1-10-2016

Anchored Discussion: Development of a Tool for Creativity in Online Collaboration

Georg J.P. Link
University of Nebraska at Omaha, glink@unomaha.edu

Dominik Siemon
Technische Universität Braunschweig, d.siemon@tu-bs.de

Gert-Jan de Vreede
University of South Florida, Tampa, gdevreede@usf.edu

Susanne Robra-Bissantz
Technische Universität Braunschweig, s.robra-bissantz@tu-bs.de

Follow this and additional works at: https://digitalcommons.unomaha.edu/isqafacpub
Part of the Computer Engineering Commons, and the Computer Sciences Commons

Recommended Citation
Link, Georg J.P.; Siemon, Dominik; de Vreede, Gert-Jan; and Robra-Bissantz, Susanne, "Anchored Discussion: Development of a Tool for Creativity in Online Collaboration" (2016). Information Systems and Quantitative Analysis Faculty Publications. 49.
https://digitalcommons.unomaha.edu/isqafacpub/49
Abstract: Open innovation and crowdsourcing rely on online collaboration tools to enable dispersed people to collaborate on creative ideas. Research shows that creativity in online groups is significantly influenced by the interaction between group members. In this paper, we demonstrate how theory can be effectively used to design and evaluate a tool for creative online collaboration. Anchored discussions represent a new mode for creative interaction. In anchored discussion every comment is tied to some aspect of an idea. We evaluated the anchored discussion tool in a laboratory experiment, which generated insights for additional and refined research. Our results indicate that anchored discussion leads to a more structured discussion amongst group members and consequently to more creative outcomes. In a post session survey, participants made several suggestions on how to improve anchored discussion. This paper concludes that anchored discussion is promising as a new tool to aid online groups in creative collaboration. This paper extends a previous version presented at CRIWG 2015 [Link, 2015].

Keywords: theory-driven software design, creativity support systems, anchored discussion, idea generation, ideation, creativity, online collaboration, Team Creativity Model

Category: L.6.2

1 Introduction

Modern companies operate in a competitive and ever changing environment. New solutions are needed to meet the new paradigm of global competitiveness, which leads to rapid innovation in entirely new ways, e.g. [Levy, 2015]. The innovation process relies on the ability to generate, evaluate, and refine ideas. Therefore companies are constantly searching for ways to develop innovative ideas.
One approach to seek and develop innovative ideas concerns open innovation, which is the purposeful usage of internal and external sources for ideas [Chesbrough, 2006]. More generally speaking, ideas can originate from a company’s employees, suppliers, partners, customers, or other external parties. One way to allow these people to collaboratively innovate and make the resulting ideas accessible to the organization is the usage of online collaboration systems. These systems have the potential to evoke and enhance the creativity of a diverse set of participants [Van Osch, 2009].

Organizations that design and deploy online collaboration systems should adopt a theory-driven design method and account for the many factors that influence the creativity of individuals and groups [e.g. Amabile, 1983; Vreede, 2012] so that they stimulate rather than discourage the creative process [Briggs, 2006]. This is, for example, evident in the area of small group brainstorming. Early research demonstrated that blocking effects significantly reduced group productivity and creativity in traditional brainstorming groups where only one person can talk at a time [Diehl, 1987]. Information System researchers developed collaboration systems where group members generate ideas simultaneously leading to higher levels of creativity and productivity [Fjermestad, 1998].

A key challenge in open innovation and crowdsourcing is to enable online groups to generate creative ideas. Many open innovation systems have features for idea competitions or for collecting ideas straightforwardly, yet little support is offered to evaluate and refine ideas as a group effort [Hrastinski, 2010]. However, when people collaborate and improve each other’s ideas, the resulting ideas are likely to be more creative [Blohm, 2010; Yu, 2011]. This is evident in an organization’s use of a collaboration tool where engineers collaborate with employees from the field (e.g. support technicians who observe customers’ use of products) in the creative task of developing products that better suit customers’ needs [Bertoni, 2011].

We propose that new tools for online collaboration that are grounded in creativity and ideation theory have the potential to foster group creativity. Specifically, we follow a theory-driven design approach to design a tool that enables us to evaluate online anchored discussions as a mechanism to stimulate creative idea development. Anchored discussion originates from the field of education [Van der Pol, 2006]. Students’ comments are tied to specific sections of the academic text and are shown next to it. Research shows that this design increases collaboration, knowledge sharing, and student engagement [Alrushiedat, 2012]. We propose to replace the academic text with a shared idea editor. Generated ideas can then be the center of an anchored discussion, which has the potential to stimulate a more structured collaboration process for elaborating and evaluating ideas.

## 2 Theoretical Foundation

### 2.1 Creativity in Groups

According to [Amabile, 1983, p. 33] “a product or response will be judged as creative to the extent that it is (a) both a novel and appropriate, useful, correct, or valuable response to the task at hand, and (b) the task is heuristic rather than algorithmic in nature.” Research on creativity is often grouped into the Four P’s of Creativity:
person, process, product, and press [Rhodes, 1961]. Early creativity research focused on describing attributes of creative people. More recent research focuses on the creative process and its products. [Mumford, 2012] provide a model of the creative process, which consists of eight core processes: (a) problem definition, (b) information gathering, (c) information organization, (d) conceptual combination, (e) idea generation, (f) idea evaluation, (g) implementation planning, and (h) solution monitoring. These eight core processes, which we will refer to as phases, can occur in different orders and loops due to the heuristic nature [Amabile, 1983] of the creative process. Especially within a group, it can be assumed that several of these phases occur simultaneously in different people.

