), and (2) how mental spaces can be modified by problem-related information contributed after a prior idea, which stimulates cognitive effort toward blending. Our representation of crowd ideation as modification and blending of mental spaces allows us to account for how prior ideas may affect subsequent ideas, as we explain below.

The intent of an online ideation contest is for crowd participants to generate ideas that meet the requirements of a problem statement in terms of being both novel and offering competitive advantage. Since ideas represent mental spaces that can serve as inputs to subsequent blended mental spaces, the degree to which the prior idea meets the problem requirements is likely to also serve as input to subsequent blended mental spaces. While ideas generated in online ideation contests generally are evaluated by sponsors after the contest has concluded (Dahlander et al., 2019), crowd participants will also evaluate ideas extemporaneously (Bayus, 2013; Kohn et al., 2011; Liu et al., 2014; Wang et al., 2018). Therefore, as a crowd participant examines an idea, their subjective assessment of this idea may influence their behavior in a manner similar to what online brainstorming has found: that prior ideas stimulate new ideas (Dennis et al., 1996; Seeber et al., 2017). Blending theory, however, also suggests that a blended mental space can be modified by additional information providing new concepts and/or relationships. Since, in some ideation contests, information about the problem rather than just solutions can be provided by crowd participants, it is possible that this additional information coupled with the subjective assessments of the quality of prior ideas may affect the quality of new ideas produced.

Problem-Related Information Contributions

The creative cognition literature has identified various information contributions that modify mental spaces in ways that increase cognitive effort toward generating creative outcomes (Paulus & Nijstad, 2003). This increased cognitive effort may help to explain the type of information that needs to be contributed for subsequent ideas to be of high quality. Four of the most commonly described contributions likely to modify existing mental spaces include assumptions that can add new concepts to an existing mental space, analogies that can shift a mental space, questions that can lead to productive inquiry into a mental space, and paradoxes that examine tensions in relationships between concepts within a mental space (Cronin & Weingart, 2007; Girotra et al., 2010; Leonard-Barton & Swap, 2005; Paulus & Nijstad, 2003).

Assumptions refer to information such as new facts, observations, or beliefs (Dane, 2010; Gentner & Stevens, 2014). They provide concrete backing or validation for an idea (Tsoukas, 2009) and help to stimulate creativity (Hender et al., 2001; Santanen et al., 2004). Hence, assumptions modify an already existing mental space by adding new concepts to it, which signals that any subsequent ideation efforts need to expend cognitive effort to consider these additional concepts as well.

Analogies refer to information about a different context (Ansburg & Hill, 2003; Choi & Kim, 2017; Nijstad et al., 2003) compared to the context of the mental space pertaining to a prior idea. For example, when trying to generate ideas about how to best design a car, one can use an analogy from the zoological context, for example, drawing inspiration from the physical shape of jaguars to stimulate the emergence of high-quality ideas. Analogies may, therefore, shift the prior mental space, signaling that cognitive effort needs to be expended to consider this new context (Fiske & Taylor, 1984; Hargadon & Bechky, 2006) in addition to the concepts and relationships in the mental space pertaining to the prior idea.

Questions are inquiries into the mental space pertaining to the prior idea (Baralou & Tsoukas, 2015; Boland & Tenkasi, 1995; Garrison, 2015). Therefore, questions modify the mental space pertaining to a prior idea by making definitions of concepts and their relationships clearer. This signals to the crowd that subsequent cognitive efforts to generate further ideas need to consider such inquiries into the mental space pertaining to the prior idea.

Paradoxes provide information on tension or conflict between different concepts. Paradoxes, therefore, identify interrelated yet contradictory requirements that coexist (Hargrave & Van de Ven, 2017; Jay, 2013), for example, trade-offs between multiple objectives (Cronin & Weingart, 2007). Since paradoxes involve contradiction and tension, paradoxes can foster new linkages within a mental space by forcing re-evaluation of a problem. Paradoxes may expose a new way of seeing a problem and, therefore, also its solution (Miron-Spektor et al., 2011; Poole & Ven, 1989). Hence, identifying a paradox forces cognitive effort to resolve the paradox during the course of generating a subsequent idea through blending.

We know little, however, about how the different types of problem-related information contributions need to be aligned with prior ideas of either low or high quality to contribute toward generating subsequent high-quality ideas. Each of these contributions modifies the mental space pertaining to the prior idea in different ways and, therefore, also stimulates cognitive effort differently on the part of the crowd participant who posts the next idea. It may be that high-quality ideas emerge from some form of an alignment between the quality of the prior idea and the type of information contribution made subsequently.

Method

The literature on online ideation contests says relatively little about the sequencing of ideas and problem-related contributions, therefore we sought to generate new theory. To do this, we choose a computationally intensive data-driven exploratory research approach (Berente et al., 2019; Lindberg 2020; Shrestha et al., 2021) enabling us to draw on digital trace data and computational methods specifically geared toward analyzing sequences. Below we describe the dataset, the measures of idea quality, measures of problem-related information contributions, and our usage of VLMC (Gabadinho & Ritschard, 2016; Mächler & Bühlmann, 2004).

Dataset

The dataset consists of 20 different online ideation contests, each with a different problem statement sponsored by either a company from the U.S., Scandinavia, or China operating in manufacturing or service sectors (e.g., software, hardware, financial services, entertainment, sports, and industrial products). Each contest had between 48 and 179 unique participants. This dataset is unique in being one of the few datasets of crowd-based ideation on strategic, ill-structured problems where the crowd was encouraged to share unstructured information about the problem statements in addition to their ideas. The design of the ideation contests is described in Majchrzak & Malhotra (2020).

Of particular note here is that the ideation contests had problem statements sufficiently similar to be aggregated in the analysis. Each problem statement was (1) ill-structured, requiring multiple perspectives such that solutions would benefit from conversations and creative input with no single correct answer, (2) important to the executive in that it affected processes, products, or services that the executive thought were critical to the organization's future, and (3) one which the executive had personally tried to solve, unsuccessfully, so that the executive would be able to assess the quality of proposed ideas.

The 20 ideation contests followed similar protocols. The contests all lasted about the same amount of time (7–10 days). All participants were allowed to contribute anonymously or quasi-anonymously (i.e., they could choose a screen name other than their real name). All ideas and problem-related information were contributed through discussion forums. The following protocol for incentivizing participation was used: small in-kind brand or impact awards were provided for (1) top ideas as judged by sponsors and (2) posting on others' posts (Hutter et al., 2011; Zheng et al., 2011). A total of 1,982 posts were obtained including 1,176 ideas.

