Comparing the Generativity of Problem Solving and Appreciative Inquiry: A Field Experiment

Gervase R. Bushe1 and Neelima Paranjpey2

Abstract
Appreciative inquiry (AI) theorists claim AI is a more generative form of inquiry than problem solving; this study uses a classical field experiment to test that claim. We test three different processes for producing generative ideas defined as new ideas that motivate new actions. Why AI may be better at producing such ideas is explored and a method for amplifying those qualities (synergenesis) is described. Hypotheses are tested by assessing ideas produced from groups of employees at an urban transit organization. Synergenesis-based groups scored significantly higher than either of the other groups on ratings of generative ideas. Examination of participant’s pre- and post semantic maps show predictable differences in the effects of problem solving and appreciative approaches on engagement of employees in the ideation phase of a change process, consistent with AI claims. Implications for practitioners and suggestions for future research are discussed.

Keywords
appreciative inquiry, generativity, brainstorming, field experiment

Appreciative inquiry (AI) was initially conceptualized and offered as a method for producing what Gergen (1978) called “generative theory” (Cooperrider, 2013). The seminal article on AI (Cooperrider & Srivastva, 1987) contended that there was a need
to view action science in light of the theory by which it is supported. They critiqued
the positivist assumptions and stance underlying much organization development and
organization research and proposed, instead, that action research be focused on its
“generative capacity.” This recommendation rested on Gergen’s (1978) proposal to
shift social research away from a scientific search for prediction and control and,
instead, focus it toward “the capacity to challenge the guiding assumptions, to raise
fundamental questions, to foster reconsideration of that which is taken for granted, and
thereby to generate fresh alternatives for social action” (p. 1344). In this study, we
extrapolate from Gergen to define generative ideas as new ideas which are compelling
and thereby increase the options for change and the probability that change will occur.

Generative process, therefore, is one that results in generative ideas.

While the impact of AI in the past 25 years has been immense, very little empirical
research exists explicating how one increases the generativity of any organizational
change effort, nor has there been much research to study any of the claims made about
it. This study is an initial effort to address both those questions. In this study, we use a
naturally occurring field experiment to test the claim that AI is a more generative pro-
cess than problem solving, and we explore the processes that make AI generative. In
this article, we begin by describing how the idea of generativity has emerged in the
organization development literature. We review the rather sparse literature on stimu-
lating generativity in organizations and describe a method, synerogenesis, first pro-
posed by Bushe (1995) for increasing the generativity of the Discovery phase of AI.
We then describe an experiment conducted in a metropolitan transit authority that
engaged employees in designing an employee recognition system and the methodol-
gy we used to test the generativity of three different approaches to engaging employ-
ees in the ideation stage of organizational change.

**Generativity in Organization Development**

After Cooperrider & Srivastva’s (1987) initial article on AI, the notion of generativity
receded behind the attention given to the focus on “the positive” in AI (e.g., Barge &
Oliver, 2003; Cooperrider & Sekerka, 2006; Fineman, 2006; Fitzgerald, Oliver, &
Hoxsey, 2010). Recently however, the importance of generativity to the transforma-
tional change potential of AI has regained attention (Bright, Powley, Fry, & Barrett,
2013; Bushe, 2010; Bushe & Kassam, 2005; van den Nieuwenhof, 2013; Zandee,
2013), exemplified in the most recent World Conference on Appreciative Inquiry in
2012 being titled “Scaling up the Generative Power of AI.” Generativity is ascending
as an important meme in organization studies. A simple scan of journal articles where
the words “generative” or “generativity” appeared using Business Source Complete
found only 38 articles in 2003 but 519 articles in 2013. Recently, Bushe and Marshak
(2014, 2015) have argued that generativity is one of three transformational processes
underlying any successful Dialogic OD effort.

Yet in organization development the words generative and generativity often get
used without much definition. Generative is an adjective and, according to the
Merriam-Webster Dictionary means having the power or function of generating,
originating, producing, or reproducing. Sometimes it is used to describe a process of interaction that produces something more than the process itself. For example, Isaacs (1999) writes about generative dialogue, and Marshak (2004) of generative conversations. In these instances, generative is used to indicate that the interaction produces a salutary outcome of some kind. Other times, the term is used to describe the property of something that, in turn, produces an effect on social interactions. For example, in Schön’s (1979) work on “generative metaphor,” he proposed that all decision-making processes are powerfully influenced by the underlying metaphors decision makers relied on to think about complex issues. He argued that all social policy discourses rested on metaphors that focused perceptions and influenced narratives in specific ways. Thus, they were generative, though not necessarily salutary, and he advised policy makers to become aware of the metaphors at work in the narratives guiding their decisions. Gergen’s (1978, 1982) “generative theory” was of a similar nature, used to depict a kind of theory that, in turn, affects the social construction of reality.

Application of the idea of generativity to organization development was first presented in a series of publications by Frank Barrett, David Cooperrider, and Suresh Srivastva. Cooperrider and Srivastva’s (1987) first article on AI built from Gergen (1978) to support their argument that the main barrier limiting organization development had been its romance with action at the expense of theory. This separation of theory and action, they argued, was supported by an underlying generative metaphor of “organizations as problems to be solved” and the consequent view of OD as primarily a process of problem solving. To them, too many in the discipline had underestimated the power of new ideas for changing social systems. Theories “may be among the most powerful resources human beings have for contributing to change and development in the groups and organizations in which they live . . .” (Cooperrider & Srivastva, 1987, p. 132). To the extent that action is based on ideas, beliefs, meanings, and intentions, organizations can be transformed by changing idea systems or preferred ways of talking. How do we inquire in a way that is more likely to create new, generative images, ideas and theories? AI was initially conceptualized and offered as a method for producing generative theories. Almost no empirical research exists examining this claim.