For example, [Sawyer, 2009] found that the interaction among group members can be a substantial source of creativity. This is evident in the reported case of a theater play in which actors construct the narrative in taking turns. One actor can make a new proposal for a development but the actual development of the narrative depends on how the next actor chooses to respond and what parts of the proposal they choose to continue. [Sonnenburg, 2007] specifically views creativity as something that results from the interaction of many factors. He reframed the 4P’s of creativity research and developed the 1-5P-Model which describes the creative potential to be influenced by the factors: person, place, process, problem, and (proto-)prototype. The (proto-)prototype is the work-in-progress or preliminary outcome of the group process, which is transformed into the actual prototype - the creative outcome. From this point of view, the group members represent one factor (person) that influences the potential for creativity. Sonnenburg recognized that only people can start the creative process, but their interaction is influenced by other factors.

The interaction amongst group members themselves is a key source of creative potential. This is demonstrated by the Team Creativity Model (TCM), which was developed in an exploratory field study [Vreede, 2012]. According to TCM, the two main antecedents to group creativity are individual creativity and shared mental model. Especially the shared mental model is a direct result from the interaction of the group members. A shared mental model represents the extent to which group members have a shared understanding of the group situation and the task [Cooke, 2000]. Shared understanding is the result of knowledge sharing, which requires clear communication [Bischof, 2011]. Clear communication is concise, has a logical structure, provides an explicit context and is low in ambiguity. To increase shared understanding, the interaction should be focused on a goal and aimed to help group members understand. Therefore, the Team Creativity Model proposes that shared mental model acts as a mediator between knowledge sharing (i.e. interaction) and group creativity.

Thus, following the logic of TCM, it is important to support diversity and independence to evoke creativity in individual participants [Van Osch, 2009]. Also, it is essential to lower the barriers for interaction and knowledge sharing between participants [Lugger, 2001] to enhance the formation of shared mental models. Based on past research and the logic of TCM we therefore hypothesize:

**Hypothesis 1a:** Group members reporting more similar shared mental models will have more creative products.

**Hypothesis 1b:** Groups with a more structured and goal oriented interaction will report higher similarity in their shared mental model.
2.2 Individual Creativity and Group Creativity

According to TCM, individual creativity is a key antecedent to group creativity. It is also influenced by the interaction of the group members. This can be explained with the search for ideas in associative memory (SIAM) model [Nijstad, 2006]. SIAM assumes that links between ideas in people’s memory exist and that usually only ideas with a strong link to current thoughts are activated. A creative thought is characterized by two distinct, simultaneously active ideas that were not or only loosely linked before [Santanen, 2004]. This activation can result from the interaction within the group or other outside stimulation.

A common form of stimulation is to change perspective, which can occur in one of three ways [Knoll, 2010]: (a) by searching for similar situations (analogy) and generating ideas for those situations; (b) by challenging assumptions (provocation) and using resulting situations for idea generation; or c) by using random elements and knowledge about them to generate ideas. Due to the social norm of not changing the subject and the tendency of building on other ideas, people have the tendency to voice similar ideas [Satzinger, 1999] and to not explore changes of perspective. Paradigm preserving (similar) statements elaborate on previous ideas and have value in building understanding. However, novel ideas are more likely to be paradigm modifying by introducing new elements or explore different relationships.

It can therefore be argued that in the context of a creativity task the goal is to gain paradigm modifying ideas. Groups can be stimulated to explore a wider variety of ideas when giving diverse stimuli, for example from a computer algorithm [Siemon, 2014] or from a facilitator [Connolly, 1993; Santanen, 2004]. The latter intervention method was informed by the Cognitive Network Model of Creativity, which draws a connection between the stimuli, cognitive load, and resulting creative thoughts [Santanen, 2000]. Simply put, the model extends SIAM and posits that creative thought results from the combination of previously unlinked ideas, which is aided by a variety of stimuli, but impeded if the individual experiences high cognitive load. Cognitive load increases when many different ideas are active at once but cannot be combined into a chunk. Giving stimuli increases group creativity provided the prompts are well timed [Santanen, 2004].

The research by Santanen and colleagues applies to co-located groups that have access to dedicated facilitation support. We are interested in determining how online groups can be supported without a professional facilitator who provides stimuli. When a group uses an electronic system to generate ideas, then only ideas that the group generated earlier can serve as stimuli [Satzinger, 1999]. In terms of Sonnenburg’s 1-5P-Model the (proto-)prototype is a collection of all previous ideas [Sonnenburg, 2007]. The way in which the system displays previous ideas can influence the creative process [Javadi, 2013]. Attention diversion and lack of attention to other people’s ideas can both be the result of excessive exposure to other people’s ideas. Following the logic of the Cognitive Network Model of Creativity, this is due to the increase in cognitive load among the group members. Javadi et al. propose to limit the number of displayed ideas based on a ranking derived from user votes. The same effect could be accomplished through a differently designed user interface that would separate different aspects of a problem into different dialogs (similar to different pages of a book) and thus separating the discussion [Dennis, 1997]. Splitting idea generation into subcategories resulted in the generation of more ideas, more
high-quality ideas, and more novel ideas. However, if a complex problem has multiple areas that could be subdivided but have a high interdependence, then the dependencies might not be considered if the problem is only viewed in subcategories and never as a whole.