The quality of an idea was measured by ratings conducted by the executive with regard to the novelty of the idea and the competitive advantage of the idea if it were to be implemented. The ideas were rated as high- or low-quality, where high-quality was defined as an idea being *both* novel and offering a competitive advantage. All other ideas were rated as low-quality.

To reduce the cognitive load of the executives, we followed the advice of Lamastra (2009). Only the first version of non-duplicative ideas was evaluated by the executives. This led to executives rating between 3 and 21 unique ideas. Across the 20 contests, there were 297 non-duplicative ideas for which assessments about the quality of the ideas were obtained from the executives. Independent raters were trained based on the assessments of the executives, and then they rated the remaining ideas, obtaining an inter-rater reliability for categorical data of 77% (Cohen's Kappa Coefficient $\kappa = 0.77$; $p < 0.001$; see Landis & Koch, 1977).[2] The dataset also

[2]When multiple ideas occurred directly adjacent to each other, they were considered together, as a "set" of ideas. Because of their proximity, any of these ideas may have been generated by previously contributed posts. In order to discern the effect that previous posts may have on any idea inside of an idea set, we recoded such adjacent sequences using a singular marker.

included 806 posts that were not ideas.[3] These posts were coded to identify if they contained any of the four problem-related information contributions defined above: assumptions, analogies, questions, or paradoxes. Two independent raters scored an inter-rater agreement within the acceptable range (Cohen's Kappa Coefficient $\kappa = 0.74$; $p < 0.001$). Any disagreements were resolved through discussion. Posts not containing any of these were ignored, leading to removal of 17% of the posts. Such posts commonly contained vague statements such as "OK," "Yes," or "Agree."

Overview of Analysis Strategy

To examine the effect of prior idea quality and subsequent problem-related information contributions on the quality of a subsequent idea, we analyzed sequences of ideas and contributions. Amongst several alternative analytic methods for sequence analysis, we choose VLMC (Gabadinho & Ritschard, 2016; Mächler & Bühlmann, 2004). While traditional Markov chains, as well as hidden Markov models, are "memoryless" in the sense that they predict future events based on *only* the current event (see, for example, Singh et al., 2011), VLMC has a memory of variable length, thus allowing for a sequence of prior events to be used to predict the next event. This means that when predicting high- and low-quality ideas, we can examine a variable length of prior sequences—consecutively ordered sets of ideas and problem-related information contributions—to predict the quality of a subsequent idea.

VLMC has several advantages compared to other techniques that may be used to analyze sequences of ideas and problem-related information contributions. First, in contrast to optimal matching techniques (Abbott & Tsay, 2000; Gaskin et al., 2014), VLMC does not make *a priori* assumptions with regard to where and when relevant sequences occur (and therefore when a sequence starts) within a larger flow of posts. Instead, VLMC can inductively identify repeated sequences using computational means. Second, VLMC also does not assume that sequences are independent of each other, as a regression-based research design would (see Majchrzak & Malhotra, 2016), but rather allows for overlapping sequences (Gabadinho & Ritschard, 2016). Third, and last, VLMC recognizes the recursive nature of sequences. This means that sequences related to prior ideas can also lead to subsequent ideas.

VLMC has been used to analyze protein classifications and English language texts (Ron et al., 1996), music pieces (Begleiter et al., 2004), and DNA sequences (Bühlmann & Wyner, 1999; Mächler & Bühlmann, 2004). Most recently, in the IS literature, Wikipedia edits (Arazy, Lindberg, Lev, et al., 2020) were analyzed using VLMC, revealing the presence of "emergent routines" (i.e., stabilized patterns of editing wiki articles emerging in a bottom-up fashion, rather than being centrally coordinated). We use VLMC to achieve "analytical generalization" from certain empirical statements to theory (Lee & Baskerville, 2003; Shrestha et al., 2021), rather than generalization to a population.

Using the VLMC approach, a sequence was defined as a consecutively ordered set of ideas and problem-related information contributions. Hence, we specified as a condition that any sequences identified had to contain prior ideas. A major assumption that the VLMC method makes is that the identified sequences consist of directly adjacent (i.e., occurring right after each other in time, with no intermediate contributions in between) contributions and that these contributions directly precede subsequent ideas.

To identify repeated sequences of ideas and problem-related information contributions frequently preceding high- versus low-quality ideas, VLMC constructs and prunes probabilistic suffix trees (Gabadinho & Ritschard, 2016) containing all those sequences that have a non-zero probability of preceding an idea of either low or high quality. The lengths of the sequences are not specified *a priori* but rather are computationally determined. The probability distributions yielded by this analysis then lead to varying conditional probabilities for the relationship of each sequence to high- and low-quality ideas.

While VLMC generates sequences, it does not explicitly specify the statistical relationships between such sequences and outcomes of interest (i.e., low- and high-quality ideas). Hence, to arrive at patterns that have theoretical value, further analysis was required. Based on the VLMC results, we computed descriptive statistics, including the frequency of each sequence (i.e., how often that particular sequence of ideas and problem-related information contributions was observed in the data), the conditional probabilities of a high- or low-quality idea given a specific sequence, and the lift of each type of

Hence, if a set of ideas contained a high-quality idea, the whole set was coded as high-quality; if such a set did not contain a high-quality idea, the whole set was coded as low-quality. This enabled us to identify high-quality ideas embedded immediately next to several low-quality ideas. Our dataset included a total of 134 high-quality idea sets and 255 low-quality idea sets.

[3]In coding the posts, we used a coding scheme that constrained the post to be of one category type. Hence, if the dominant content type in the post was an assumption, then we categorized it as an assumption even though a small portion of the text may have been used to describe an idea. This strict single-code categorization scheme resulted in 22 posts not being considered as containing ideas, but rather being categorized strictly as non-idea contributions.

sequence preceding a high- or low-quality idea (Lenca et al., 2008).