A key premise they offered was that theories are generative when they expand the realm of the possible and point toward an appealing future. In so arguing, they moved beyond Schön’s and Gergen’s use of the notion, and proposed that generativity has to do with our shared and desired futures and our ways of making these futures possible. In its simplest terms, these are new ideas that are compelling to people—they offer a new way of thinking and acting, and people want to act on them (Bushe, 2013b). After reviewing a number of theorists and research studies they made the point that a method of inquiry that would generate such ideas would have to proceed from an affirmative stance. This “focus on the positive” captivated practitioners and academics alike, and AI has mainly come to be associated with it in both supportive and critical descriptions. The idea of generative capacity is hardly mentioned at all. In this dominant narrative, the transformational effects of AI are attributed to its focus on the positive attributes of social systems (Cooperrider, 1990; George & McLean, 2002; Ludema,
Whitney, Mohr, & Griffin, 2003) its ability to engender “positive emotions” (Cooperrider & Sekerka, 2006; Fry & Barrett, 2002) and its refocusing of managerial attention away from problems to strengths and capabilities (Watkins, Mohr, & Kelly, 2011; Whitney & Trosten-Bloom, 2010).

One exception to this has been empirical research by Bushe and his colleagues. An early article described how AI in teams could lead to the emergence of a generative image that helped the group get unstuck from whatever was causing it problems (Bushe, 1998). A meta-analysis of 20 published cases of AI found that in all seven cases that showed transformational changes, new ideas and a generative metaphor had emerged while in the 13 incremental change cases none seemed to produce new ideas and only one described the emergence of a generative metaphor (Bushe & Kassam, 2005). A later study of eight appreciative inquiries in a metropolitan school district (Bushe, 2010) comparing four sites that experienced transformational change with four that did not found that new ideas which captured people’s energy and enthusiasm and led to changes emerged in the transformational sites while none did in the other four. That study also found that all the transformational sites used AI to address problems that were widely seen as important problems, while that was not true in the non-transformational sites. This led Bushe (2013b) to argue that people would not put in the effort required for transformational change unless they addressed real, pressing problems, but that AI addressed problems through generative ideas, not problem solving. Survey measures in that (2010) study showed participants at all sites had high levels of positive feelings and positive anticipations about the future after their AI summits, suggesting that positivity may not be the core of AI’s transformational potential (Bushe, 2013b) and that more attention needs to be placed on understanding the nature and sources of generativity.

While all this supports Bush and Marshak’s (2014, 2015) argument that the capacity of any Dialogic OD process to produce new ideas that people find compelling and want to act on, is of central importance to the success of that change effort, does very little to explain where new ideas come from nor does it test whether AI is a better way of enhancing creativity and producing ideas than more conventional problem-solving approaches to change. A key conceptual scaffolding for AI has been research showing the beneficial effects of positive images and positive affect on creativity and performance. Cooperrider (1990) proposed the “heliotropic hypothesis” to explain the power of positive images to influence performance in individuals and groups. Recently, Bright and Cameron (2009) have argued that research on positive organizational climates, positive energy networks and high-quality relationships substantiate the proposition that heliotropism exists in social organizations. They also point out that since “bad is stronger than good” (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001) an emphasis on the positive must be sufficiently pervasive and strong enough to overcome the natural tendency of people and organizations to be more affected by negative events, situations and interactions than positive ones, which is consistent with studies showing effects from ratios of more positive than negative statements on creativity in groups (Fredrickson & Losada, 2005).
More attention in the AI literature, however, has been placed on the role of positive emotions, than positive images or ideas, for increasing creativity in groups (Cooperrider & Sekerka, 2006; Cooperrider & Whitney, 2001). Studies that show positive feelings lead people to be more flexible, creative, integrative, open to information and efficient in their thinking (e.g., Amabile, Barsade, Mueller, & Staw, 2005; Isen, Daubman, & Nowicki, 1987; Isen, Johnson, Mertz, & Robinson, 1985) are often cited, along with Fredrickson’s (2001, 2006) research showing that people experiencing positive affect are more resilient and able to cope with occasional adversity, have an increased preference for variety, and accept a broader array of behavioral options. Closely aligned is Ludema’s articulation of the nature and importance of hope for encouraging creative engagement in organizations (Ludema, Wilmot, & Srivastva, 1997) and the way in which AI can provide hope (Ludema, 2000).

As George (2007) points out, however, recent research is providing a much more nuanced understanding of how internal states interact with context to produce new ideas. Under positive affect people are less likely to see a need for novelty and creativity while negative affective states lead people to be more likely to identify situations requiring creative solutions and to focus more carefully on the facts on hand than rely on preexisting ways of thinking about things (Kaufmann, 2003; Martin & Stoner, 1996; Schwarz, 2002). Negative mood states have been shown to lead to more creativity when people expect to get recognition and rewards for creativity and clarity of feelings are high (George & Zhou, 2001). It may well be that, positive and negative emotions, and probably any internal state, can contribute to producing new ideas depending on context and how people make meaning of the situation they are in.