In summary, it can be argued that the way a group interacts will impact individual creativity and group creativity. The interaction is a stimulus for individual creativity. As a result individuals share their ideas and thereby stimulate creative thoughts in other individuals and overall impact group creativity. A better-structured and more goal-oriented interaction can reduce cognitive load and stimulate more diverse ideas. We thus formulate the following hypotheses:

Hypothesis 2a: The higher the creativity of individual group member, the higher the group creativity.

Hypothesis 2b: The more structured and goal-oriented the interaction among group members, the higher the individual creativity of the group members.

2.3 Anchored Discussion for Creative Online Collaboration

When a group uses an electronic system to collaborate, the design of the system shapes their interaction. Group members collaborating from separate locations and interacting only through a digital collaboration system are discouraged from social interaction but their writing efficiency may be boosted [Shah, 2015]. Early electronic brainstorming research demonstrated that the negative effects of face-to-face groups, such as production blocking and evaluation apprehension [Diehl, 1987, 1991], could be mitigated. Using electronic brainstorming can also enable larger groups to be productive, especially because they need not wait for each other to finish voicing ideas [Dennis, 1993; Gallupe, 1992]. To achieve positive results with such collaboration technology, designers have to ensure that the user interface supports the specific creative task. For example, an online group that works on the interconnection of ideas benefits from a diagram based collaboration tools that seamlessly links textual and graphical information [Azevedo, 2013].

We build on previous research for designing a tool for creative online collaboration. Recently, [Voigt, 2013] conducted a literature review to compile an integrated framework for designing group creativity support systems. To aid the development of group creativity support systems the framework provides 13 design principles and six components with different tasks within one system. For our research focus, the two components that are tasked to allow users to collaboratively generate ideas are relevant: shared idea editor and communication component. The related design principles are as following: (a) Provide the possibility to share ideas to foster mutual inspiration, (b) provide session histories and dialogue mapping, to support idea reflection and information storage to build trust within the group, and (c) support group awareness to avoid coordination problems and foster reciprocal inspiration. While these design principles address some of our hypotheses, they are not comprehensive. Additional design principles can be derived from the above literature review: (d) Limit the number of displayed ideas to reduce cognitive load [Javadi, 2013], (e) allow discussion and idea generation for different aspects of a problem to be separated [Dennis, 1997], and (f) provide a way to maintain the big picture, to allow groups to consider the relationships between different aspects [Azevedo, 2013].
Based on these design principles, we argue that anchored discussion represents a viable system and user interface design. As explained earlier, anchored discussion originates from the education research literature and aims to allow students to collaboratively process academic literature [Van der Pol, 2006]. The user interface for an anchored discussion displays the article that students discuss on one side and the discussion on the other. “Anchoring is a process of creating reference points between parts of a document and comments in the discussion space to help prevent drifting within the context” [Alrushiedat, 2012, p. 12].

The advantages of anchored discussion are argued to include more meaning-oriented discussion, more frequent referring to content, fewer self-clarifications, fewer words needed to express ideas, increased sharing of ideas, enhanced participation, and improved engagement [Alrushiedat, 2012; Van der Pol, 2006]. Anchored discussion is often compared to a forum discussion, which typically has participants focusing more on establishing social relationships and regulating the collaborative process. Also in forums there are more argumentations and confirmations. While the actual level of shared understanding (shared mental model) may not increase through anchored discussion, researchers claim that less effort is required to reach the same level of shared understanding [Van der Pol, 2006]. This is because anchored discussion increases the efficiency and clarity of communication.

To the best of our knowledge, there is no prior research on anchored discussions to stimulate creativity in problem solving tasks. We suggest to replace the academic literature frame with a shared idea editor. The original idea would always be present in the shared idea editor, so that the big picture is not lost. At the same time, using anchored discussion, separate subcategories of the overall problem can be discussed without interference of discussion of other subcategories. In other words, an anchored discussion for online idea generation will consist of a text editor and an area for comments. In the comments area only comments related to the currently selected text in the editor would be displayed. Thus, in-depth discussions can occur separately for each aspect of the whole idea. Additionally, the desired change of perspective occurs every time a user considers comments tied to a different aspect of an idea. The group would make changes to the original idea, based on the discussion. All comments should remain accessible, regardless of changes to the text in the shared idea editor. Anchored discussion solves issues of clear communication, specifically related to the following [Bischof, 2011]:

- Writing concise content. Participants have to identify the most important aspects, leave out what is not required, and find simple and easily understood language. Anchored discussion focuses detailed discussion (including history) on only the one element of an idea that is in focus and thus encourages participants to rework this element based on the discussion.
- Building logical structure. This is a matter of how parts need to be connected. Anchored discussion always provides a connection between comments and elements of an idea.
- Providing an explicit context for discussion. By showing comments related to one element of an idea at a time, the context of the comment is explicit without the participant having to explain the context, thus reducing the effort needed for clear communication.
In summary, anchored discussion reduces cognitive load and fosters the emergence of shared mental models by providing an environment that elicits clear knowledge communication. We propose that anchored discussion is well suited to enable an online group to better structure discussions of larger ideas. Thus, we hypothesize:

Hypothesis 3a: Groups that use anchored discussion will report higher shared mental model scores than groups that do not.

We further propose that anchored discussion reduces cognitive load because it limits the displayed comments to only those relevant to the subcategory a user is currently focusing on. Reduced cognitive load should result in higher levels of individual creativity [Santanen, 2004]. Thus we hypothesize that:

Hypothesis 3b: Groups that use anchored discussion will produce higher levels of individual creativity than groups that do not.