The lift is the conditional probability of generating a high- or low-quality idea, given a specific sequence, divided by the unconditional probability of a high- or low-quality idea. This helped us to interpret the conditional probabilities in light of how common high- and low-quality ideas were in the dataset overall. We used the lift calculations to compare whether a sequence is statistically more likely to precede a high- or a low-quality idea, thus providing additional validation of the VLMC results. Through calculating the Z-value of the difference between the lift for a specific sequence leading to a high-quality versus a low-quality idea (Lenca et al., 2008, p. 619), we determined which sequences are significantly more likely to precede a high-quality idea rather than a low-quality idea, or vice versa. Sequences for which the difference in lifts was not statistically significant were not considered (i.e., where the Z-score of the difference was between -1.645 and 1.645, indicating a lack of significance at the 0.05 level).[4]

To understand what sequences generated high-quality ideas, but not low-quality ideas, and vice versa, we examined sequences preceding ideas of different quality, seeking to identify systematic differences. To do this, we examined the sequences with significant lift statistics differentiating between those generating either high or low-quality ideas. Since multiple sequences were significant, we looked for patterns among the sequences. The patterns were iteratively compared to the creative cognition literature (e.g., Fauconnier & Turner, 1998, 2003) to seek theoretical explanations for why a specific sequence would precede a particular outcome. Through such an iterative theoretical justification process, we generated a classification of the sequences. Our approach, consisting of induction from data and deduction from literature, was therefore abductive (Holland et al., 1989; Lindberg, 2020).

Findings

The analysis described above led to the identification of 10 types of sequences for which the lift scores relative to low- versus high-quality ideas were significantly different. Table 1 shows the 10 types of sequences, which collectively occurred 271 times. Six types of sequences (N = 123) occurred before high-quality but not low-quality ideas, and four types of sequences (N = 148) occurred before low-quality but not high-quality ideas. To read these results, one looks across the rows. For example, the first sequence, HQ-QU, occurs 39 times in the dataset, has a 0.35 conditional probability of generating a high-quality idea and a 0.20 conditional probability of generating a low-quality idea. This translates into a lift relative to high-quality ideas of 3.09, and a lift relative to low-quality ideas of 0.95. The Z-score of the difference between these two lift values is 4.05, which is significant at the 0.001 level.

Of the 10 sequence types, 4 types (N = 27) shared significant commonality between high- and low-quality ideas (HQ-QU-AS and AN-HQ-AS before low-quality ideas; HQ-AS-LQ-QU and QU-HQ-AS before high-quality ideas), thus failing to differentiate between outcomes. These sequences were dropped from further interpretation. The remaining six types of differentiating sequences (N = 244) were categorized by similarity. As shown in Table 2, we have labeled these sequences based on how problem-based information contributions modify the mental space pertaining to the prior idea: shifting prior ideas through analogies, expanding prior ideas through questions, and expanding prior ideas through paradoxes.

Shifting prior ideas through analogies. This sequence category encompasses two variations: AN-AS-LQ-AN (N = 3) and LQ-AN (N = 49) representing 52 sequences out of the total of 244 differentiating sequences. Both of these sequence variations include a low-quality idea followed by an analogy, which then is followed by a high-quality idea. The prior low-quality idea may be neither novel nor offer competitive advantage, and when assessed as such, crowd participants form a motivation to redirect away from the prior idea (Fiske & Taylor, 1984; Hargadon & Bechky, 2006). This motivation leads to the contribution of an analogy, thus shifting the mental space of the prior idea in such a way that subsequent cognitive efforts to see the problem statement in a new way are enabled. The crowd participant reading the prior idea and subsequent contributions then blends the modified mental space of the prior idea with a mental space of their own, which then yields a high-quality idea.

In the example below, a prior low-quality idea with regard to incentive programs and their relationship to success in the shipping industry is offered, which is then followed by an analogy that shifts the mental space of the prior idea toward a focus on digital technologies. As a crowd participant reads the prior idea and subsequent analogy, they blend the shifted mental space of the prior idea with a mental space of their own, yielding a high-quality idea (i.e., RFID to provide better package tracking).

[4]Sequences for which lifts in relation to both high- and low-quality ideas were significant were also excluded because they did not differentiate between high- and low-quality ideas, which occurred for two sequences.

Table 1. Sequence Statistics

Sequence Type	# of Sequences of this Type	Conditional Probabilities Relative to HQ	Conditional Probabilities Relative to LQ	Lifts Relative to HQ	Lifts Relative to LQ	Difference in Z-score*
Significant Sequences Preceding High-Quality Ideas						
HQ-QU	39	0.35	0.20	3.09	0.95	4.05
HQ-PA	17	0.45	0.29	4.02	1.34	3.39
HQ-AS-LQ-QU	5	0.39	0.01	3.43	0.05	2.58
QU-HQ-AS	10	0.29	0.01	2.60	0.05	2.43
AN-AS-LQ-AN	3	0.64	0.01	5.66	0.05	1.95
LQ-AN	49	0.20	0.20	1.82	0.95	1.83
Significant Sequences Preceding Low-Quality Ideas						
HQ-QU-AS	3	0.01	0.95	0.09	4.44	-2.33
LQ-PA	38	0.08	0.42	0.70	1.97	-2.37
LQ-QU	98	0.09	0.39	0.82	1.81	-3.03
AN-HQ-AS	9	0.01	0.74	0.09	3.46	-3.55

HQ = High-Quality Idea, LQ = Low-Quality Idea, PA = Paradox, AN = Analogy, AS = Assumption, QU = Question. Sequence types marked in bold are those which were theoretically classified to distinguish between the generation of high- and low-quality ideas.

*The Z-score indicates the difference between lifts relative to high-quality ideas vis-à-vis low-quality ideas, thereby indicating whether a certain sequence is significantly more likely to be followed by high- or low-quality ideas.

Table 2. Categorization of Sequences

Category	Sequences Preceding High-Quality Ideas (N = 108)	Sequences Preceding Low-Quality Ideas (N = 136)
1. Shifting prior ideas through analogies	Shifting prior low-quality ideas (AN-AS-LQ-AN, N = 3; LQ-AN, N = 49)	
2. Expanding prior ideas through questions	Inquiring into prior high-quality ideas (HQ-QU, N = 39)	Inquiring into prior low-quality ideas (LQ-QU, N = 98)
3. Expanding prior ideas through paradoxes	Examining tensions amongst concepts related to prior high-quality ideas (HQ-PA, N = 17)	Examining tensions amongst concepts related prior low-quality ideas (LQ-PA, N = 38)

Incentive programs are always good for both the small business and the shipping service business. I guess an example would be if you ship 50+ items, get 10–20% off your total shipping costs. If the job is done well, it's a good thing. On the flipside, the courier service may be offering an incentive program to regain credibility for some debacle that could've happened on a previous shipment. I don't think this is likely, but it's just a wild thought. (Low-Quality Idea)

