ICIS 2022 Proceedings

Sharing Economy, Platforms, and Crowds

Dec 12th, 12:00 AM

Examining the Impact of Citation Network Embeddedness on Crowdsourced Idea Refinement: The Moderating Role of Idea Breadth and Depth

Vincent Z.W. Mack
Singapore Management University, vincentmack.2019@phdis.smu.edu.sg

Yimei Zhou
Tsinghua University, zhouym.18@sem.tsinghua.edu.cn

Qian Tang
Singapore Management University, QIANTANG@smu.edu.sg

The Impact of Citation Network Embeddedness on Crowdsourced Idea Refinement: The Moderating Role of Idea Breadth and Depth

Short Paper

Vincent Z.W. Mack
Singapore Management University
80 Stamford Road, Singapore 178902
Singapore
vincentmack.2019@phdis.smu.edu.sg

Yimei Zhou
Tsinghua University
30 Shuangqing Rd, Haidian District,
Beijing 100084, China
zhouym.18@sem.tsinghua.edu.cn

Qian Tang
Singapore Management University
80 Stamford Road, Singapore 178902 Singapore
qiantang@smu.edu.sg

Abstract

Idea refinement is a crucial step in the ideation process. Given that ideas are embedded within a network of other related ideas, how the network structure fluctuates throughout the idea refinement process may affect its end outcome. Through studying proposed ideas on TVTropes.org, this research seeks to understand how network embeddedness of an idea's citation network influences the idea quality perception, particularly through degree centrality, betweenness centrality and eigenvector centrality. We also study the boundary conditions of their impact on quality and analyze the moderating role of topic breadth and depth which capture the overall content of the idea. Our results suggests that the positive effect is stronger when the idea topic breadth is lesser, or the idea topic depth is greater. We further explore the mechanism behind these effects by analyzing their effect on change of positive votes and negative votes, showing the influence path might be different.

Keywords: Network embeddedness, centrality, idea refinement, topic breadth, topic depth

Introduction

Ideation is the creative process of generating new ideas from existing ideas through search and knowledge recombination (Katila and Ahuja 2002; Burt 2004; Kaplan and Vakili 2015). Effective ideation is an integral part of innovation, and crucial to scientific advancement and technological development. From a network perspective, ideation involves exploring the search space defined by connected ideas in a global idea network. In recent years, crowdsourcing has been gradually adopted for ideation. By inviting a crowd to perform a task, crowdsourcing can overcome the limitations and knowledge gaps of individuals or small groups in ideation by leveraging the collective intelligence. This resulting parallelization speeds up the iterative network search and recombination. And thus, crowdsourcing ideation has been attached great importance and widely used by both individuals and firms (Huang et al. 2014). For example, Amazon

Studios, Amazon's crowdsourcing TV and movie entertainment platform, allows users to submit and edit scripts and even vote for which pilots to produce, contributing to content production. On many crowdsourcing platforms, ideas get evaluated for suitability and promise before getting launched (Seeber et al. 2017). For example, the Lego review board decides whether to put an idea into production only if more than 10,000 people support the idea. Usually ideas undergo a process of revision and refinement where idea creators build up their ideas through critical inspection and bounce them to other collaborators for testing and feedback (Perry-Smith and Mannucci 2017). Given that firms and individuals invest heavily in crowdsourcing ideation, it is of great practical importance to provide some guidance for the idea elaboration or refinement phase, as ideas do not usually arrive at their finalized form in the first iteration.

However, the idea refinement phase has largely been ignored in the literature (Seeber et al. 2017). Existing literature in the management field mainly focuses on idea generation or implementation (e.g., Deichmann et al. 2020; Dennis and Valacich 1997; Huang et al. 2014). And idea refinement is usually alluded to or studied implicitly as an extension of the idea generation process, or subsumed under the entire ideation process (Perry-Smith and Mannucci 2017). The few pieces of literature that study idea refinement and its impact on idea quality tend to focus on correlational predictions rather than causal relationships, most of which belong to the computer science field (Betancourt et al. 2016; Dalip et al. 2011; De la Calzada and Dekhtyar 2010). These studies usually aim at predicting high-quality Wikipedia articles through generating features from article edit history, ignoring the underlying causal relationships between features in idea refinement and idea quality.