We propose that shared mental model and individual creativity mediate between group interaction and group creativity. Consequently, if anchored discussion improves the group interaction, we hypothesize that group creativity will be improved as well:

Hypothesis 3c: Groups that use anchored discussion will produce higher levels of group creativity than groups that do not.

3 Method

Following the principles of theory-driven design, we designed and developed a technical artifact to evaluate our hypotheses [Briggs, 2006]. Our artifact approaches our defined problem to increase creativity in teams based on the theoretical foundations described above and proposes a solution in terms of adding the anchored discussion functionality to a shared idea editor [Peffers, 2007]. Our web-based artifact is a real-time synchronous editor with functionality that allows a group to work on the same text with all changes immediately visible to all members. The anchored discussion functionality provides the users with the ability to add comments to certain aspects of an idea. In our case, comments were tied to a specific line of the text. When a line of the idea was removed, comments remained accessible by merging them with the next line. Based on a user’s current cursor position, the comments for this line would be visible, as depicted in [Fig. 1]. The actual idea is written within the shared idea editor and comments are arranged in a separate area to the right. The idea and comments are real-time synchronized so that everyone in the group can see changes and new comments immediately.

For the control groups, the prototype was modified by disabling the anchored discussion functionality. Instead, a discussion function was implemented, which is similar to a one-threated forum. It basically represented a normal chat. Therefore, the control groups had no ties between comments and an idea, but still had a way of interacting, reach consensus by building on each other’s ideas, and group learning [Garcia, 2014]. The interface looked the same as in [Fig. 1] with the exception that the comment area behaved like a chat.

The prototype was implemented as an extension of the open source software Firepad. Firepad is a reference implementation for the Firebase database by Firebase Inc., which is currently owned by Google. The MIT License allowed us to use and change Firepad without restrictions. Firepad is a real-time collaborative text editor,
using the Firebase database to store and sync data in real-time. Firebase is available for many platforms, including iOS, Android, and Web. We used an API provided by a JavaScript library for web applications.

3.1 Experiment Structure

A laboratory experiment was conducted with 98 participants (64 male and 32 female). The participants were between the ages of 18 and 30 (M = 24.1, SD = 2.87). All participants were university students in a Western European university. They were undergraduate (44), graduate (52), and post-graduate (2) students with majors in different fields (48 management information systems, 32 industrial engineering, 9 computer science, and 9 others). Students were assigned to one of 26 groups of 3 to 5 participants each.

The task for all groups was to use the prototype to collaborate as board members of a fictitious organization that wants to redesign the city parks. The city council requires the organization to organize volunteers, to collect donations, and to ask the people in the affected neighborhoods for their ideas on the park redesign. The group was provided with a few initial ideas of an absent board member. They were then tasked to refine these ideas and develop additional ideas.

The experiment sessions were structured in three phases. During the introduction phase the participants were randomly assigned to groups and received an introduction into the prototype and the task. The second phase of the experiment was the collaboration phase, which lasted for 45 minutes. In the final phase a survey was issued to capture participants’ perceptions and feedback.
Two treatments were given. The first treatment, Treatment AD, included the prototype with the anchored discussion feature. This treatment was randomly assigned to 13 groups (48 participants). The second treatment, Treatment CD, was our control group and included the prototype with chat discussion. This treatment was also randomly assigned to 13 groups (50 participants).

3.2 Measures

Many researchers use an evaluation of the creative product as a proxy to assess the quality of the creative process [Dean, 2006]. A creative product is “a product or response [that] will be judged as creative” [Amabile, 1983, p. 33]. Since, there is no objective measurement, Amabile [1982] suggested to use subjective measurements by asking judges to rate the creativity based on provided dimensions. She found that judges reliably agree on what can be considered as creative, even without prior training. The dimensions to be rated by the judges can be tailored to the task that the creative outcome resulted from, but certain elements are predominantly present: novelty and quality [Dean, 2006]. Dean et al. reported high inter-judge reliability on different problems when using their ordinal scales on eight different dimensions.

For our study, four independent judges from diverse backgrounds evaluated separately the creativity of each group in random order. The creative outcome was rated on Dean et al.’s eight dimensions (1-4 scale for first six dimensions, 1-3 scale for last two): a) originality, b) paradigm relatedness, c) acceptability, d) implementability, e) applicability, f) effectiveness, g) implicational explicitness, and h) completeness. Dean et al. demonstrated that these dimensions reliably measure novelty (a + b), workability (c + d), relevance (e + f), and specificity (g + h). The overall creativity is a combination of novelty and quality, which is measured by the other three constructs. Dean et al. recommend using thresholds on novelty and the three dimensions of quality, because a strength in one dimension cannot compensate for a weakness in another area. The goal was to compare creativity and thus a thresholds based on the respective means was created which separated the group outcomes into two groups on each dimension. Only when the creative outcome is in the higher rated group on at least three of the four dimensions, was the group considered to have created a creative outcome. The judges also rated how well-structured and goal-oriented the discussion within a group was. These dimensions were informed by the benefits that anchored discussion offers. The means of the dimensions formed an anchored discussion score (AD-score).

A standard survey instrument, based on [Johnson, 2007], was used to capture the shared mental model (SMM). After each session, every participant was asked to rate 26 items (1 = strongly agree, 5 = strongly disagree). A group’s SMM-score was calculated as the mean across participants’ means across all items. The post-experiment survey also assessed participants’ perceptions of effectiveness and satisfaction. These measures were adapted from previous research [Dennis, 1996].