**Problem Solving Versus Appreciative Inquiry as Generative Processes**

Since Osborn (1953) laid down the rules of brainstorming, the technique has become an ubiquitous feature of the ideation or divergent phase of group problem solving. Hundreds of studies have examined the claims made by Osborn for the superiority of free-wheeling groups, where criticism is suppressed, and quantity of ideas is sought, for producing new ideas. A lot of this research is not that relevant to issues of generativity because until a few years ago brainstorming research focused almost exclusively on idea quantity. Recent studies, however have begun to study the impact of brainstorming on idea quality (Stroebe, Nijstad, & Rietzschel, 2010). As Girotra, Terwiesch, and Ulrich (2010) point out, in organizations the success of idea generation usually depends on the quality of the best idea generated and that in most situations, 99 bad ideas and 1 outstanding idea is preferable to 100 merely good ideas. Certainly this is true in the case of generative organizational change processes since it is the capacity of an idea to generate new meaning and action that makes it generative.

Most brainstorming research, however, has operated on the assumption that the more ideas generated, the more likely a creative idea will emerge. A key, consistent finding, is that when you aggregate the output of individuals working alone (nominal groups), they produce more ideas than a group working together (cf. Diehl & Stroebe,
Yet, despite dozens of critical studies, group brainstorming remains a popular idea generation technique (Coskun, 2005; Girotra et al., 2010; Rietzschel, Nijstad, & Stroebe, 2007) and some studies suggest that contextual and motivational influences may result in teams being more creative than individuals in organizational settings (Singh & Fleming, 2010; Sutton & Hargadon, 1996; Taylor & Greve, 2006). Many of the explanations for the poorer results of groups over individuals in experimental studies involve impairments to cognitive processes like production blocking (Diehl & Stroebe, 1987, 1991—the problems that arise from turn taking and having to manage social interaction at the same time as one is thinking), problems with group interaction getting in the way of information retrieval (Hinsz, Tindale, & Vollrath, 1997), and problems from fixating on what previous people have said (Kohn & Smith, 2010; Smith, 2003). While research on the impact of emotion on group brainstorming is virtually nonexistent, the most robust explanation for poorer results in groups than individuals working alone is evaluation apprehension (Camacho & Paulus, 1995; Diehl & Stroebe, 1987), the concern that one’s ideas will be negatively judged by others. One of the few studies of brainstorming in organizations (Rickards, 1975), found that organization members refrained from freely speculating in brainstorming sessions, presumably due to evaluation apprehension.

There is no reason to suspect that brainstorming, of itself, creates evaluation apprehension. Indeed, the ground rule of no criticism is intended to ensure that does not happen. But it seems our brains are so wired for social comparison and status threats (Boksem, Kostermans, & De Cremer, 2011; Zink et al., 2008) that any kind of unstructured group interaction can create it. We argue that evaluation apprehension is likely to produce negative emotional states and that AI may be a more generative process that produces more generative ideas (ideas that are new and compelling) by both reducing evaluation apprehension and increasing positive affect, which, as noted earlier, has been associated with increased cognitive flexibility and creativity. Amelioration of evaluation apprehension is one of three reasons to suspect that the initial “Discovery phase” of AI creates contexts that make it more likely to result in more generative ideas than problem solving.

Recall that in the Discovery phase of AI, participants listen to each other’s stories of peak experiences related to the focus of the inquiry and then discuss what they have learned from listening to the stories. Numerous writings provide anecdotal evidence that this process increases positive affect and personal bonds, increasing trust and a sense of community (e.g., Khalsa, 2002; Ludema, Cooperrider, & Barrett, 2001; Whitney & Trosten-Bloom, 2010). Bushe’s (2010) study of eight appreciative inquiries found that positive affect was so high in all cases the lack of variance failed to explain why some sites had transformational changes while others did not. Thus, AI appears to create social conditions more conducive to group creativity.

The possibility-centric focus of AI is a second reason to suspect it is more likely to create more relaxed, cognitive conditions conducive to creativity. Cooperrider and Barrett (2002) describe a training program in which some managers were instructed to analyze the organization and report on the issues it needed to address, while others were instructed to do an AI and report on the successes it was achieving. Apparently
the impact on the managers in each group was so stark, with those in the former group appearing depressed, deflated and demoralized, while the latter group appeared excited, energized and inspired, it convinced senior managers to try an AI throughout their organization. We propose that focusing on desired futures, rather than analyzing current problems, sets the stage for more generative ideas to emerge.

During the Discovery phase of AI, participants trade concrete stories that they then explore to uncover ideas related to the focal concern. For example, if an inquiry is about increasing customer satisfaction, stories of exemplary customer service might be examined to identify the sources of customer satisfaction. We contend that situating inquiry first in concrete experience makes it more likely that people will have new ideas. This is a third reason to suspect that AI will lead to more compelling new ideas, than problem solving. As long as discussions are based on the current mental models of individuals, it’s unlikely that they will have a new idea. Focusing on lived experience, and being questioned about it by others, increases the possibility for things not currently in one’s mental model to be revealed.

**Hypothesis 1:** The Discovery phase of AI will be more generative, producing more ideas that are new and compelling, than brainstorming in a problem-solving process.

**Increasing Generativity With Synergenesis**

If it is true that transformational change hinges on the generativity of a change process, and one advantage of AI over problem solving is that it is more generative in producing new ideas that people want to act on, then it is useful to ask what might support or increase the generativity of AI itself. Synergenesis, a technique developed by Bushe (1995, 2007), claims to increase the generativity of the Discovery phase of AI. It requires participants to not only describe and listen to each other’s peak experience stories but also to write up another person’s story, written in the first person. These written stories are then used to prime a group brainstorming session about the focal topic as follows. A small group meets and is presented with a specific question to answer based on the focal concern driving the change process. All members read one story together and discuss ideas the story evoked in them for answering the question. They are encouraged to not simply analyze the story but to vocalize any idea that comes to mind during the conversation. All ideas are boarded and once the energy runs out, the group reads another story and continues on until further stories produce no new ideas.