Motivated by the above research gaps and practical problems, our study aims to understand how idea refinement affects eventual quality from the perception of crowds. And crowdsourced idea refinement process begins when an idea is proposed and exposed to the online crowd for revisions and comments and ends with the acceptance or rejection of the idea. Recognizing the criticality of the connected ideas in ideation (Katila and Ahuja 2002; Kaplan and Vakili 2015), we explore what role the citation network plays in idea refinement. Specifically, altering the set of source ideas for recombination during the refinement phase may cause perturbations in the network embeddedness of the new idea in the global network of ideas. Past researchers have highlighted the importance of embeddedness in access information or other resources in the network (e.g., Grewal et al. 2006; Ransbotham et al. 2012). We thus focus on the impact of changes of citation network embeddedness which captures the connection of the idea with other ideas through citations (Granovetter 1985; Grewal et al. 2006; Ransbotham et al. 2012). Also, we are interested in the moderating roles of topic breadth and depth which capture the overall content semantics of the idea. Therefore, we propose the following research questions:

RQ1: How do changes of citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) in the refinement process impact the crowd perception of idea quality?

RQ2: How do topic breadth and depth of the idea moderate the relationships between changes in citation network embeddedness and the crowd perception of idea quality?

To answer these questions, we collected data from TVTropes.org, an online ideation community and wiki platform where members of this community interact to propose and critique novel ideas. Using panel data on crowdsourcing idea refinement, we find that increasing citation network embeddedness has a significant positive effect on idea quality. We also find that the positive effect is stronger when the idea topic breadth is smaller or the idea topic depth is greater. Furthermore, we explore the mechanism behind these effects by analyzing their effect on the changes of positive votes and negative votes, showing the influence path might be different.

Our research aims to make several important contributions. First, this paper extends the literature on crowdsourcing ideation by considering the impact of features in the refinement process while most of the previous literature focuses on idea generation or implementation (e.g., Deichmann et al. 2020; Dennis and Valacich 1997; Huang et al. 2014). Second, our research studies the moderating role of idea content in the relationship between structural position changes and idea quality and thus providing insights into the interactive effect of the sematic aspect and structural aspect of ideas in crowdsourcing. Third, our study explores the relationship between idea refinement and the change of positive and negative votes, enriching the understanding of the process of crowdsourced idea refinement. Finally, our findings offer practical implications for individuals and firms engaging in crowdsourcing ideation.

Literature Review

Existing literature on crowdsourcing ideation lags. First, previous literature on crowdsourcing ideation in the management field mainly focuses on idea generation or implementation (e.g., Deichmann et al. 2020; Dennis and Valacich 1997; Huang et al. 2014), ignoring the process of refinement. Only a few of them study idea refinement implicitly as an extension of the idea generation process while most of them include just idea refinement in the entire ideation process (Perry-Smith and Mannucci 2017). The research implicitly studying idea refinement usually focuses on studying the collaboration patterns between idea creators and evaluators and try to understand the effect of idea refinement on certain outcomes such as idea diffusion (Deichmann et al. 2020) or market success (Burt 2004). The implicit assumption behind these studies is that idea commentators are not merely passive participants, oftentimes their comments may directly influence the author, or in certain cases are gatekeepers evaluating the suitability of an idea for launch. This leads to the modification of the idea content during this idea refinement phase. Many of these studies investigate the role of feedback in increasing participant idea contribution behavior and productivity (Satzinger et al. 1999; Garfield et al. 2001; Jung et al. 2010; Ransbotham and Kane 2011; Zhang and Wang 2012; Huang et al. 2019; Piezunka and Dahlander 2019; Dargahi and Namin 2020) or end-product quality (Hildebrand et al. 2013; Wooten and Ulrich 2017; Dargahi and Namin 2020). However, unlike the idea generation phase where the goal is usually to generate as many ideas as possible, idea refinement seeks to improve and select the most promising ideas (Perry-Smith and Mannucci 2017; Seeber et al. 2017). To make it more clear, crowdsourced idea refinement is the collaborative idea development process between idea generation and idea implementation (Perry-Smith and Mannucci 2017).

Second, there are also some research on idea refinement and its impact on idea quality in the computer science field (Betancourt et al. 2016; Dalip et al. 2011; De la Calzada and Dekhtyar 2010), but these studies also ignore the underlying relationships between features in idea refinement and idea quality. These pieces of literature usually focus on predicting, ignoring the rationale behind selecting the relevant features and not further investigating the causality underlying relationships between content development factors, and article quality.