Another construct we measured was individual creativity. A common way to measure creativity is the good-idea count, by rating every submitted idea [Reinig, 2006]. However, the shared idea editor captures entries from every person as changes to the idea set. In other words, the final result represents the final transformation of the idea set. The prototype does not contain an algorithm to keep track of individual contributions or separate contributions. Thus it was not possible to determine who
entered what idea or changed another idea in the text editor. Until we extend our prototype, individual creativity can only be measured by proxy. To this end, it was assumed that individual creativity can be expressed as the number of ideas presented to the group and assuming that this is related to the amount of text a participant enters into the prototype [Briggs, 2007]. Contributions were counted in terms of the number of characters added to the idea editor, the number of changes made in the idea editor, and the number of characters captured in comments.

After the quantitative analysis, the first two authors evaluated the use of the prototype by analyzing the presentation of the final document, the discussions, and comments provided in the post-experiment survey [Thomas, 2006]. We analyzed the use of different structuring elements including headlines, bullet points, and enumerations and their support for clarity and comprehensiveness of the idea. We further analyzed 2084 comments from the discussion within each group, coding for idea-related, organizing and task-unrelated comments. The first and second author coded independently to achieve inter-coder reliability and discussed their differences of judgment [Thomas, 2006]. Lastly, we analyzed five pages of comments provided by 51 participants in the post-experiment survey. We highlighted comments addressing multiple aspects with different colors and recorded emergent themes.

4 Results

An independent samples t-test was conducted to compare the group creative score (sum of creativity dimensions) for groups that reported more similar SMM and groups with less SMM (H1a). We split the groups by median SMM. There was no significant difference in group creativity score for groups with more similar SMM (M=22.6, SD = 1.85) and groups with less similar SMM (M = 22.1, SD = 1.48; t(24) = -.78, p =.44, two tailed). The magnitude of the differences in the means (mean difference = -.51, 95% CI: -1.85-.83) was moderate (eta squared = .02). H1a is not supported.

An independent samples t-test was conducted to compare the reported similarity of SMM for groups that had more structured and goal oriented interaction and groups that had less (H1b). We split the groups by median anchored discussion score (AD-score). The AD-score captures how well structured and goal oriented the discussion within a group was (from 1 = unstructured to 4 = well structured). Inter-judge reliability is high on four items (Cronbach’s alpha: .89). There was no significant difference in reported SMM similarity for groups that had a higher AD-score (M=49.7, SD = 11.2) and groups that had a lower AD-score (M = 50.7, SD = 11.3; t(24) = .23, p =.82, two tailed). The magnitude of the differences in the means (mean difference = 1.01, 95% CI: -8.13-10.16) was small (eta squared = .002). H1b is not supported.

To determine which groups produced a creative outcome, we counted the creativity dimensions for which the rating was above the dimension mean over all groups. Groups that reached at least the mean on three dimensions were considered to be more creative [see Tab. 1]. The inter-judge reliability on group creativity ranged from mediocre to acceptable, given the Cronbach’s alpha for the dimensions novelty (.56), workability (.60), relevance (.51), and specificity (.74). On workability we excluded the ratings of one judge as they showed a clearly different rating scheme.
Table 1: AD groups were rated on more dimensions above means

<table>
<thead>
<tr>
<th></th>
<th>AD</th>
<th>CD</th>
<th>More creative</th>
<th>Less creative</th>
<th>t-test</th>
<th>p</th>
<th>MD</th>
<th>95% CI</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>t</td>
<td></td>
<td>p</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Creativity</td>
<td>22.6</td>
<td>1.66</td>
<td>22.2</td>
<td>1.67</td>
<td>-.68</td>
<td>.50</td>
<td>-.44</td>
<td>-1.79-90</td>
<td>.02</td>
</tr>
<tr>
<td>Individual Creativity</td>
<td>2081</td>
<td>659</td>
<td>2426</td>
<td>563</td>
<td>1.44</td>
<td>.16</td>
<td>345</td>
<td>-181-190</td>
<td>.21</td>
</tr>
<tr>
<td>Similar SMM</td>
<td>55.1</td>
<td>9.1</td>
<td>45.1</td>
<td>10.8</td>
<td>-2.55</td>
<td>.02</td>
<td>-10.0</td>
<td>-181-190</td>
<td>.21</td>
</tr>
<tr>
<td>Perceived Satisfaction</td>
<td>2.76</td>
<td>.61</td>
<td>2.36</td>
<td>.57</td>
<td>1.72</td>
<td>.10</td>
<td>.40</td>
<td>-.87-.08</td>
<td>.11</td>
</tr>
<tr>
<td>Perceived Effectivity</td>
<td>2.49</td>
<td>.45</td>
<td>2.13</td>
<td>.41</td>
<td>2.17</td>
<td>.04</td>
<td>.37</td>
<td>.02-.71</td>
<td>.16</td>
</tr>
<tr>
<td>AD-score</td>
<td>3.32</td>
<td>.60</td>
<td>2.84</td>
<td>.35</td>
<td>2.50</td>
<td>.02</td>
<td>.48</td>
<td>-.87-.08</td>
<td>.21</td>
</tr>
</tbody>
</table>

Note. M = Mean. SD = Standard Deviation. MD = Mean Difference. AD = Anchored Discussion (treatment). CD = Chat Discussion (control group). SMM = Shared Mental Model ranges from 26 (most similar) to 130 (least similar). Individual creativity by proxy number of contributions. Group Creativity ranging from 8 to 30 (most creative). Perceived Satisfaction and Perceived Effectivity range from 1 to 5 (best). AD-score ranges from 1 (unstructured) to 4 (well structured).