There are a number of reasons to suspect that such a procedure will produce even more novel and compelling ideas than the traditional AI. First, the AI Discovery process can be run more as an attempt to identify preexisting strengths and the “positive change core” of the system, than to surface new and compelling ideas. Ideation may be held off for the later phases of Dream or Design. Bushe’s (2010) study found, however, that the genesis of many of the generative ideas in the transformational sites could be traced back to the Discovery phase. It may be that procedures that encourage participants to be alert to new and compelling ideas during Discovery will produce more generative appreciative inquiries.
Second, it is likely that the synergenesis procedure ensures participants hear/read more of the stories generated during the Discovery phase. In classical AI, participants are often paired to do interviews and then meet in small groups to share the stories that emerged. The only stories they are sure to hear are those from the members of those small groups, and there is no control over the extent to which any of those stories are actually attended to. In synergenesis, however, participants consider a large number of stories, each in turn, continuing to read and discuss stories until doing so no longer produces any new ideas. In addition, these stories have been preselected from all those produced during the Discovery phase because of their novel or provocative potential. As a result, more new and compelling ideas are likely to be produced.

Third, requiring people to write up another person’s story forces them to pay closer attention to the story and to think about the important lessons in the story for addressing the focal concern. This step ensures that people are taken out of their own stories and are considering the focal issue from new angles, which may increase subsequent creativity.

There is also evidence from the brainstorming research to support the assertion that a procedure like synergenesis will be more generative. While interacting in groups can reduce idea quantity due to production blocking, there is evidence that sharing ideas can be stimulating to the creative quality of ideas produced (Dugosh, Paulus, Roland, & Yang, 2000; Nijstad, Stroebe, & Lodewijks, 2002; Stroebe et al., 2010). Furthermore, the impact of shared ideas is enhanced when individuals are motivated to attend to those ideas (Kohn, Paulus, & Choi, 2011; Paulus & Yang, 2000), which we argue will occur as a result of having to write up another person’s story, and reading those stories together, in a group.

**Hypothesis 2:** Synergenesis will be more generative, producing more new and compelling ideas, than conventional AI or problem solving.

**Methodology**

Seventy-six employees of a Midwest urban transit organization of 10,000 employees who volunteered to participate were assigned to one of six meetings where they were tasked with producing ideas for an employee recognition program. In each meeting one of the three ideation processes being studied was used. This activity was part of an effort from the human resources department to improve employee recognition among a largely unionized, long-tenured workforce. An Employee Recognition Advisory committee guided the inquiry. They spread the word about the initiative and encouraged employees to participate by joining in a group discussion. An email was sent to the employees who were willing to participate. Managers were notified about their employees’ involvement and were provided information about the intent of the process and the duration of participation.

Participants were categorized into operations, nonoperations, exempt employees, and nonexempt employees, as these were considered important demographic
differences in this organization, and members from each category were assigned in roughly equal numbers to each of the six meetings, each tasked with coming up with ideas to improve employee recognition. In each case the session was led by the second author. All meetings took place within a 2-month period. There were two meetings for each experimental condition, that is, two sessions used problem solving, two did a conventional AI Discovery process, and the other two did synergenesis. To ensure experimental effects could not be simply attributed to cognitive stimulation from conducting interviews prior to group ideation, each condition began with people interviewing each other in pairs (sometimes, because of numbers, a trio was formed) even though that is probably unusual in problem-solving situations. In each condition, participants were given interview scripts they were asked to follow. Each session lasted between 2 and 3 hours. We describe each in turn, and the procedures for each are summarized in Table 1.

Problem Solving

Participants were asked to form pairs and interview each other on their thoughts about employee recognition programs. The first question in this script was, “Why do employee recognition programs fail?” with the follow-up probes “What are some problems that we need to overcome at (organization) to recognize employees?” and “What solutions can you think of for having a long-term recognition program?” Further questions asked how they would like to be recognized, how to improve communication in the organization and how they thought absenteeism could be reduced. After the interviews participants formed into groups of 5 to 7 to brainstorm answers to the questions, “What are the next steps in order to start a recognition program at the organization? How can the recognition program be useful for all employees, from management, nonmanagement, union, nonunion, all locations?” The ideas generated were collected for later analysis.

Classical AI

There are many ways one can shape the Discovery phase of an AI. We suggest that managers and facilitators often look for ways to condense the 4 D process into only 1 or 2 days. A Discovery process that lasts from 2 to 3 hours is probably very common, and we adopt it here with the label “Three Hour Discovery” to acknowledge that Discovery can be designed to last for a much longer time and include more generative activities than our Design.

Again, participants were asked to pair up and interview each other, this time using an AI style interview guide. The first question was, “Will you please think back on your career/work history and tell me about a time when you received recognition, appreciation or acknowledgement for your work; the time where because of that recognition, you felt a great sense of satisfaction?” Further questions asked for their best experiences of communication in the organization, what most inspired
them to come to work every day, and to imagine a future in which they felt recognized and appreciated by the organization and what that would look like. Participants were encouraged to be genuinely curious about their partner’s experiences and not provide opinions about their stories. If necessary, they could probe for more details regarding the story.

Table 1. Summary of the Three Experimental Conditions.