Hypothesis Development

Effect of Network Embeddedness on Idea Quality

We first consider the impact of citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) changes in idea refinement on its quality. New ideas are constructed from the recombination of older ideas (Katila and Ahuja 2002; Burt 2004; Kaplan and Vakili 2015). This set of base ideas are cited by ideators during the conceptualization phase to serve as reference points for fellow ideators and reviewers and they may change during each refinement iteration. These connections within idea networks indicate flows of knowledge elements between them (Wang et al. 2014). Ideas are cited for several reasons: (1) to pay homage to core or central ideas in whichever tradition the new idea is part of; (2) to compare and contrast similar or different ideas ; and (3) to serve as stimuli (Wang and Nickerson 2019) or respond to questions raised in previous ideas (Fan 1985). Given the networked structure of interconnected ideas, research into effective ideation has focus on identifying how structural characteristics of ideas affect its perceived quality (Burt 2004; Björk and Magnusson 2009; Zhang and Wang 2012; Wang et al. 2014; Deichmann et al. 2020).

Citation network embeddedness is the degree to which an idea is connected to other ideas in the citation network, showing the structural position of the idea (Granovetter 1985; Grewal et al. 2006; Ransbotham et al. 2012). The previous literature suggests that there are three dimensions of network embeddedness, i.e., degree centrality which captures the extent to which an idea is entrenched in the citation network, betweenness centrality which captures the extent to which an idea connects other cited ideas, and eigenvector centrality which captures the extent to which an idea is connected with other ideas entrenched in the citation network (Grewal et al. 2006; Björk and Magnusson 2009; Ransbotham et al. 2012). When increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) in idea refinement, the proposed idea is more likely to improve in quality. First, the proposed idea can involve more knowledge and resources when organizing its own content for more cited ideas (degree centrality). Second, the proposed idea can improve the quality of the knowledge and resources involved when generating its own content for the fact that it has good connectivity with other ideas in general (betweenness centrality), especially the ideas of great importance (eigenvector centrality). To conclude, increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) could contribute to greater resources and better organization of knowledge in the idea refinement process, and thus have a positive impact on idea quality (Freeman 1979; Grewal et al. 2006). In short, we hypothesize as follows:

Hypothesis 1 (H1): *Increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) in idea refinement will positively influence the idea quality.*

Moderating Effect of Idea Topic Breadth and Depth

We then consider the moderating role of idea topic breadth and depth in the impact of changes to the network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) of the citation network. Previous literature (Dane 2010; Mannucci and Yong 2018; Taylor and Greve 2006) highlights how idea topic breadth and depth impact the structure of domains and schemas within the idea itself and their combinations, suggesting that both idea topic breadth and depth will probably influence how the idea content is generated through the structural organization of the ideas cited, moderating the relationship between the changes of its citation network embeddedness and its quality.

Idea topic breadth is the diversity of the content in the idea; i.e., the number of different domains and schemas in the idea (Mannucci and Yong 2018; Taylor and Greve 2006). While idea diversity is associated with greater novelty and then higher perceived quality (Björk 2012; Kaplan and Vikili 2015), too much novelty may lead to feelings of uncertainty and striking a balance between novelty and conventionality may achieve better perceptions of idea quality instead (Toubia and Netzer 2017; Deichmann et al. 2020). When idea topic breadth is greater, the idea will have more domains and schemas within the idea itself, which might generate information load (Gavetti et al. 2005; Mannucci and Yong 2018). And thus, the idea will carry a greater information load when trying to create linkages between a greater number of existing domains and schemas and new domains and schemas outside the idea brought by increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality), harming the positive effect of increasing citation network embeddedness (Mannucci and Yong 2018; Taylor and Greve 2006). In short, we hypothesize as follows:

Hypothesis 2 (H2): *The positive effect of increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) on the idea quality in idea refinement is weaker when the idea topic breadth is greater.*

Idea topic depth is the degree to which an idea concentrates on the main topic (Mannucci and Yong 2018; Resch and Kock 2021). When idea topic depth is greater, the idea will have more corresponding linkages within domains and schemas, which could contribute to making more proficient use of the knowledge and resources it has (Gavetti et al. 2005; Mannucci and Yong 2018). And thus, the positive effect of increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) through involving more and better knowledge and resources could be amplified when the idea topic depth is greater. In short, we hypothesize as follows:

Hypothesis 3 (H3): *The positive effect of increasing citation network embeddedness (degree centrality, betweenness centrality, and eigenvector centrality) on the idea quality in idea refinement is stronger when the idea topic depth is greater.*

Data and Variables

Research Context and Data Collection

Our research focuses on the context of TVTropes, an online crowd ideation wiki platform for recognizing and documenting new trope ideas i.e., recurrent meaningful patterns in storytelling. New trope ideas (henceforth referred to as TLPs) are proposed via the Trope Launch Pad sub-wiki, which are subject to the scrutiny of the wiki community. For each TLP, there is a working title, synopsis, and a description of the idea, usually containing citations of other (already launched) TVTropes pages. Tropes are cited to establish the TLP's relative position in the trope ontology, to compare or contrast similar or different tropes, or to elaborate the TLP's content with examples. The community interacts with the TLP through comments in the discussion section and casting anonymous positive or negative votes.