An independent samples t-test was conducted to compare the average individual creativity for groups that produced creative outcomes and groups that did not (H2a). There was no significant difference in individual creativity for groups that produced more creative outcomes (M=2425, SD = 502) and groups that had less creative outcomes (M = 2082, SD = 708; t(24) = -1.42, p =.17, two tailed). The magnitude of the differences in the means (mean difference = -341, 95% CI: -838-155) was moderate (eta squared = .07). H2a is not supported.

An independent samples t-test was conducted to compare the individual creativity for groups that had more structured and goal oriented interaction (i.e. higher AD-score) and groups that had less (H2b). There was no significant difference in individual creativity for groups that had a higher AD-score (M=2290, SD = 649) and groups that had a lower AD-score (M = 2212, SD = 623; t(24) = -.31, p =.76, two
The magnitude of the differences in the means (mean difference = -78.1, 95% CI: -596-439) was small (eta squared = .004). H2b is not supported.

An independent samples t-test [Tab. 2] was conducted to compare SMM similarity for groups of the different treatments (H3a). There was a significant difference in similar SMM between the treatments. The magnitude of the differences in the means was large. H3a is not supported, since lower SMM values indicate more similarity.

An independent samples t-test [Tab. 2] was conducted to compare the average individual creativity for groups of the different treatments (H3b). There was no significant difference in individual creativity between the treatments. The magnitude of the differences in the means was moderate. H3b is not supported.

An independent samples t-test [Tab. 2] was conducted to compare the levels of group creativity for groups of the different treatments (H3c). There was no significant difference in levels of group creativity between the treatments. The magnitude of the differences in the means was moderate. Even within the four dimensions of creativity (i.e. novelty, workability, relevance, and specificity) there is no significant difference between the treatments.

Additionally, we compared the creativity dimensions for each group. A group would be counted as 1 on a dimension if their rating was above the dimension mean. Groups that reached at least the mean on three dimensions were considered to be more creative. 8 AD and 5 CD groups met this criterion. The results in [Tab. 1] show that groups with anchored discussion produced more creative outcomes. This supports H3c.

We conducted further t-tests for the treatments, found in [Tab. 2]: Perceived satisfaction (Cronbach’s alpha: .82) is not significantly different between the groups with a moderate magnitude of the differences in the means. Perceived effectivity (Cronbach’s alpha: .75) is significantly better for CD groups with a large magnitude of the differences in the means. The AD-score is much higher for AD groups with a large magnitude of the differences in the means.

### 4.1 Qualitative Findings

The analysis of the presentation of the final document, the discussions, and the comments in the post-experiment survey revealed interesting findings that shed light on the unexpected quantitative results.

First, comments about structuring the idea with headings or bullet points were found early in all CD group chats. Sample comments read (translated): “I would say we structure this first” “How about basing the layout off of the three points from the assignment?” “Should we reorder the list?” In comparison, with the exception of one group, none of the AD groups discussed structure to begin with. Only after realizing that deleting lines would merge comments and that moving text did not move comments did some AD groups consider structuring to prevent discussions from becoming diluted later. The presentation of the final document reflects this difference. While all thirteen CD groups used headings and bullet points to structure their final document only eight AD groups did so. The remaining AD groups produced a single text block with little formatting.

Second, the discussion within CD groups focused sequentially on the different elements of the assignment. Before moving to a new topic, group consensus was often
conveyed by confirmation messages such as (translated) “yes”, “agreed”, or a supportive emoticon (i.e. :-) or ^^). In comparison, within AD groups the discussion on different elements occurred and far fewer confirmation messages were observed. Rather, having made changes to the idea text seemed to have been sufficient and required no further comments.

Third, our analysis of all 2084 comments supports that AD groups engaged in more clear and focused discussion. We coded each comment in one of three categories as either related to the idea, organizing the team, or non-related. Cohen’s Kappa (κ) was run to determine if there was agreement between the first and second author’s judgment. There was moderate agreement for the CD treatment, κ = .539, p < .0005, and good agreement for the AD treatment, κ = .878, p < .0005. For the CD treatment, both coders agreed that of 1218 comments 47.5% were idea-related, 20.0% were organizing the team, and 5.7% were off-topic. Of the 26.7% disagreement, the majority was about whether comments were idea-related or off-topic (11.0%), and idea-related or organizing (10.9%). This finding highlights that CD groups used many ambiguous comments. One common example was (translated): “I cannot think of anything more.” This comment could be a closing remark relating to an idea or an organizing statement that the discussion should move on. One example of an idea/off-topic disagreement is found in one group that went off tangent to discuss providing free beer to volunteers, which has the potential to spark ideas but is not clearly related to the idea. These ambiguous comments were most prevalent in CD groups. For the AD treatment, both coders agreed that of 866 comments 78.4% were idea-related, 7.2% were organizing, and 10.3% were off-topic. This finding supports that AD groups were engaged in a more focussed discussion on the ideas. The surprisingly high rate of off-topic comments resulted from discussions about the prototype.