The best services I've ever used was [ShippingCo], and I only discovered them after my second cross continent move. You fill out your application online, the deliver your boxes, with tape/markers everything you need to box up. They come back 24/48 hours later to pick up and ship off. This company is also very inexpensive compared to other companies and local post offices. When your boxes arrive at their destination they personally deliver them to the address chosen by you. (all inclusive in their low cost) PLUS I received them in tip top condition. (Analogy)

What about using RFID chips that would talk to the delivery plane and/or truck's GPS device, for real-time, anytime, en-route tracking. So technically you'd be tracking the plane/truck, not the package (you'd scan the RFID chip on the package upon placing it on the vehicle, and the vehicle's move-

ment can be tracked). I think this would be a huge opportunity, because I would think one of the bigger concerns with international shipments is ensuring your package is safely and indeed on its way. (High-Quality Idea)

Expanding prior ideas through questions. The second sequence category encompasses two variations: HQ-QU (N = 39) and LQ-QU (N = 98) for a total of 137 out of the 244 differentiating sequences. The sequence variation that leads to high-quality ideas (HQ-QU, N = 39) includes a prior high-quality idea, followed by a question. When the prior idea is assessed as high-quality (i.e., both novel and offering a competitive advantage), crowd participants form a motivation to emulate the prior idea (Kohn et al., 2011). This motivation leads to the contribution of a question, thus seeking to clarify the mental space of the prior idea. This encourages incremental cognitive effort on the part of the next poster to blend a mental space of their own with the modified mental space of the prior idea in such a way that another high-quality idea is generated.

In the example below, the prior high-quality idea concerns a "portable check printer." The mental space pertaining to this idea is then expanded by a question, which highlights specific concepts regarding mobility in general within the already introduced mental space. When the next poster reads the prior idea and the subsequent question, they blend the expanded mental space of the prior idea with a mental space of their own, yielding a new high-quality idea with regard to a mobile app.

It's such a hassle for frequent travelers who prefer the security of checks when it comes to international payments. [PaymentCo] can create a totally innovative product to fix this problem for its clients: the portable check printer! Think of a calculator sized check printer which prints currency-customized checks with a swipe of your credit card. This is perfect for middle/high income travelers, whether on vacation or business. (High-Quality Idea)

Some questions on feasibility, if you could be more specific on size, mobility (ATM tracks?), security, efficiency (cash carried). (Question)

Mobile app that tracks the the moving ATMs (ATM on wheels). Whenever they are close by, we can just give them a call or poke on the app and they'll come so we can deposit or withdraw cash at the convenience of our current locations. (High-Quality Idea)

Expanding prior ideas through questions also contains a sequence variation that leads to low-quality ideas: LQ-QU (N = 98). As we have seen above, a prior low-quality idea tends to give rise to a motivation to redirect, which when coupled with an analogy that shifts the mental space to a new context may support blending in such a way that a high-quality idea is generated. At times, however, crowd participants may seek to emulate an idea, even if it is judged as being of low-quality. In such a situation, the subsequent problem-related information is not properly aligned with the quality of the idea, thus leading to incremental information contributions (i.e., expansion through inquiry) toward modifying the mental space of the prior, low-quality idea. This then suggests to the next poster that emulation of the prior idea is appropriate, which yields another low-quality idea.

Expanding prior ideas through paradoxes. This sequence category consists of two variations accounting for 55 out of the 244 differentiating sequences: HQ-PA (N = 17) and LQ-PA (N = 38). In the sequence preceding high-quality ideas (HQ-PA, N = 17), the prior idea is assessed as being both novel and offering a competitive advantage, and therefore crowd participants form a motivation to emulate the prior idea. This motivation leads to the contribution of a paradox (Miron-Spektor et al., 2011; Poole & Van de Ven, 1989), thus examining tensions between concepts, at least some of which already exist in the mental space of the prior idea. This encourages incremental cognitive effort on the part of the next poster to blend a mental space of their own with the modified mental space of the prior idea in such a way that another high-quality idea is generated. Pointing out paradoxical relationships that involve at least some concepts that already exist within the mental space of the prior idea signals to the next poster that resolving the identified paradoxes through blending may be sufficient for generating a new high-quality idea.

As the example below illustrates, the prior high-quality idea about a platform for 3D printing of cars is followed by a paradox that expands the mental space pertaining to the prior idea through examining tensions between the design capacities of humans and machines. The crowd participant reading the prior idea and the subsequent paradox then blends the expanded mental space of the prior idea with a mental space of their own, generating a new high-quality idea by identifying specific parts (i.e., a center console) of a car which may be particularly suited to 3D printing.

It will be cool to see members of the community participate in creating their own cars through 3D printing. This could be done in an online platform where people can share their car designs. Everyone eventually becomes a car designer! 3D printing will

give a more realistic picture of the car designs and drive feedback as to improve the models. (High-Quality Idea)

It would be very neat to have the cars designed more by 3D program than by humans, but it might put human designers out of work. And it is unknown if computer design can be superior to human [design]. (Paradox)

Many times, in the car, there's no room for storage because it just isn't the right shape or size: you can't put large sized McDonalds cup or something like that. Or maybe you want to put your tissue box and wallet and have compartments for that. With 3D printed center console, you can have the right shape and size just to fit you. (High-Quality Idea)

Expanding prior ideas through paradoxes also contains a sequence variation that leads to low-quality ideas: LQ-PA, N = 38. When a prior low-quality idea is followed by a paradox (i.e., expansion through examining tensions between concepts), the subsequent problem-related information is not properly aligned to the quality of the idea. This then suggests to the next poster that emulation of the prior idea is appropriate, which therefore yields another low-quality idea.

Summary of Theory Development

In sum, we suggest that participants in an online ideation contest read prior contributions and judge prior ideas subjectively for idea quality (based on meeting the problem requirements). These judgments then motivate participants to create high-quality ideas—but only provided that the problem-related information contributed by the crowd is aligned with whether the prior idea is assessed as high- or low-quality. We suggest that both high- and low-quality ideas can foster high-quality ideas through blending—but only if the crowd contributes information that expands the mental space of a prior high-quality idea and shifts the mental space of a prior low-quality idea.

We suggest that these judgments of prior ideas occur even in the absence of crowd-voting mechanisms or official judgments made by the sponsor. Moreover, we suggest that assessments of the quality of the prior idea give rise to one of two motivations: emulation or redirection.