Using a Python scraper, we collected daily vote data for TLP tropes proposed between 14th May 2020 to 10th March 2021, as well as their edit and discussion timelines. Using snowball sampling, we recursively collected the tropes cited by TLPs, and the tropes cited by the already cited tropes to generate the TLP citation network. Our samples consist of 1,198 TLP tropes among which, 581 were launched, 318 were discarded, and 299 remained proposed during the above period.

Dependent Variable –Idea Quality Perception

The crowd's perception of the quality of the TLP idea is reflected in their voting behavior. Users cast either a positive or negative vote, and the final voting result determines whether the idea is launched or discarded. We use both positive and negative votes in our model as indicators of idea quality.We also calculated the positive vote ratio using equation (1).

$$PosVoteRatio_i = \frac{PosVote_i+1}{PosVote_i+NegVote_i+2} \tag{1}$$

Independent Variables

Network Embeddedness

Using the TLP citation network, we measure the network embeddedness via various centrality measures; in particular, degree centrality, betweenness centrality, and eigenvector centrality. We calculate the measures based on the undirected TLP citation networks and the bidirectional relationships between trope ideas in the global trope network. Degree centrality of the TLP i measures the number of connections between the TLP and the trope ideas on the citation network, and is calculated using equation (2):

$$C_D(i) = \frac{\sum_{j=1}^{N} a_{ij}}{N-1}, \text{ where } i \neq j \tag{2}$$

Where N is the number of tropes in the citation network, and a_{ij} is 1 if there is a link between TLP i and trope j, and 0 if otherwise. Betweenness centrality calculated by equation (3) is the number of shortest paths between tropes that go through TLP i:

$$C_B(i) = \sum_{i \neq j \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \tag{3}$$

Where σ_{jk} is the number of shortest paths between tropes j and k, and $\sigma_{jk}(i)$ is the number of shortest paths between j and k that pass through i. Betweenness centrality of TLP i measures how the addition of TLP could alter an existing trope network by creating bridges between distantly or unconnected tropes. Eigenvector centrality of TLP i measures how well i is connected to highly influential tropes, and is calculated in equation (4) as follows:

$$C_E(i) = \frac{1}{\lambda} \sum_{j \in G} a_{ij} C_E(j), \tag{4}$$

Where λ is the largest eigenvalue, and $M(i)$ is the set of adjacent tropes to TLP i, and G is the TLP citation network.

Topic Breadth and Depth

To compute the topic breadth and depth of TLPs, we use Latent Dirichlet Allocation (LDA), a natural language processing technique for probabilistic topic modelling (Blei et al. 2003). LDA assumes that each document is a mixture of topics generated from a Dirichlet distribution. In our context, each document is the description of a TLP or trope that is directedly cited by the TLP.

We first extract the edit timelines for each TLP or trope i at time t into a corpus of documents. Next, we preprocessed the text of each document by tokenization and removing the irrelevant wiki text markup, stop words (such as "a", "the", "of") and tokens containing only single characters as they have no significant semantic value for topical categorization. Third, to determine the optimal number of topics for each TLP, we ran the LDA model for 5 to 95 topics, calculating the respective coherence and perplexity scores for each model. Based on the trade-off between coherence and perplexity scores, we selected the LDA model with 20 topics for analysis. To confirm the meaningfulness of each topic, we manually went through the top 20 words of each topic. Finally, we calculated the breadth and depth of each TLP using the 20-topic LDA model. The breadth of TLP i is operationalized as the total number of topics assigned to TLP. The depth of the TLP i is operationalized as the score of its domain and chose the assigned topic with the highest score as its domain based on existing literature (e.g., Mannucci and Yong 2018; Resch and Kock 2021).