Fourth, many AD group members commented that at first they did not understand how to use anchored discussion. This was evident in the discussion as well as the comments in the post-experiment survey. One AD user commented afterwards (translated): “A longer introduction to the web-app would have been good (maybe demonstrate the use with projector). It was not 100% clear, how the thing worked, that took a few minutes. That could have been demonstrated and everyone could have started on the task right away.” One AD group went as far as to discuss the bugs of the prototype towards the end of the session. In the post-experiment survey, emergent themes included the excitement for anchored discussion (9 participants), a wish that new comments would be indicated for each line (7 participants), a bug where the prototype stops synchronization and ideas entered are lost (5 participants), and unintuitive use of the tool (3 participants). In contrast, CD groups discussed the prototype far less during the session or afterwards and focussed more on the group dynamics. In the control group, emergent themes from the post-experiment survey addressed the sense of unplanned, unstructured and chaotic interaction (3 participants), a preference for face-to-face group meetings because it reduces chaotic interaction (2 participants), and the unstructured nature of the task description (2 participants).

Last, four CD groups created a custom anchored discussion within the text editor. Comments directly related to a specific text were added into the shared idea editor, often marked as a comment by putting “->” in between text and comment. This was unique to CD groups.
Based on our theoretical foundation regarding group creativity in an online environment, we designed and implemented a first version of a tool with anchored discussion. The results seem to not support most of our hypotheses. Below we will explore possible explanations.

In hypotheses 1a and 2a, we proposed that the shared mental model (SMM) of the group and the individual creativity (estimated by contributions as a proxy) of the group members would influence group creativity. These hypotheses are derived from the Team Creativity Model [Vreede, 2012], which proposes a relationship between SMM and individual creativity to group creativity. Both antecedents differed in the results between the treatments: the control group had the more favorable results in both cases. However, group creativity appears to be unaffected by this. This could indicate that some other factor was involved and that the relationship as stated in the Team Creativity Model is different from how we tested it. Furthermore, [Vreede, 2012] mentioned that a less similar SMM could be beneficial to group creativity, which might partly explain why the anchored discussion groups had a less favorable SMM-score but had slightly better creative outcomes.

In theory, a more structured and goal orientated group interaction was supposed to positively influence SMM and individual creativity, as stated in hypotheses 1b and 2b. The results did not support these hypotheses. The same observations from the previous paragraph apply.

Anchored discussion did enable the groups to have a more structured and goal orientated discussion with less needs to create references to the text, fewer interactions organizing the teamwork, and as a result require less discussion. The AD-score and qualitative findings clearly indicate that anchored discussion produced a more efficient interaction amongst the group members. However, hypotheses 3a (higher SMM-score for groups using anchored discussion) and 3b (higher individual creativity for groups using anchored discussion) are not supported by the results.

Interestingly, the groups using anchored discussion had a more structured interaction, but SMM, individual creativity, satisfaction, and perceived effectiveness scores were less favorable than in control groups’ scores. However, in support of hypothesis 3c, but contrary to our hypotheses 1a and 2a, the group creativity was slightly better in anchored discussion groups. A possible explanation could be limitations to our measurements. However, we used validated measures for group creativity, SMM, satisfaction, and perceived effectiveness. We acknowledge that using the contribution count as a proxy for individual creativity is arguably not preferable, and results in not being able to reliably support or refute hypotheses 2a, 2b, and 3b.

The qualitative analysis of the final document, the discussions, and comments from participants on the post-survey can aid the search for explanations. Participants were asked to comment on the experiment afterwards. A reoccurring theme was that participants were overwhelmed by anchored discussion and did not intuitively understand how to use it, which clearly showed in analysis of the discussion. Participants commented that they would have preferred a more detailed explanation of how to use anchored discussion and maybe even have an additional training before the actual creative task was assigned. It can thus be argued that the unknown interface
of anchored discussion caused a cognitive load that impeded on the collaboration task. However, giving an introduction on how to use the anchored discussion tool effectively could influence the reported levels of perceived effectiveness and thereby interfere with the measured results.

The negative effect of our prototype on satisfaction and perceived effectiveness could be further explained by some unexpected behavior. One anchored discussion group noticed early that deleting a line in the shared idea editor results in the attached comments to be combined with the next line and thus diluting that other discussion. The group’s solution was to write into a line “[deleted]” and thereby maintaining all the comments separate. Other groups interrupted their creative process to structure the document and thereby structuring subsequent discussion threats. We therefore conclude that a software design in which comments are moved to a different space if they become orphaned due to a deleted anchor would be better. Also participants asked for comments to be tied to the text and thus be moved with the text, if the text is rearranged. One group left the lines with comments in the original order to maintain all comments and compiled the fully developed plan below in new lines. The inhibiting effect of the prototype is most evident in the format of AD groups’ final documents, which often lack structuring headings and bullet points. In summary, this unintuitive behavior also appeared to increase cognitive load and reduce perceived effectiveness and satisfaction. While the effect was strongest on the antecedents, the more structured discussion appeared to compensate for this resulting in the group creativity being slightly above the control group. This is remarkable in the light of the off-topic discussion concerning the prototype within AD groups, which further distracted from the task.

Continuing with participants’ suggestions for improvements, the most requested feature was a visual indication that an unread comment existed. In our prototype the number of comments were displayed next to the line number, and when this number increased no visual cue was provided. One participant explained the downside of this implementation (translated): “Some comments were never noticed.” We had not implemented this feature to avoid distraction and to not disrupt the creative thought process. However, it seems that a non-obtrusive change of color or font style for the comment count could aid the collaboration of the group members by reducing the time in search for new comments.