An idea assessed as being high-quality motivates a participant to emulate this idea, using it as an "exemplar" that suggests "promising directions for uncovering creative ideas" (Harvey, 2014, p. 331). Under such circumstances, a question or paradox posed by the crowd after the idea is useful in expanding the mental space of the prior idea in such a way that subsequent blending (Fauconnier & Turner, 1998, 2003) will yield a high-quality idea.

Assessing an idea to be of low-quality, in contrast, redirects participants away from the prior idea, encouraging them to shift the mental space pertaining to the prior idea (Fiske & Taylor, 1984; Hargadon & Bechky, 2006). Under such circumstances, an analogy posed by the crowd after the idea is useful in shifting the mental space of the prior idea to a new context, thereby enabling subsequent blending (Fauconnier & Turner, 1998, 2003) to generate a high-quality idea. As such, both prior low- and high-quality ideas can foster high-quality ideas—depending on what the crowd posts following the prior idea. Our theorizing is illustrated in Figure 1.

Discussion

Our theorizing suggests implications for future research on crowd ideation processes. Further, our methodological approach holds implications for research using sequence analysis.

Theoretical Implications

Our research extends prior work on ideation contests (Bayus, 2013; Kohn et al., 2011; Liu et al., 2014; Wang et al., 2018) and online brainstorming (Dennis & Williams, 2003; Seeber et al., 2017). As with that research, we find that the quality of the prior ideas influences subsequent ideas. Of greater importance, though, we find that, unlike past research, the quality of the prior idea is not enough to explain the quality of the subsequent idea produced. Instead, what is needed is an alignment of the quality of the prior idea with the problem-related information contributed by the crowd afterwards. This suggests that future research should focus not only on the ideas themselves but also on the sequencing and alignment of specific types of problem-related information contributed by the crowd.

These findings suggest avenues for further research. Much of our theorizing is about latent unobserved processes—crowd participants subjectively assess ideas when they read them, different assessments motivate crowd participants' creativity in different ways, and crowd participants modify and blend various mental spaces as they develop a new idea. This begs questions such as whether crowd participants reliably can judge the quality of ideas or not. Surfacing these latent cognitive processes is an important next step in research on crowd behavior.

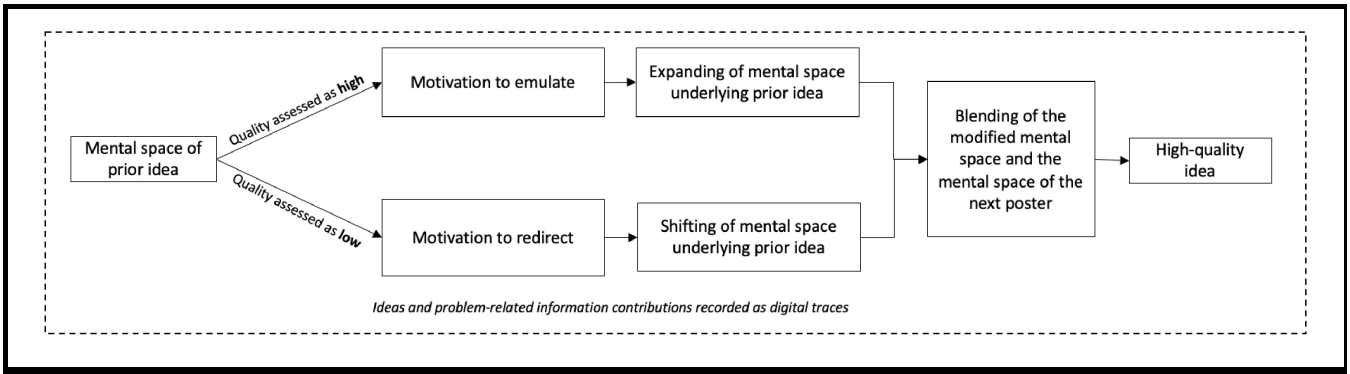

Figure 1. How Expanding and Shifting Enable the Generation of High-Quality Ideas through Blending

Our theorizing describes the need for alignment between assessment of the quality of a prior idea and the type of problem-related information subsequently contributed by the crowd. We do not theorize if the crowd is aware of the need for such alignment, and then contributes accordingly. Is it possible that some members in the crowd intuit the value that such alignment provides, and then contribute per such alignment? Alternatively, is it possible that the alignment is simply serendipitous—that the crowd contributes what they are interested in contributing and it happens that at times the crowd's contributions align with the quality of prior ideas and at other times they do not? Research is needed to uncover whether there is any intentionality at play here, or whether the alignment is a foundational stigmergic principle that crowd participants unconsciously adhere to (Bolici et al., 2015; Majchrzak et al., 2021).

Our findings also suggest that the particular ways in which mental spaces are modified by the information content of subsequent problem-related contributions matter for producing high-quality ideas. This implies that further research is needed on different mechanisms of how mental spaces can be modified and blended over time to generate various outcomes of interest. Specifically, researchers may want to ask what are the cognitive processes underlying emulation and redirection? Answering this question may help identify routines and practices (Lindberg et al., 2013) within crowds. This represents a potential contribution, as crowds have traditionally been seen as fluid (Faraj et al., 2011) and lacking in well-defined processes. Recent work has, however, suggested that structured processes (Faraj et al., 2016) may emerge in crowds in a bottom-up fashion (Lindberg et al., 2016), but we know little about how such processes may be structured or influenced. To support the emergence of structured processes that are helpful in generating high-quality ideas, future research may want to investigate how specific platform affordances (see Dennis et al., 1999; Dennis et al., 1996) may motivate crowd participants to respond to prior ideas by contributing information that is likely to stimulate the contribution of subsequent high-quality ideas.

Alternative theories for explaining the findings should be developed. For example, is it possible that the results indicate a particular form of modification and blending of mental spaces that can be described as "reinterpretation" of earlier contributions through particular sequences of contributions? Majchrzak and Malhotra (2016) found that paradoxes followed by assumptions were more likely to be generative than paradoxes not followed by assumptions. They suggest that this was because the assumptions posted by a crowd participant after another crowd participant posted the paradox helped validate the paradox (i.e., helped to reinterpret the paradox in a positive light). Is it possible, then, that a sequence, such as a question following a low-quality idea, can only generate another low-quality ideas because low-quality ideas lead to cognitive inertia (Dane, 2010; Dennis et al., 1997), which blocks reinterpretation and therefore prevents the generation of a high-quality idea? Research on such alternative theories about sequences of crowd contributions is needed.