Model Specification and Preliminary Results

To quantify the impact of citation network embeddedness changes in idea refinement on its quality, we use a panel data approach for crowdsourcing idea during the refinement process. We specify the model in Equation (5) to test H1 and the model in Equation (6) to test H2 and H3.

$$\Delta\, PosRatio_{it} = \beta_0 + \boldsymbol{\beta}_{\mathbf{e}} \boldsymbol{\Delta\, Embeddedness_{it}} + \beta_r PosRatio_{i(t\text{-}1)} + \boldsymbol{\gamma\, Controls_{it}} + \theta_t + \theta_i + \epsilon_{it} \quad (5)$$

$$\Delta\, PosRatio_{it} = \beta_0 + \boldsymbol{\beta}_{\mathbf{e}} \boldsymbol{\Delta\, Embeddedness_{it}} + \boldsymbol{\beta_{eb}} \boldsymbol{\Delta\, Embeddedness_{it}} \times TopicBreadth_{it} + \boldsymbol{\beta_{ed}} \boldsymbol{\Delta\, Embeddedness_{it}} \times TopicDepth_{it} + \beta_r\, PosRatio_{i(t\text{-}1)} + \boldsymbol{\gamma\, Controls_{it}} + \theta_t + \theta_i + \epsilon_{it} \quad (6)$$

To further explore the mechanism of the impact, we specify the models in Equation (7) and (8) to explore how citation network embeddedness changes impact positive votes and negative votes, respectively. The models in Equation (9) and (10) are used to explore the moderating roles of idea topic breadth and depth. We allow There may exist unobservables that can affect both positive votes and negative votes We choose simultaneous equations models when estimating (7) and (8), and .

$$Ln[E\ (\Delta PosVote_{it} \mid X_{it}, \beta_p^{(+)})] = \beta_0^{(+)} + \boldsymbol{\beta}_{\mathbf{e}}^{(+)} \boldsymbol{\Delta\, Embeddedness_{it}} + \beta_{nvt}^{(+)} \Delta NegVote_{it} + \beta_{pvpt}^{(+)} PosVote_{i(t\text{-}1)} + \boldsymbol{\gamma}^{(+)} \boldsymbol{Controls_{it}} + \theta_t^{(+)} + \theta_i^{(+)} + e^{(+)}{}_{it} \quad (7)$$

$$Ln[E\ (\Delta NegVote_{it} \mid X_{it}, \beta_p^{(-)})] = \beta_0^{(-)} + \boldsymbol{\beta}_{\mathbf{e}}^{(-)} \boldsymbol{\Delta\, Embeddedness_{it}} + \beta_{pvt}^{(-)} \Delta PosVote_{it} + \beta_{nvpt}^{(-)}\, NegVote_{i(t\text{-}1)} + \boldsymbol{\gamma}^{(-)} \boldsymbol{Controls_{it}} + \theta_t^{(-)} + \theta_i^{(-)} + e^{(-)}{}_{it} \quad (8)$$

$$Ln[E\ (\Delta PosVote_{it} \mid X_{it}, \beta_p^{(+)})] = \beta_0^{(+)} + \boldsymbol{\beta}_{\mathbf{e}}^{(+)} \boldsymbol{\Delta\, Embeddedness_{it}} + \boldsymbol{\beta}_{\boldsymbol{eb}}^{(+)} \boldsymbol{\Delta Embeddedness_{it}} \times TopicBreadth_{it} + \boldsymbol{\beta}_{\boldsymbol{ed}}^{(+)} \boldsymbol{\Delta\, Embeddedness_{it}} \times TopicDepth_{it} + \beta_{nvt}^{(+)} \Delta NegVote_{it} + \beta_{pvpt}^{(+)}\, PosVote_{i(t\text{-}1)} + \boldsymbol{\gamma}^{(+)} \boldsymbol{Controls_{it}} + \theta_t^{(+)} + \theta_i^{(+)} + e_{it}^{(+)} \quad (9)$$

$$Ln[E\ (\Delta NegVote_{it} \mid X_{it}, \beta_p^{(-)})] = \beta_0^{(-)} + \boldsymbol{\beta}_{\mathbf{e}}^{(-)} \boldsymbol{\Delta\, Embeddedness_{it}} + \boldsymbol{\beta}_{\boldsymbol{db}}^{(-)} \boldsymbol{\Delta\, Embeddedness_{it}} \times TopicBreadth_{it} + \boldsymbol{\beta}_{\boldsymbol{ed}}^{(-)} \boldsymbol{\Delta\, Embeddedness_{it}} \times TopicDepth_{it} + \beta_{pvt}^{(-)}\, \Delta PosVote_{it} + \beta_{nvpt}^{(-)}\, NegVote_{i(t\text{-}1)} + \boldsymbol{\gamma}^{(-)} \boldsymbol{Controls_{it}} + \theta_t^{(-)} + \theta_i^{(-)} + e_{it}^{(-)} \quad (10)$$