Only participants in the control groups argued that they would have liked additional support to structure their collaboration process. The analysis of the discussion demonstrated that CD groups put much effort into structuring the document and coordinating the discussion. Headings and bullet points were mostly used by CD groups, but apparently did not satisfy the need for coordination. Suggestions for more support included to-do-lists and voting features. The communication was described in some comments as unstructured, having too many conflicts, and not being goal oriented. One comment described in detail how the constant stream of comments in the chat area caused attention diversion and prevented creative work in the shared idea editor.

Interestingly, four of the 13 control groups helped themselves and created their own anchored discussion within the text. Thus we can see how anchored discussion is a promising solution for structuring text related discussions. The effectiveness is supported by the fact, that three of the four groups were rated as having created a
6 Limitations

Our research has several limitations. First, the standard limitations of a lab experiment apply. For example, students were encouraged to participate but may not have had any intrinsic motivation to do so. When anchored discussion is used on social media platforms where people collaborate only voluntarily, different motivations and incentives may apply so a field evaluation of anchored discussions is recommended. Also, the fictitious experiment task may not have stimulated creativity and participation equally for all students.

Additional limitations are related to the prototype. Many participants complained about the performance and attributed slower collaboration to this aspect, but this was consistent for both treatments. Additionally, as argued above, the functionality of anchored discussion was unknown to the subjects. Prior training might have aided performance and yielded different results. Better results might have also been achieved if the participants’ suggestions were implemented and the stability was improved.

The design of our prototype did not allow us to perform personalized idea counts, which would be a preferred measurement for creativity [Reinig, 2007]. Using number of characters typed as a proxy for individual creativity has several drawbacks. Contribution counts could be high for other reasons, like correcting spelling mistakes, or communicating unrelated messages. However, using the count of good-ideas would be a more correct measurement, especially since the ratio of good-ideas to total ideas can serve as an indicator for cognitive inertia experienced by the group [Briggs, 2007], which anchored discussion should reduce. Thus, comparing contribution counts has limited value in comparing individual creativity between control groups and anchored discussion groups.

Finally, the dimensions from [Dean, 2006] appear to be a better fit for evaluating many ideas that are created after another, and not so much for evaluating the overall creative outcome. This can be demonstrated on the dimensions originality and paradigm relatedness, which both are better judged in comparison to previously created ideas, but are difficult to judge without reference points. Another limitation in our evaluation of the creative group outcome is that we did only a short training with the judges, because [Amabile, 1982] suggested it was not needed, and not as extensive as was suggested [Dean, 2006].

7 Conclusions

We successfully demonstrated that the body of knowledge on creativity provides a sufficient foundation to inform the development of new tools to support creativity. Based on the body of knowledge we designed and developed a prototype that used anchored discussion and evaluated it in an initial experiment.
Our experiment revealed that anchoring is an adequate tool to structure an ideation process and has the potential to enhance performance. The participants used anchored discussion or created their own anchored discussion-like functionality. This suggests that anchoring is effective to structure the communication in a creative process. Further, the results suggest that anchored discussion leads to better overall ideas and that groups using anchored discussion are working in a more structured way. However, our study revealed the most important issue of any creativity tool, which is usability. As the design of the system shapes the group interaction, the implementation of anchored discussion also influences group performance, even impairing creativity, if inappropriately designed. Our work provides tested design guidelines and suggestions for improvements. Most importantly, structuring group work using anchored discussion should be intuitive and not overstrain the user, leading to high cognitive load.

To summarize our findings, anchored discussion seems promising. From the experiment we expected stronger support for our hypotheses, yet it has to be considered that this was the first prototype to explore the use of anchored discussion for creativity in online groups. We have gained a better understanding of what can be improved.

Our contribution is twofold. First, we demonstrated how theory can serve as a foundation to inform the design and development of new collaboration tools. Second, we advise future research on how to build better prototypes using anchored discussion to increase the creative quality of online group ideation. In addition, we initiated the idea of implementing anchored discussion into a variety of group support systems, where individuals work together in a shared editor. Anchored discussion has the potential of improving group work, such as joint decision-making, joint evaluation, joint problem solving, or other group processes. Even so, the implementation of anchored discussion probably needs appropriation for different contexts. Our findings reveal the most important issues and provide valuable insights for designers and developers of creativity support systems.

8 Future research

There are several promising avenues for future research. To date, anchored discussion has not been evaluated for creative tasks. Our first prototype was positively received by participants, who made some suggestions for an improved design. Further research could incorporate and improve the usage of anchored discussion in creativity support tools and other group support technology. Comprehensive tasks and problems can be separated into different aspects to reduce cognitive load, while still maintaining the big picture. Our research results can therefore be seen as informing future design guidelines on how to implement anchored discussion into group support systems.

Additional insights could be generated, if the divergent and convergent creativity process were measured independently. The influence of anchored discussion on the different creativity processes could aid in understanding the underlying principles of anchored discussion. This could reveal whether the structuring of content is desirable for either the divergent stage or the convergent stage of the creativity process. This could also increase a more comprehensive understanding of anchored discussion.
In social media, crowdsourcing, and open innovation processes, people are more likely to contribute when a task is enjoyable, easy to understand, rewarding, and when it creates a state of flow. Further research could evaluate the effect of anchored discussion on these dimensions.

Acknowledgements

Funding for the first author was provided by a scholarship from the Hanns-Seidel-Foundation as part of its mission to support highly talented students with financial support from the German Federal Ministry of Education and Research (BMBF).

The National Science Foundation Grant #1322285 provided funding for the third author’s efforts to this research.

References