Finally, our theorizing may be of interest to scholars working outside of the immediate context of online ideation contests, especially those who focus on peer-production settings such as open source software (Lindberg et al., 2016; Singh et al., 2011) or wikis (Arazy, Lindberg, Lev, et al., 2020; Arazy Lindberg, Rezaei et al., 2020; Keegan et al., 2015). In such contexts, what patterns of mental space modifications may lie behind creative problem-solving and the production of high-quality artifacts?

Methodological Implications

To support a research program on the sequential patterns of crowd behavior and its effect on the emergence of high-quality ideas, sophisticated methods that are sensitive to the sequential

ordering of contributions varying in content need to be developed further. This involves, for example, developing VLMC methods that can use numerical variables as predictors, as well as multi-level nesting of observations. We encourage predictive work since there is already substantial amounts of descriptive and classificatory research on sequences (Klarner & Raisch, 2013; Sabherwal & Robey, 1993), such as optimal matching-based studies and attendant methods (Abbott & Tsay, 2000; Gaskin et al., 2014). By supplementing the results of VLMC with lift calculations and their statistical significance, we have moved VLMC a step closer to predictiveness. To make VLMC broadly useful as a predictive tool for IS researchers, however, much more analytics work is needed.

Limitations

Several limitations of our research can serve to spur on future studies. It is possible the sequences we uncovered could be specific to the contexts of the problem statements presented to the crowd (Seidl et al., 2019). Although high-quality ideas may emerge in all contexts of idea crowdsourcing, the sequencing of ideas and information contributions required may be different depending on, for example, the complexity of the problems posed to the crowd. In peer-production contexts focused on more complex artifacts, such as software, ideas may not be contributed as isolated units but may rather consist of multiple pieces of interrelated code. Drawing on the mirroring hypothesis (Colfer & Baldwin, 2016), we may expect that the sequences that yield high-quality outcomes in such contexts are more complex than the sequences we have identified.

Similarly, iteration on ideas and artifacts may differ because of their degree of path dependence (Van Driel & Dolfsma, 2009) based on the types of tools used to create said artifacts (Zhang et al., 2021) as well as the flexibility of the artifacts themselves (Kallinikos et al., 2013; Yoo et al., 2010). Therefore, a range of problem types based on the degree of flexibility, digitalization, and automation in a particular context should be examined to ascertain whether the types of sequences that we have identified apply across a wide range of problems.

Further, we assumed that the subjective assessments of quality made by crowd participants matched those of the executives rating the ideas after the ideation contests concluded. There may, however, be multiple sources of variance here: crowd participants may rate ideas differently from executives, and ideas may be rated differently depending on whether they are rated in real-time or at the end of an ideation contest. Future research may want to examine the dynamics of assessing ideas from different perspectives and at different times.

Finally, the VLMC method employed to investigate the sequencing of ideas and additional information, while associated with benefits relative to other methods, such as the ability to use subsequences to predict outcomes embedded within longer sequences, also comes with limitations. The most severe of these limitations is the inability to use control variables. Future work needs to develop methods which allow for analysis of sequences of variable length while incorporating control variables.

Conclusion

Through focusing on sequences of ideas and problem-related information contributions, this research provides some suggestions for why past research on online ideation contests have provided mixed findings with regard to the effects of prior ideas on subsequent ideas. We find that the effect of prior ideas on subsequent ideas depends on how subjective assessments of the prior ideas' quality are aligned with subsequent problem-related information contributed by the crowd. We also illustrate, and provide improvements to, variable-length Markov chains, thus expanding opportunities for future research to use this method to predict outcomes of various sequences of events and activities. Future research should extend its focus beyond obtaining a range of ideas to choose from and examine types and sequences of information contributions made by the crowd in-between ideas. Doing so will shed light on the dynamics of how crowds can yield high-quality ideas through emulating prior high-quality ideas, and even use prior low-quality ideas as stepping-stones toward subsequent high-quality ideas.

References

Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology: Review and prospect. *Sociological Methods & Research*, 29(1), 3–33.

Amabile, T. M., Barsade, S. G., Mueller, J. S., & Staw, B. M. (2005). Affect and creativity at work. *Administrative Science Quarterly*, 50(3), 367–403.

Ansburg, P. I., & Hill, K. (2003). Creative and analytic thinkers differ in their use of attentional resources. *Personality and Individual Differences*, 34(7), 1141–1152.

Arazy, O., Lindberg, A., Lev, S., Wu, K., & Yarovoy, A. (2020). Emergent routines in peer-production: examining the temporal evolution of wikipedia's work sequences. *ACM Transactions on Social Computing*, 3(1), 1–24.

Arazy, O., Lindberg, A., Rezaei, M., & Samorani, M. (2020). The evolutionary trajectories of peer-produced artifacts: Group composition, the trajectories' exploration, and the quality of artifacts. *MIS Quarterly*, 44(4), 2013–2053.

Baralou, E., & Tsoukas, H. (2015). How is new organizational knowledge created in a virtual context? An ethnographic study. *Organization Studies*, 36(5), 593–620.

Bayus, B. L. (2013). Crowdsourcing new product ideas over time: an analysis of the Dell IdeaStorm community. *Management Science*, 59(1), 226–244.

Begleiter, R., El-Yaniv, R., & Yona, G. (2004). On prediction using variable order Markov models. *Journal of Artificial Intelligence Research*, 22), 385–421.

Berente, N., Seidel, S., & Safadi, H. (2019). Research Commentary—Data-driven computationally-intensive theory development. *Information Systems Research*, 30(1), 50–64.

Boland Jr., R. J., & Tenkasi, R. V. (1995). Perspective making and perspective taking in communities of knowing. *Organization Science*, 6(4), 350–372.

Bolici, F., Howison, J., & Crowston, K. (2015). Stigmergic coordination in FLOSS development teams: Integrating explicit and implicit mechanisms. *Cognitive Systems Research*, 38, 14–22.

Bühlmann, P., & Wyner, A. J. (1999). Variable length Markov chains. *The Annals of Statistics*, 27(2), 480–513.

Chesbrough, H. W. (2006). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Choi, H. H., & Kim, M. J. (2017). The effects of analogical and metaphorical reasoning on design thinking. *Thinking Skills and Creativity*, 23, 29–41.

Colfer, L. J., & Baldwin, C. Y. (2016). The mirroring hypothesis: Theory, evidence, and exceptions. *Industrial and Corporate Change*, 25(5), 709–738.

Cronin, M. A., & Weingart, L. R. (2007). Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review*, 32(3), 761–773.