<table>
<thead>
<tr>
<th>Initial interview question asked in pairs</th>
<th>Problem solving</th>
<th>Classical AI</th>
<th>Synergensis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why do ER programs fail?</td>
<td>Will you please think back on your career/ work history and tell me about a time when you received recognition, appreciation, or acknowledgement for your work; the time where because of that recognition, you felt a great sense of satisfaction?</td>
<td>Will you please think back on your career/ work history and tell me about a time when you received recognition, appreciation, or acknowledgement for your work; the time where because of that recognition, you felt a great sense of satisfaction?</td>
<td>Individually asked to write up the story they found most inspiring. Stories with the most stimulating and inspiring ideas selected by facilitators.</td>
</tr>
<tr>
<td>Ideation process</td>
<td>Pairs asked to form groups of 5-7 to brainstorm answers to the question, “What are the next steps in order to start a recognition program at the organization? How can the recognition program be useful for all employees, from management, nonmanagement, union, nonunion, all locations?”</td>
<td>Pairs asked to form groups of 5-7 to talk about what the stories tell you about an organization of employee recognition at its best. Find the patterns and themes and look for the root causes of success.</td>
<td>Pairs asked to form groups of 5-7 and to read one story together and then brainstorm ideas about “What types of recognition make people feel valued and appreciated, inspiring them to come to work every day and do their best?” Once there were no more ideas, read another story and repeat. Continue until reading another story produces no more new ideas.</td>
</tr>
<tr>
<td></td>
<td>Individuals asked to write up the story they found most inspiring. Stories with the most stimulating and inspiring ideas selected by facilitators.</td>
<td>Then brainstorm all of the root causes of success of employee recognition programs.</td>
<td></td>
</tr>
</tbody>
</table>

Once the interviews were over each pair met with two other pairs to form a group of 5 to 7 and given the following instructions:

As a group, talk about what the stories tell you about an organization of Employee Recognition at its best. Find the patterns and themes lying within those stories, and in particular what were the circumstances (situations, people, processes, etc.) that enabled those experiences. Look for the root causes of success that made them possible. It may take some time before the themes and patterns become clear to you as a group so do not feel pressured to come up with instant answers. Now as a group, brainstorm all of the “root causes of success” you found in the stories and write them on a flip chart.

Once again ideas were collected for later analysis.

Synergenesis

Once again participants were asked to pair and up for interviews, and they were given the same interview guide as the AI condition. After the interviews were over, participants were asked to identify the story that most inspired them and to write it up. They were asked to write it as a story (not a chronological recounting of what was said in the interview, but as a story written in the first person with a beginning, middle, and end). Once the story was written, it was given back to the interviewee to amend as needed and they were given a break. During the break the author and two members of the Employee Recognition Advisory committee selected the best stories for the synergenesis sessions. Best stories were those that included stimulating and inspiring insights and experiences.

Once again, pairs were asked to meet with other pairs, forming groups of 5 to 7. They were asked to follow the synergenesis process of reading a story, brainstorming ideas, and then reading another story until no more new ideas were being produced. The question to focus the synergenesis was, “What types of recognition make people feel valued and appreciated, inspiring them to come to work every day and do their best?” Ideas were collected for later analysis.

Generativity Measures

The generativity of each process was assessed in two ways. One was to assess how generative each idea was. Using a 5-point Likert-type scale, expert raters (described below) rated every idea produced by every group on three questions designed to assess how new and compelling each idea was. One question simply asked (1) Novel: “Is this idea novel/ new which has not been done before in this organization?” The other two questions were designed to assess how compelling the idea was by asking about the extent to which it evoked their interest and they thought it could be implemented: (2) Interesting: Does this idea evoke interest and compel you to implement it? (3) Practical: Can this idea be implemented practically in this organization?
All ideas generated were assembled into a random order and rated by an expert panel who were blind to the experiment. Raters were chosen from varied departments: law, finance, human resources, operations, and technology to obtain a good representation from all areas of the organization. All had more than 5 years of experience with the organization (some considerably more), they had the requisite organizational knowledge required for rating the comments, and they had some background/experience with past employee recognition initiatives. The five raters independently rated all the ideas. Ratings were combined into one score for each idea on each of the three criteria. Analysis of variance was used to test hypothesized differences between the three conditions in the generativity of ideas produced.

A second measure of the generativity of the process examined the effect each condition had on the mental models of participants before and after their participation in the group discussions. Though this was not directly related to testing our hypotheses, we hoped it would provide greater insight into the effects of each condition. A more generative process should lead to more change in participant mental models, evoking new ideas for action and more interest in taking action. To study this, all participants were asked to anonymously write down their answers to the same two open-ended questions immediately before and immediately after their participation in the ideation sessions: “What are your thoughts on having an Employee Recognition Program in the organization? What should it include?” The responses were analyzed to understand whether there were changes in mental maps, whether there were patterns in these changes, and whether individuals from any particular group were more favorable toward having an Employee Recognition Program after the group discussions. To understand the sentiment and the meaning of words, open coding was conducted. In the open coding phase, the text was examined for salient categories of information using Wolcott’s (1994) coding approach, with every sentence coded. The Wordnet dictionary was used to define codes for hope, positive comments, and negative comments, with similar ideas grouped together as one comment. A comment that alluded to specific suggestions or an action was categorized as action step. For example, “The employee recognition program should include announcements of milestones and achievements on a monthly basis.” Other thoughts and ideas that did not fall into any of these four categories were classified as “other.” In a very few cases, the same ideas were coded under more than one category. For example, the comment “I am excited and hopeful the company will follow through” was coded as both a positive sentiment and a hopeful one. All comments were analyzed by using Automap and ORA to produce semantic maps of participants in the three conditions, pre– and post–group discussion.

Results

Sample

Data on demographic variables that might affect participants’ ideas about employee recognition were collected from the sample and examined to insure the three conditions were composed of equivalent types of employees. Table 2 shows the number of
participants in each condition, their years of service, management or nonmanagement, exempt or nonexempt, and whether located in the field or at the general office. An ANOVA across each variable in the sample found no significant differences in the three conditions.