$\Delta\, PosRatio_{it}$ is the daily change in the positive ratio of TLP *i* on day *t* compared to day *t-1*, which indicates the idea quality. $\boldsymbol{\Delta\, Embeddedness_{it}}$ is a three-dimensional vector representing the daily change in network embeddedness of TLP *i* on day *t* from day *t-1,* consisting of degree centrality ($\Delta DegrCentrality_{it}$), betweenness centrality ($\Delta BetwCentrality_{it}$) and eigenvector centrality ($\Delta EigeCentrality_{it}$). $TopicBreadth_{it}$ and $TopicDepth_{it}$ are the topic breadth and depth of TLP *i* on day *t.* In addition, the control variables $\boldsymbol{Controls}_{it}$ identify heterogeneity in time-varying TLP features including other citation network features and content features, and the fixed effect θ_t and θ_j in the above equations identifies heterogeneity between TLPs and days.

Our estimated sample size is 26,883. The descriptive statistics show that, on average, ideas have the largest eigenvector centrality (mean = 0.007; std dev. = 0.020), followed by betweenness centrality (mean = 0.002; std dev. = 0.005), and finally, degree centrality (mean = 0.001; std dev. = 0.001). In addition, the average number of positive and negative votes are 4.387 (std dev. = 5.100) and 2.395 (std dev. = 3.571), and the average idea topic depth and breadth are 0.285 (std dev. = 0.007) and 13.178 (std dev. = 4.700), respectively.

As is shown in Table 1, Model 1 demonstrates the effect of citation network embeddedness changes on its quality. Consistent with H1, increasing degree centrality (β = 41.974, p < 0.01) and eigenvector centrality (β = 0.861, p < 0.01) have a significant positive impact on positive ratio. H1 is supported. Also, the results show that the influence of citation network embeddedness varies across three different measures, suggesting the criticality of further research at the subconstruct level. And this finding is also consistent with prior literature on network embeddedness (e.g., Gavetti et al. 2005).

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Dependent Variable	*Δ* $PosRatio_{it}$	*Δ* $PosRatio_{it}$	*Δ* $PosVote_{it}$	*Δ* $NegVote_{it}$	*Δ* $PosVote_{it}$	*Δ* $NegVote_{it}$
$\Delta DegrCentrality_{it}$	41.974*** (15.124)	209.930*** (49.084)	289.457** (123.097)	-174.556* (98.034)	339.611 (401.350)	-556.282* (319.680)
$\Delta BetwCentrality_{it}$	0.099 (1.190)	-8.014 (6.617)	-9.280 (9.696)	-6.887 (7.721)	-82.171 (50.171)	-51.548 (39.938)
$\Delta EigeCentrality_{it}$	0.861*** (0.136)	3.334*** (0.912)	3.271*** (1.104)	-1.916** (0.878)	15.320** (7.422)	-7.988 (5.909)
$\Delta DegrCentrality_{it} \times TopicBreadth_{it}$		-10.438*** (2.896)			-2.275 (23.684)	34.880* (18.870)
$\Delta BetwCentrality_{it} \times TopicBreadth_{it}$		0.180 (0.377)			0.701 (3.069)	3.068 (2.443)
$\Delta EigeCentrality_{it} \times TopicBreadth_{it}$		-0.159*** (0.051)			-0.492 (0.415)	-0.424 (0.330)
$\Delta DegrCentrality_{it} \times TopicDepth_{it}$		-0.138** (0.057)			-0.142 (0.464)	-0.168 (0.370)
$\Delta BetwCentrality_{it} \times TopicDepth_{it}$		0.024*** (0.007)			0.248*** (0.055)	0.004 (0.044)
$\Delta EigeCentrality_{it} \times TopicDepth_{it}$		0.002* (0.001)			0.278*** (0.009)	0.012 (0.008)

Notes: *** p < 0.01, ** p < 0.05, * p < 0.1. Control variables are included but not shown in the table.

Table 1. Effect of Citation Network Embeddedness Changes on Idea Quality

Model 2 reports the moderating effect of idea topic breadth and depth. As shown in Model 2, the positive effect of increasing citation degree centrality (β = -10.438, p < 0.01) and eigenvector centrality (β = -0.159, p < 0.01) is weaker when the idea topic breadth is greater. H2 is supported. Consistent with H3, the positive effect of increasing citation betweenness centrality (β = 0.024, p < 0.01) and eigenvector centrality (β = 0.002, p < 0.10) is stronger when idea topic depth is greater. However, the positive effect of increasing citation degree centrality (β = -0.138, p < 0.05) is stronger when the idea topic depth is greater. H3 is partially supported.

