# From Diversity to Creativity: Stimulating Group Brainstorming with Cultural Differences and Conversationally-Retrieved Pictures

**Hao-Chuan Wang[1], Susan R. Fussell[1,2], Dan Cosley[1,2]**
[1]Department of Information Science, [2]Department of Communication
Cornell University, Ithaca, NY 14850 USA
[hw346, sfussell, drc44] @cornell.edu

**ABSTRACT**
Group brainstorming, or collaboratively generating ideas through idea sharing, demands diverse contributions to spark more ideas and improve creativity. One approach to supporting group brainstorming is to introduce conceptual diversity. In this study, we evaluate the effects of two sources of diversity on group brainstorming: cultural differences internal to multicultural groups and pictures related to the conversation retrieved by a computer agent. The pictures generally enhanced performance as measured by both originality and diversity of ideas. The pictures also helped to convert cultural diversity into a creative outcome, the diversity of ideas generated. We argue that with appropriate technology mediation, cultural diversity may be used strategically to enhance task outcomes.

**Author Keywords**
Intercultural collaboration, group brainstorming, group creativity, cultural diversity, creativity support tools

**ACM Classification Keywords**
H5.3. Group and Organization Interface: Computer-supported cooperative work

**General Terms**
Design, Experimentation, Human Factors

## INTRODUCTION

Generating ideas is an integral component to work in many domains. Designers propose designs of products with improved functions or appearance. Engineers think of strategies to solve technical problems. Scientists generate hypotheses and models to advance understanding of phenomena. Yet creativity is difficult, with the limits in individuals' perspectives, experiences and knowledge making individual idea generation a challenging task. Distributed social and technological support that shares the agency of creativity is essential [1][10].

Group brainstorming is one widely practiced approach that uses social means to address the cognitive bottleneck of idea generation [23]. Group brainstorming engages multiple individuals to collaborate by communicating and sharing ideas in groups. This broadens the knowledge base available for idea generation and allows creative effort to be aggregated. Further, overhearing others' ideas can help individuals think of ideas that they would not explore on their own [19][21]. Ideas in group brainstorming, then, are not only the *products*, but also the *inputs* for stimulating and moving the brainstorming cycle forward.

But two (or more) heads are not necessarily better than one. Structural and contextual factors may influence group brainstorming outcomes. Consider a group of people who possess exactly the same knowledge and generate ideas in the exact same manner (e.g., proposing ideas in the same order). Collaboration in this case is unlikely to enhance ideation because group members will not be able to access extra concepts beyond those already accessible individually. Though this is an extreme example, it highlights the crucial role of diversity in brainstorming. Empirical work also shows that when shared ideas are semantically similar, such as those from only a few homogeneous topics,


**Figure 1. Brainstorming with two sources of diversity. Internal diversity: Diverse concepts individuals hold. External diversity: Stimuli provided by an external agent.**

brainstorming performance is worse than when topical diversity is present [20][21][24].

Thus, an important issue for supporting group brainstorming is to consider ways of supplying sufficient *diversity* of concepts. These sources of diversity may be either internal or external to the group. Internal diversity refers to background differences between group members that originate from long-term learning and socialization, such as knowledge, experiences, and cultural background [1]. In international workgroups, cultural differences may naturally serve as a source of diversity.

External diversity, on the other hand, refers to stimulation supplied by external agents. One approach shown to be effective for sparking ideas is to automatically retrieve and present pictorial stimuli based on the content of the ongoing conversation [31]. The injection of pictures relevant to the conversation adds both new stimuli and a second, visual channel for stimuli. Perceptual processes of selective attention and subjective interpretation may then provide diversity above and beyond the original verbalizations.

Figure 1 shows an illustrative scenario of how the two sources of diversity may both contribute to creative ideation. Based on their conceptual models of typical pets, the person on the left might be more likely to say “cat” while her partner might think “rabbit” when brainstorming about pets to adopt. Neither is likely to think “gerbil” on their own. However, if the computer agent recognizes that the person on the left said “cat”, it might retrieve a picture of a cat staring at a gerbil. This might then help the group suggest a gerbil as a pet—and activate concepts related to other possible pets that are related to gerbils.

In the current study, we examine how one internal and one external source of diversity—cultural differences and retrieved pictures, respectively—affect the productivity and diversity of ideas generated in group brainstorming. Diversity is not only an important input to groups, but also an important output that people demand in work, such as design alternatives and multiple solutions to a problem. We compare intercultural and intracultural dyads consisting of people with American and Chinese cultural backgrounds. We also evaluate two external stimulation strategies, one that presents pictures closely related to the topics of conversation and another that presents pictures that are less related but possibly more stimulating. We find that picture support in general enhanced brainstorming outcomes, and that intercultural groups especially benefited from this external stimulation. The results suggest that technological design grounded in a systematic understanding of cultures can play a valuable role in converting cultural diversity into creative outcomes.