Dahlander, L., Jeppesen, L. B., & Piezunka, H. (2019). How organizations manage crowds: define, broadcast, attract, and select. *Research in the Sociology of Organizations*, 64, 239–270.

Dane, E. (2010). Reconsidering the trade-off between expertise and flexibility: A cognitive entrenchment perspective. *The Academy of Management Review*, 35(4), 579–603.

Dennis, A. R., Aronson, J. E., Heninger, W. G., & Walker, E. D. (1999). Structuring time and task in electronic brainstorming. *MIS Quarterly*, 23(1), 95–108.

Dennis, A. R., Valacich, J. S., Carte, T. A., Garfield, M. J., Haley, B. J., & Aronson, J. E. (1997). Research Report: The effectiveness of multiple dialogues in electronic brainstorming. *Information Systems Research*, 8(2), 203–211.

Dennis, A. R., Valacich, J. S., Connolly, T., & Wynne, B. E. 1996. Process Structuring in Electronic Brainstorming. *Information Systems Research*, 7(2), 268–277.

Dennis, A. R., & Williams, M. L. (2003). Electronic brainstorming. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 160–178). Oxford University Press.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge collaboration in online communities. *Organization Science*, 22(5), 1224–1239.

Faraj, S., von Krogh, G., Monteiro, E., & Lakhani, K. R. (2016). Online community as space for knowledge flows. *Information Systems Research*, 27(4), 668–684.

Fauconnier, G., & Turner, M. (1998). Conceptual integration networks. *Cognitive Science*, 22(2), 133–187.

Fauconnier, G., & Turner, M. (2003). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.

Finke, R. A., Ward, T. B., & Smith, S. M. (1996). *Creative cognition: Theory, research, and applications*. MIT Press.

Fiske, S., & Taylor, S. (1984). *Social cognition: From brains to culture* (2nd ed.). SAGE Publications.

Füller, J., Hutter, K., & Faullant, R. (2011). Why co-creation experience matters? Creative experience and its impact on the quantity and quality of creative contributions. *R&D Management*, 41(3), 259–273.

Gabadinho, A., & Ritschard, G. (2016). Analyzing state sequences with probabilistic suffix trees: The PST R package. *Journal of Statistical Software*, 72(3), 1–39.

Garrison, D. R. (2015). *Thinking collaboratively: Learning in a community of inquiry*. Routledge.

Gaskin, J., Berente, N., Lyytinen, K., & Yoo, Y. (2014). Toward generalizable sociomaterial inquiry: A computational approach for zooming in and out of sociomaterial routines. *MIS Quarterly*, 38(3), 849–871.

Gentner, D., & Stevens, A. L. (2014). *Mental models*. Psychology Press.

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea generation and the quality of the best idea. *Management Science*, 56(4), 591–605.

Hargadon, A. B., & Bechky, B. A. (2006). When collections of creatives become creative collectives: A field study of problem solving at work. *Organization Science*, 17(4), 484–500.

Hargrave, T. J., & Van de Ven, A. H. (2017). Integrating dialectical and paradox perspectives on managing contradictions in organizations. *Organization Studies*, 38(3-4), 319–339.

Hart, C. (2007). Mental spaces, blended spaces and discourse spaces in the British national party. In C. Hart & D. Lukeš (Eds.), *Cognitive linguistics in critical discourse analysis: Application and theory* (pp. 107–131). Cambridge Scholars Publishing.

Harvey, S. (2014). Creative synthesis: Exploring the process of extraordinary group creativity. *Academy of Management Review*, 39(3), 324–343.

Hender, J. M., Rodgers, T. L., Dean, D. L., & Nunamaker, J. F. (2001). Improving group creativity: Brainstorming versus non-brainstorming techniques in a GSS environment. In *Proceedings of the 34th Hawaii International Conference on System Sciences*. IEEE Computer Society Press.

Holland, J., Holyoak, K., Nisbett, R., & Thagard, P. (1989). *Induction: Processes of inference, learning, and discovery.* MIT Press.

Hutter, K., Hautz, J., Füller, J., Mueller, J., & Matzler, K. (2011). Communitition: The tension between competition and collaboration in community-based design contests. *Creativity and Innovation Management*, 20(1), 3–21.

Jay, J. (2013). Navigating paradox as a mechanism of change and innovation in hybrid organizations. *The Academy of Management Journal*, 56(1), 137–159.

Kallinikos, J., Aaltonen, A., & Marton, A. (2013). The ambivalent ontology of digital artifacts. *MIS Quarterly*, 37(2), 357–370.

Keegan, B. C., Lev, S., & Arazy, O. (2015). Analyzing organizational routines in online knowledge collaborations: A case for sequence analysis in CSCW. In *Proceedings of the 18th ACM conference on computer-supported cooperative work and social computing*, pp. 1065–1079.

Kellogg, K. C. (2009). Operating room: Relational spaces and microinstitutional change in surgery. *American Journal of Sociology*, 115(3), 657–711.

Klarner, P., & Raisch, S. (2013). Move to the beat—Rhythms of change and firm performance. *Academy of Management Journal*, 56(1), 160–184.

Kohn, N. W., Paulus, P. B., & Choi, Y. H. (2011). Building on the ideas of others: An examination of the idea combination process. *Journal of Experimental Social Psychology*, 47(3), 554–561.

Lamastra, C. R. (2009). Software innovativeness. A comparison between proprietary and free/open source solutions offered by Italian SMEs. *R&D Management*, 39(2), 153–169.

Landis, J. R., & Koch, G. G. (1977). The Measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.

Latané, B., & Liu, J. H. (1996). The intersubjective geometry of social space. *Journal of Communication*, 46(4), 26–34.

Lee, A. S., & Baskerville, R. L. (2003). Generalizing generalizability in information systems research. *Information Systems Research*, 14(3), 221–243.

Lenca, P., Meyer, P., Vaillant, B., & Lallich, S. (2008). On selecting interestingness measures for association rules: User oriented description and multiple criteria decision aid. *European Journal of Operational Research*, 184(2), 610–626.

Leonard-Barton, D., & Swap, W. (2005). *When sparks fly: Harnessing the power of group creativity*. Harvard Business Press.

Lindberg, A. (2020). Developing theory through integrating human and machine pattern recognition. *Journal of the Association for Information Systems*, 21(1), 90–116.

Lindberg, A., Berente, N., Gaskin, J., & Lyytinen, K. (2016). Coordinating interdependencies in online communities: A study of an open source software project. *Information Systems Research*, 27(4), 751–772.