Model 3 and 4 explores the mechanisms behind the effects of citation network embeddedness changes. Results show increasing degree centrality (positive: β = 289.457, p < 0.05; negative: β = -174.556, p < 0.1) and eigenvector centrality (positive: β = 3.271, p < 0.01; negative: β = -1.916, p < 0.05) significantly increase positive votes and decrease negative votes, which suggests increasing citation network embeddedness could increase highlights and improve deficiencies, supporting H1.

Model 5 and 6 explore the mechanisms behind the moderating role of idea topic breadth and depth. The results show idea topic breadth reduces the negative effect of increasing degree centrality on negative votes (β = 34.880, p < 0.10), and then reduces the positive effect of increasing degree centrality on the overall positive ratio. And idea topic depth increases the positive effect of increasing network embeddedness on positive votes (betweenness: β = 0. 248, p < 0.01; eigenvector: β = 0.278, p < 0.01), and then increases the positive effect of increasing network embeddedness on the total positive ratio supporting H2 and H3.

To further validate the robustness of our analysis, we conduct robustness checks by altering the measurement of variables. First, we change the measure of dependent variable and use the original positive ratio. Second, we lag the dependent variable data by one day and estimate the model using both raw and log-transformed data, to control for the possibility of reverse causation. These analyses yield similar results, indicating the reliability of our results.

Discussion and Next Steps

Our preliminary results provide potential practical implications in innovation management and ideation. First, our findings suggest that network embeddedness does influence the idea quality perception. Apart from betweenness centrality, increases in both degree centrality and eigenvector centrality have positive and significant effects on idea quality perception. This suggests that ideators and managers should pay attention to how the structural connections between the set of ideas may be reconfigured when a new idea based upon them is proposed. Second, our results support the idea that the level of knowledge specialization moderates the effect of network embeddedness on the perceived idea quality. More specifically, the positive effect is stronger when the idea topic breadth is smaller, or when the idea topic depth is greater. Finally, the effects of network embeddedness and topic specialization on change of positive votes and negative votes indicate that the influence path through changes of positive and negative votes might be different. This suggests that ideators should prioritize depth-focus exploration instead of focusing on expanding the idea scope.

While our initial results are promising, there may still exist endogeneity issues in our model such as other unobserved factors that may confound the influence the network embeddedness. Second, the topic breadth and depth of the tropes cited by the TLPs and how they may influence the idea quality perception of TLPs. Third, further exploration on the influence paths of the effects can be conducted.

References

Betancourt, G. G., Segnine, A., Trabuco, C., Rezgui, A., and Jullien, N. 2016. “Mining Team Characteristics to Predict Wikipedia Article Quality,” in *Proceedings of the 12th International Symposium on Open Collaboration - OpenSym ’1*6, Berlin, Germany: ACM Press, pp. 1–9.

Blei, D. M., Andrew Y. N., and Michael I. J. 2003. "Latent Dirichlet Allocation," *Journal of Machine Learning Research* (3). pp. 993-1022.

Björk, J. 2012 “Knowledge Domain Spanners in Ideation: Knowledge Domain Spanners in Ideation.” *Creativity and Innovation Management* (21:1), pp. 17–27.

Björk, J., and Magnusson, M. 2009 “Where Do Good Innovation Ideas Come From? Exploring the Influence of Network Connectivity on Innovation Idea Quality.” *Journal of Product Innovation Management* (26:6) pp. 662–670.

Burt, R. S. 2004. “Structural Holes and Good Ideas,” *American Journal of Sociology* (110:2), pp. 349–399.

Dalip, D. H., Gonçalves, M. A., Cristo, M., and Calado, P. 2011. “Automatic Assessment of Document Quality in Web Collaborative Digital Libraries,” *Journal of Data and Information Quality* (2:3), pp. 1–30.

Dane, E., 2010. “Reconsidering the Trade-Off Between Expertise and Flexibility: A Cognitive Entrenchment Perspective,” *Academy of Management Review* (35:4), pp. 579–603.

Dargahi, R., and Namin, A. 2020. “Making More in Crowdsourcing Contests: A Choice Model of Idea Generation and Feedback Type,” *Journal of Marketing Communications*, pp. 607–630.

De la, C., Gabriel., and Alex, D. 2010. “On Measuring the Quality of Wikipedia Articles,” In *Proceedings of the 4th Workshop on Information Credibility - WICOW ’10*, 11. Raleigh, North Carolina, USA: ACM Press.

Deichmann, D., Christine, M., Julie, M. B., Adina, N., Peter, G., and Shenghui, W. 2020. “Ideas with Impact: How Connectivity Shapes Idea Diffusion,” *Research Policy*(49:1), pp. 1-11.