### Generative Ideas

The interrater reliability among the raters, overall, was computed using the intraclass correlation coefficient, which is a measure of agreement among two or more raters. The intraclass correlation of 0.80 obtained in this sample is considered to be a high to almost perfect agreement (Altman, 1991). The average ratings of ideas on each of the three criteria were used for further analysis. The ratings of all ideas generated in each condition were combined and averaged, producing one score for each measure of generativity for each of the three approaches to generating ideas. Table 3 shows the means and standard deviations of average ratings on innovative, interesting, and practical for each condition.

The number of ideas generated in each condition were almost similar, \( n = 51, n = 56, n = 50 \) for Synergensis, Classical AI, and Problem Solving, respectively. One-way ANOVA was used to test for significant differences in the scores across experimental conditions. AI was expected to score higher than problem solving, and synergenesis higher than AI. All means in all cells were in the expected direction. No significant differences, however, were found for ratings of novelty, \( F(2, 154) = 1.72, p = .182 \).
The results were significant for the two measures of compelling: interesting: $F(2, 154) = 7.592, p = .01$; and practical: $F(2, 154) = 10.074, p = .00$. As shown in Table 3, ideas from groups using synergenesis were rated as significantly more interesting and practical than both the other conditions. No other statistically significant differences were found.

**Impact on Mental Maps**

The semantic maps created by coding the written responses to the open ended questions “What are your thoughts on having an Employee Recognition Program in the organization? What should it include?” are shown in Figures 1 to 6. They show that the number of positive, hopeful, negative, and action step ideas were almost similar in all three conditions in the pretest.

![Figure 1. Pre-synergenesis semantic map.](https://example.com/figure1.png)

<table>
<thead>
<tr>
<th></th>
<th>SYN</th>
<th>AI</th>
<th>PS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel</td>
<td>3.27 ± 0.36</td>
<td>3.30 ± 0.41</td>
<td>3.17 ± 0.35</td>
</tr>
<tr>
<td>Interest</td>
<td>4.02a,b,***</td>
<td>3.71a,0.56</td>
<td>3.57b,0.71</td>
</tr>
<tr>
<td>Practical</td>
<td>3.95a,b,***</td>
<td>3.65a,0.61</td>
<td>3.42b,0.68</td>
</tr>
<tr>
<td>N of ideas</td>
<td>51</td>
<td>56</td>
<td>50</td>
</tr>
</tbody>
</table>

Note. SYN = synergenesis; AI = appreciative inquiry; PS = problem solving. Means that share the same subscript are significantly different from each other.

*p < .05. **p < .01.
The posttest show a narrowing, or convergence, with less ideas in every category in almost all cases. Most striking, however, is the steep drop in ideas after the problem-solving discussions. Table 4 shows the number of ideas in each of the categories by condition, pre and post. While those in the synergenesis and AI conditions provided 25% and 37% less ideas post-group discussion, those in the problem-solving condition provided 65% less. Members of problem solving groups only provided two ideas that fell into the action steps category and less than half the positive comments of people in the other two conditions. Accounting for the differences in number of people providing written responses, the average number of ideas in the minds of participants in the problem-solving condition dropped from 2.42 to 1.14 per participant. The drops

![Figure 2. Post-synergenesis semantic map.](image)

### Table 4. Number of Ideas in Each Category, in Each Condition, Pre– and Post–Group Discussion.

<table>
<thead>
<tr>
<th></th>
<th>SYN Pre</th>
<th>SYN Post</th>
<th>AI Pre</th>
<th>AI Post</th>
<th>PS Pre</th>
<th>PS Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>N of people supplying written responses</td>
<td>18</td>
<td>16</td>
<td>21</td>
<td>20</td>
<td>19</td>
<td>14</td>
</tr>
<tr>
<td>Positive</td>
<td>7</td>
<td>10</td>
<td>15</td>
<td>8</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>Negative</td>
<td>6</td>
<td>4</td>
<td>11</td>
<td>3</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Hope</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Action</td>
<td>11</td>
<td>6</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>4</td>
<td>12</td>
<td>10</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
<td>27</td>
<td>52</td>
<td>33</td>
<td>46</td>
<td>16</td>
</tr>
<tr>
<td>Average number of ideas</td>
<td>2.19</td>
<td>1.69</td>
<td>2.47</td>
<td>1.65</td>
<td>2.42</td>
<td>1.14</td>
</tr>
</tbody>
</table>

Note. SYN = synergenesis; AI = appreciative inquiry; PS = problem solving.
The data provide limited support for Hypothesis 1. While all the averages in Table 3 are in the expected directions, the average rating of ideas from the AI group are not statistically different from those coming from the problem-solving group. This may be

Discussion

The data provide limited support for Hypothesis 1. While all the averages in Table 3 are in the expected directions, the average rating of ideas from the AI group are not statistically different from those coming from the problem-solving group. This may be
a result of the small sample size. Support for hypothesis one does come from the semantic map results, which show the problem-solving condition having a dramatic negative impact, postdiscussion, on the number of ideas compared with the AI group, particularly the number of positive and action ideas. The ratings of ideas do provide support for Hypothesis 2, however, with two of the three measures of idea generativity
in Table 3 being significantly higher than either of the two other conditions. The semantic map data, however, does not show much difference between the synerogenesis condition and the AI condition while the postsemantic maps show quite a bit of difference between those who participated in problem-solving and those in the other two “appreciative” conditions. The second author, who facilitated all the sessions, found the most noticeable difference was that people in the problem-solving session expressed more frustration with the company and talking about problems appeared to make them even more frustrated. They appeared to want to finish the session quickly, many saying that they could not stay to complete the postmeeting surveys. In general, these participants spent little time writing answers to the open-ended questions. After the AI and synergenesis meetings, however, people were more interested and took more time to complete the surveys. Many hung around even after completing the surveys to chat in groups and get to know each other and talk about the session. Consistent with claims made by AI advocates, the AI and Synergenesis sessions appeared to stimulate more engagement and dialogue among participants.