Lindberg, A., Gaskin, J., Berente, N., Lyytinen, K., & Yoo, Y. (2013). Computational approaches for analyzing latent social structures in open source organizing. In *Proceedings of the 34th International Conference on Information Systems.*

Liu, T. X., Yang, J., Adamic, L. A., & Chen, Y. (2014). Crowdsourcing with all-pay auctions: A field experiment on Taskcn. *Management Science*, 60(8), 2019–2037.

Mächler, M., & Bühlmann, P. (2004). Variable length Markov chains: Methodology, computing and software. *Journal of Computational and Graphical Statistics*, 13(2), 435–455.

Majchrzak, A., & Malhotra, A. (2016). Effect of knowledge-sharing trajectories on innovative outcomes in temporary online crowds. *Information Systems Research*, 27(4), 685–703.

Majchrzak, A., & Malhotra, A. (2020). *Unleashing the crowd: Collaborative solutions to wicked business and societal problems*. Springer.

Majchrzak, A., Malhotra, A., & Zaggl, M. (2021). How Open Crowds Self Organize. *Academy of Management Discoveries*, 7(1).

Miron-Spektor, E., Gino, F., & Argote, L. (2011). Paradoxical frames and creative sparks: Enhancing individual creativity through conflict and integration. *Organizational Behavior and Human Decision Processes*, 116(2), 229–240.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.

Nijstad, B., Diehl, M., & Stroebe, W. (2003). Cognitive stimulation and interference in idea-gererating groups. In P. B. Paulus & B. A. Nijstad (Eds.), *Group creativity: Innovation through collaboration* (pp. 137–159). Oxford University Press.

Oakley, T., & Hougaard, A. (2008). *Mental spaces in discourse and interaction*. John Benjamins Publishing Company.

Paulus, P. B., & Nijstad, B. A. (Eds.). (2003). *Group creativity: Innovation through collaboration*. Oxford University Press.

Poole, M. S., & Van de Ven, A. (1989). Using paradox to build management and organization theories. *Academy of Management Review*, 14(4), 562–578.

Rao, H., & Dutta, S. (2012). Free spaces as organizational weapons of the weak: Religious festivals and regimental mutinies in the 1857 Bengal native army. *Administrative Science Quarterly*, 57(4), 625–668.

Ron, D., Singer, Y., & Tishby, N. (1996). The power of amnesia: Learning probabilistic automata with variable memory length. *Machine Learning*, 149, 117–149.

Sabherwal, R., & Robey, D. (1993). An empirical taxonomy of implementation processes based on sequences of events in information system development. *Organization Science*, 4(4), 548–576.

Santanen, E. L., Briggs, R. O., & De Vreede, G. J. (2004). Causal relationships in creative problem solving: Comparing facilitation interventions for ideation. *Journal of Management Information Systems*, 20(4), 167–197.

Seeber, I., de Vreede, G. J., Maier, R., & Weber, B. (2017). Beyond brainstorming: Exploring convergence in teams. *Journal of Management Information Systems*, 34(4), 939–969.

Seidl, D., von Krogh, G., & Whittington, R. (2019). Defining open strategy: Dimensions, practices, impacts, and perspectives. In D. Seidl, G. von Krogh, & R. Whittington (Eds.), *Cambridge handbook of open strategy* (pp. 9–23). Cambridge University Press.

Shrestha, Y. R., He, V. F., Puranam, P., & von Krogh, G. (2021). Algorithm supported induction for building theory: How can we use prediction models to theorize? *Organization Science*, 32(3), 856–880.

Singh, P. V., Tan, Y., & Youn, N. (2011). A hidden Markov model of developer learning dynamics in open source software projects. *Information Systems Research*, 22(4), 790–807.

Terwiesch, C., & Xu, Y. (2008). Innovation contests, open innovation, and multiagent problem solving. *Management Science*, 54(9), 1529–1543.

Tsoukas, H. (2009). A dialogical approach to the creation of new knowledge in organizations. *Organization Science*, 20(6), 941–957.

Turner, M. (2014). *The origin of ideas: Blending, creativity, and the human spark*. Oxford University Press.

Van Driel, H., & Dolfsma, W. (2009). Path dependence, initial conditions, and routines in organizations: The Toyota production system re-examined. *Journal of Organizational Change Management*, 22(1), 49–72.

Wang, K., Nickerson, J., & Sakamoto, Y. (2018). Crowdsourced idea generation: The effect of exposure to an original idea. *Creativity and Innovation Management*, 27(2), 196–208.

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). The new organizing logic of digital innovation: An agenda for information systems research. *Information Systems Research*, 21(4), 724–735.

Zhang, Z., Yoo, Y., Lyytinen, K., & Lindberg, A. (2021). The unknowability of autonomous tools and the liminal experience of their use. *Information Systems Research*, 32(4), 1192–1213.

Zheng, H., Li, D., & Hou, W. (2011). Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15(4), 57–88.

About the Authors

Aron Lindberg is an assistant professor of Information Systems at the School of Business, Stevens Institute of Technology. He received his Ph.D. at Case Western Reserve University (2015), and primarily studies complex, distributed digital innovation processes, often using a combination of qualitative and computational methods. His research has been published in major journals such as *MIS Quarterly, Information Systems Research*, and *Journal of the Association for Information Systems*.

Ann Majchrzak is The Associates' of USC Chair of Business Administration for the Marshall School of Business at the University of Southern California and Fellow of the Association for Information Systems. She is interested in the inherent paradox between providing structure and yet allowing perspectives to organically collide to foster co-creation in crowds. Her most recent book is with Arvind Malhotra: *Unleashing the Crowd: Collaborative Solutions for Wicked Business and Societal Problems*. She has held honorary appointments at Vrije Universiteit Amsterdam, Luiss University, Simon Fraser, and ESADE.

Arvind Malhotra is the H. Allen Andrews Professor of Entrepreneurial Education and Professor of Entrepreneurship and Strategy at the Kenan-Flagler Business School, the University of North Carolina at Chapel Hill. His research focuses on open-innovation organizational and extra-organizational structures. Arvind's work has been published in leading journals such as *Harvard Business Review, MIT Sloan Management Review, MIS Quarterly, Information Systems Research, Journal of Management, Human Resource Management,* and *Journal of Academy of Marketing Science.* He is currently serving on the editorial boards of *Organization Science, Journal of Strategic Information Systems,* and *Journal of Knowledge Management*.