Dennis, Alan R., Joseph S. Valacich, Traci A. Carte, Monica J. Garfield, Barbara J. Haley, and Jay E. Aronson. 1997. “Research Report: The Effectiveness of Multiple Dialogues in Electronic Brainstorming,” *Information Systems Research* (8:2), pp. 203–211.

Fan, David P. 1985. “Ideodynamics: The Kinetics of the Evolution of Ideas.” *The Journal of Mathematical Sociology* (11:1) pp. 1–23.

Garfield, Monica J., Nolan J. Taylor, Alan R. Dennis, and John W. Satzinger. 2001. "Research Report: Modifying Paradigms—Individual Differences, Creativity Techniques, and Exposure to Ideas in Group Idea Generation," *Information Systems Research* (12:3), pp. 322–333.

Grewal, R., Lilien, G. L., and Mallapragada, G. 2006. "Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems," *Management Science* (52:7), pp. 1043-1056.

Huang, Y., Vir Singh, P., and Srinivasan, K. 2014. "Crowdsourcing New Product Ideas Under Consumer Learning," *Management Science* (60:9), pp. 2138–2159.

Huang, N., Gordon, B., Bin, G., Yili, H., Chen, L., Kanliang, W., Dongpu, F., and Bo, Y. 2019. "Motivating User-Generated Content with Performance Feedback: Evidence from Randomized Field Experiments," *Management Science*(65:1), pp. 327–345.

Jung, J. H., Christoph, S., and Joseph, V. 2010. "Enhancing the Motivational Affordance of Information Systems: The Effects of Real-Time Performance Feedback and Goal Setting in Group Collaboration Environments," *Management Science* (56:4), pp. 724–742.

Kaplan, S., and Vakili, K. 2015. "The Double-Edged Sword of Recombination in Breakthrough Innovation: The Double-Edged Sword of Recombination." *Strategic Management Journal* (36:10) pp. 1435–1457.

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.

Mannucci, P. V. , and Yong, K. 2018. "The Differential Impact of Knowledge Depth and Knowledge Breadth on Creativity over Individual Careers," *Academy of Management Journal* (61:5) , pp. 1741-1763.

Perry-Smith, J. E., and Mannucci, P. V. 2017. "From Creativity to Innovation: The Social Network Drivers of the Four Phases of the Idea Journey," *Academy of Management Review* (42:1), pp. 53–79.

Piezunka, H., and Dahlander, L. 2019. "Idea Rejected, Tie Formed: Organizations' Feedback on Crowdsourced Ideas," *Academy of Management Journal* (62:2), pp. 503–530.

Ransbotham, and Kane. 2011. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia," *MIS Quarterly* (35:3), pp. 613-628.

Ransbotham, S., Kane, G., Lurie, N.H. 2012. "Network Characteristics and the Value of Collaborative User Generated Content," *Marketing Science* (31:3), pp.387–405.

Resch, C., and Kock, A. 2021. "The Influence of Information Depth and Information Breadth on Brokers' Idea Newness in Online Maker Communities," *Research Policy* (50:8), pp. 1-12.

Satzinger, J. W., Monica J. G., and Murli N. 1999. "The Creative Process: The Effects of Group Memory on Individual Idea Generation," *Journal of Management Information Systems* (15:4), pp. 143–160.

Seeber, I., Gert-Jan de V., Ronald, M., and Barbara, W. 2017. "Beyond Brainstorming: Exploring Convergence in Teams," *Journal of Management Information Systems* (34:4), pp. 939–969.

Taylor, A., and Greve, H. R. 2006. "Superman or the Fantastic Four? Knowledge Combination and Experience in Innovative Teams," *Academy of Management Journal* (49), pp. 723–740.

Toubia, O., and Netzer, O. 2017. "Idea Generation, Creativity, and Prototypicality." *Marketing Science* (36:1) pp. 1–20.

Wang, C., Rodan, S., Fruin, M., and Xu, X. 2014. "Knowledge Networks, Collaboration Networks, and Exploratory Innovation." *Academy of Management Journal* (57:2) pp. 484–514.

Wang, K., and Nickerson, J. V. 2019. "A Wikipedia-Based Method to Support Creative Idea Generation: The Role of Stimulus Relatedness." *Journal of Management Information Systems* (36:4) pp. 1284–1312.

Zhang, X., and Wang, C. 2012. "Network Positions and Contributions to Online Public Goods: The Case of Chinese Wikipedia," *Journal of Management Information Systems* (29:2), pp. 11–40.