These observations are also consistent with research that shows people become more analytical and focused when they are in a negative state of mind (Fredrickson & Branigan, 2005; Forgas, Laham, & Vargas, 2005). Participants in the problem-solving condition were willing to fill out close-ended surveys (not described in this study) after their meetings, but when they were asked to write answers for the open-ended questions used for the semantic map analysis, they expressed a little dissatisfaction and more than half did not want to do it. What this suggests is that engaging people in group problem solving, which is often done in organization development to increase people’s commitment to change and reduce resistance to change, may increase the salience of their frustration with the organization. If the change strategy involves heightening people’s dissatisfaction with the current state, that may be all to the good. If the change strategy, however, is to increase people’s sense of engagement and commitment to the organization, this approach may well backfire.

The lack of differences in the novelty of ideas among the three conditions bears discussion. As shown in Table 3, novelty scored lowest among the three generativity ratings in all conditions. There have been numerous recognition programs in this organization in the past. Novelty was rated according to “is this idea novel/ new, which has not been done before in this organization?” It is possible that many, if not all of these ideas had already been broached at some point in the organization’s history. Discussing these results with the raters, it surfaced that while they did not think ideas were innovative, they found some compelling and wanted to implement them. In their view, past recognition efforts failed not because of the ideas, but because they were not implemented well.

To add richness to this discussion, Tables 5 to 7 show the ideas rated as most innovative, compelling, and practical for each of the three conditions. These tables were created by showing the two or more ideas that got the highest total combined rating from all five judges. If only one idea got the highest rating, all the ideas receiving the next highest score were included. Interestingly, while there was some overlap in ideas
between the three conditions, there were also many different ideas. One of the most striking differences is the nature of the ideas from the problem-solving groups versus the more appreciative ones. For the most part, the ideas from the synergenesis and AI groups appear quite relevant to the issue of employee recognition. Many of the ideas from the problem-solving groups, on the other hand, appear to focus on general concerns (e.g., When full-time jobs become available interns should have top priority;...
jobs should be posted internally before going out). These data suggest that not only do appreciative approaches to ideation in employee engagement strategies lead to more generative ideas but they also create a greater focus on the specific area of managerial interest, while a problem-solving strategy may lead employees to use the opportunity to surface other issues that are “problems” for them.

It is worth noting interesting similarities with a study published after this research was done. Carlsen, Rudningen, and Mortensen (2014) used appreciative interviews into distinct qualities of ideal work practices when at their best to produce “cards,” each one consisting of a work practice that had emerged during the interviews. They then used these cards to stimulate small group conversations, slightly different in design but similar in intent to synergenesis sessions, as a generative ideation process. One finding in their study was the importance of participants holding and dealing the cards. They note that taking cards in hand signals a dialogic shift in genre where participants are invited to influence both categories and content. They point to the tactile nature of holding the cards and to insights into grounded cognition (Barsalou, 2008; Robbins & Aydede, 2009), to argue that tactile engagement with the card has an opening-up and seeding function in the interaction, “being vehicles for stories and

### Table 6. The Most Interesting Ideas.

<table>
<thead>
<tr>
<th>Synergenesis</th>
<th>Appreciative inquiry</th>
<th>Problem solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 Performance evaluations should be done annually. Performance evaluation should be a fair process—some departments give 5s other department do not give 5. There should be a consistent process of evaluation.</td>
<td>23 The organization should provide you with training programs for Microsoft Office or any other new software.</td>
<td>22 Provide opportunities for training and development for all employees.</td>
</tr>
<tr>
<td>24 The interns should be used to their full potential and provided challenging opportunities.</td>
<td>23 Knowledge should be transferred; there should be some form of mentoring, rather than working in dark, employees should leave institutional knowledge and share the knowledge with others.</td>
<td>22 Jobs should be posted internally before going out.</td>
</tr>
<tr>
<td>24 If you have responsibilities and if you feel committed to them; managers should foster commitment to work, interns should feel accountable to work.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The numbers to the left are the rating the idea received, out of 25.
Table 7. The Most Practical Ideas.

<table>
<thead>
<tr>
<th>Synergenesis</th>
<th>Appreciative inquiry</th>
<th>Problem solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 Performance evaluations should be done annually. Performance evaluation</td>
<td>23 There should be professional development and training; promotion because of your</td>
<td>23 Ask 3-4 areas of interests before hiring interns and place them in one of</td>
</tr>
<tr>
<td>should be a fair process—some department give 5s other department do not</td>
<td>work and not because who you know.</td>
<td>those areas of interest.</td>
</tr>
<tr>
<td>give 5. There should be a consistent process of evaluation.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Management should provide opportunities to increase cooperation among</td>
<td>23 There should be reference materials or procedure manual, formal training session</td>
<td>22 Jobs should be posted internally before going out.</td>
</tr>
<tr>
<td>colleagues/coworkers.</td>
<td>for each new employee to tell more about their job responsibilities.</td>
<td></td>
</tr>
<tr>
<td>23 Verbal recognition from leadership. For example, both my GM and VP came</td>
<td>23 Knowledge should be transformed; there should be some form of mentoring, rather</td>
<td>22 There should be a better onboarding process for the interns.</td>
</tr>
<tr>
<td>up and said congratulation—took time out of their busy schedules and stopped</td>
<td>than working in dark, employees should leave institutional knowledge and share the</td>
<td></td>
</tr>
<tr>
<td>by and said thank you for a job well done.</td>
<td>knowledge with others.</td>
<td></td>
</tr>
<tr>
<td>23 Provide opportunities for growth (provide a pathway). Career advancement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>opportunities.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Building trust and integrity in work environment.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Opportunities to communicate your ideas with experienced people in the</td>
<td>23 Manager should provide one-on-one personal thank you to employees when a job is</td>
<td></td>
</tr>
<tr>
<td>department.</td>
<td>done well.</td>
<td></td>
</tr>
<tr>
<td>23 Manager should provide one-on-one personal thank you to employees when a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>job is done well.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The numbers to the left are the rating the idea received, out of 25.

opinions to be communicated verbally as participants move from individual to collective sensemaking” (Carlsen et al., 2014, p. 305). This raises interesting areas for future research into synergenesis. What is the impact of having a deck of stories in hand, rather than simply relying on memory of stories just heard (as in classical AI)? Are there ways in which having a concrete, shared repository of stories affects the dynamics that take place in a synergenesis group? What are the effects of authoring written
stories before a synergenesis session? Does this in some way create a seeding effect prior to the ideation process that is different from classical AI Discovery? What about the ownership of the deck of stories? Does it matter if these stories were written by the participants or would any deck of good stories have the same effects?

There are limitations to the study worth noting, not the least of which is the low $N$ for a statistical study. However, the difficulty of finding sites for naturally occurring field experiments makes reporting of these findings worth considering, as does the consistency in the various data in showing distinct differences in the generativity of problem solving versus appreciative approaches to engaging employees in the ideation process during a change effort.

A second limitation worth noting is the use of paired interviews before the problem-solving groups’ brainstorming. We did this to strengthen our ability to attribute any differences in results to the factors we used to explain why AI might be more generative and not simply a result of increased cognitive stimulation from having participated in interviews prior to ideation in small groups. While we believe that is strength of the study, it also means we have compared an idiosyncratic form of problem-solving against AI. Cognitive stimulation from the interviews might have increased the generativity of the problem-solving process and caused the nonsignificant differences found for hypothesis one. Facilitating problem-solving groups as they would normally be run, without interviews prior to brainstorming, might have led to more significant differences.

Another limitation of the study concerns the issue of whether generativity in organizational change requires only one really good idea. We designed this study on the assumption that a process that produces more ideas rated to have generative potential is a more generative process, using the average generative quality of all ideas produced as the dependent measure. As a change mechanism, however, generativity may rely on only one truly generative idea to be successful. This raises another limitation of the study—the use of an “expert panel” to rate the generativity of the ideas. A better study would follow, longitudinally, what actually happened to these ideas and one could argue that studying the generativity of any idea can only be done in retrospect, since what makes an idea generative depends on the context in which it is used. Unfortunately, we did not have the ability to do that. We do not know what inherent limitations there might be in the ability of a panel of senior managers to predict how likely an idea would motivate new actions, though we are heartened by the fact that many told us, informally, that they did find some of the ideas compelling. That suggests a higher likelihood that they subsequently did lead to action.

It should also be noted that this study only compares the generativity of problem solving and synergenesis against the Discovery phase of AI. A full AI process, of course, will also include Dream and Design phases, and it may be that more generative ideas will be produced during these later phases. The results of the semantic map data, coupled with our experience of how participants reacted to being asked to provide it (described above), offer some evidence to suggest that the Discovery process of classical AI may be more of a priming process by creating higher levels of engagement and interest in the focal topic. AI’s superiority as a generative process over problem solving may be more evident when the later
Bushe and Paranjpe

stages of Dream and Design are taken into account. Though Bushe (2010) found many of the generative ideas that produced transformational change emerged during the Discovery phase, all the appreciative inquiries in that study used synergenesis. Future longitudinal studies of AI with an interest in generativity could look more closely at what actually contributes to the emergence of generative ideas, and where they emerge.

The difference in ratings of ideas produced from classical AI and synergenesis suggest that practitioners may want to consider this and other ways to increase the output of generative ideas during Dialogic OD efforts. Practitioners often appear to focus most of their effort on building better relationships and increasing the quality of dialogue, but is that enough? Can we assume that better ideas already exist, and just need a forum to be heard, or will priming produce more generative processes? Lukensmeyer’s (2013) successful civic engagement process includes a step of increasing the knowledge and complexity of thinking of participants before ideation and it might be that synergenesis, by exposing participants to more stories than they would normally be exposed to in classical AI, and by focusing conversation around each story, increases the complexity of participant thinking. Practitioners may want to include ways of priming people to increase the complexity of their thinking about the focal issue before the ideation stage of any change process. There is also the intriguing line of inquiry opened up by Carlsen et al.’s (2014) observations about the priming nature of tactile stimulation for producing generative ideas. Practitioners may want to experiment with ways of priming participants somatically, with perhaps visual metaphors and ways of “holding ideas” in their hands.

This is a small, simple study, but we hope it increases OD researchers’ interest in the nature and effects of generativity in organizational change processes. If Bushe (2013a, 2013b; Bushe & Marshak, 2014) is correct and generative images are one of the core change processes underlying the success of Dialogic OD efforts, than we need to pay more attention to them—what they are, where they come from, how and when to introduce them into a change process—in our studies of organization development and change.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


(Eds.), *Organizational generativity: The appreciative inquiry summit and a scholarship of transformation* (Vol. 4, pp. 3-67). Bingley, England: Emerald.