## BACKGROUND

This work is motivated by a theoretical and empirical understanding of how ideas are generated in brainstorming groups. In this section, we first review socio-cognitive perspectives on group brainstorming, highlighting the importance of diversity. We then discuss the properties of our sources of diversity, cultural differences and stimuli supplied by external agents. For external stimuli, we focus on picture support and discuss the properties that might make it uniquely useful. We then discuss the hypotheses to be tested in this study.

### The Socio-Cognitive Model of Group Brainstorming

Group brainstorming may be best described as using groups to perform idea generation. Group brainstorming involves both the social process of idea exchange and the cognitive process of idea generation. A complete picture of how ideas are generated through group brainstorming requires understanding both of these processes, as well as the interplay between them.

At the cognitive level, ideas are developed from individuals’ prior knowledge. Concept retrieval from the semantic memory is thus at the core of idea generation as no ideas may be generated without retrieving relevant concepts.

A prevalent model of memory posits that semantic knowledge forms a network, in which nodes of the network represent concepts, and links between nodes denote semantic associations between concepts [2]. A number of variations on the theory exist, but a common characteristic is that concept retrieval is not a standalone action. The activation of one concept will make other interconnected concepts more accessible, forming a chain reaction [2][26]. For example, thinking about the concept “pet” may activate the concept “cat,” which in turn might activate concepts like “cute” and “playful.” Thus, activation of one concept spreads through the network, with the degree of influence attenuating with successive steps across nodes.

This model helps explain why brainstorming often results in a series of semantically related ideas. Either thinking of or hearing an idea based on a given concept contributes to the retrieval of related concepts that foster a new cycle of ideation [3][21]. Ideation therefore may spread out in a chained fashion. The effort required to generate the first idea of one category may be much larger than that required to generate follow-up ideas with minor variations. Through group brainstorming, individuals hearing other people’s ideas can more easily reach concepts far away in their own semantic network. Thus, shared ideas are both products and social inputs that may stimulate thinking [9][20][21][24].

However, the potential advantage of group brainstorming may not be realized if shared ideas fail to stimulate thinking. This may happen when negative social processes, such as evaluation apprehension (fearing to express ideas because they might be viewed negatively) and production blocking (taking turns to speak up) reduce the quantity and quality of ideas (i.e., stimuli) expressed [8]. Failure to stimulate may also happen when group members possess similar knowledge or thinking styles. In this case, the stimulating utility of shared ideas is lower due to redundancy. One clear message derived from the socio-cognitive perspective is the importance of diversity among shared ideas for sustaining cognitive stimulation and productive brainstorming.

### Internal Diversity: Cultural Differences

To sustain group brainstorming, one source of conceptual diversity is variation among group members in terms of background knowledge and expertise. National cultures are one source of background differences available in today's workgroups given the increasing popularity of international and intercultural collaboration. Thus, creating multicultural groups is a promising approach for generating diversity.

Psychology research suggests that East Asians and Americans have broadly different cognitive styles [16][22]. East Asians tend to allocate greater attention to contextual information, such as background objects in a picture. At higher levels of cognition, such as interpretation and categorization, East Asians often associate concepts based on ecological relations (e.g., associating cow and grass because cows eat grass) [22][25]. Americans, on the other hand, attend primarily to focal information, such as foreground objects in a picture, and to categorize based on shared properties (e.g., associating cows and sheep, because both are farm animals). Overall, the cognitive style of East Asians tends to be more holistic, and that of Americans tends to be more analytical. Neither is "better", and group brainstorming might profit from both.

People from different cultural backgrounds may also have differences in their semantic networks that may provide diversity. For example, the concept of "turkey" may be more central to Americans' semantic networks than to Chinese due to the importance of Thanksgiving as an American holiday, while the Confucian concept of filial obedience may be completely absent from most Americans' semantic networks.

Cultures also differ in social orientations and collaborative behaviors [7][27]. East Asian cultures are generally more collectivistic and relationship-oriented [16][29]. When working in groups, East Asians may be less comfortable with dissent, and tend to conform to other people's opinions in order to avoid threats to interpersonal relations. In contrast, Americans are generally more individualistic and task-oriented. In teamwork, Americans may be more comfortable with sharing their thoughts directly without worrying as much about other people's opinions or feelings.

One threat to intercultural group brainstorming is that negative social factors like evaluation apprehension may be more prominent among some cultures (e.g., Chinese) and under some conditions (e.g., talking about sensitive topics or communicating face-to-face) [30]. The social and communicative barriers that block idea sharing thus may be even larger in intercultural groups than in intracultural groups. Relying on cultural differences as the sole source of diversity does not guarantee better outcomes.

### External Diversity: Stimuli Supplied by External Agents

Another approach to introducing conceptual diversity is to supply extra stimuli from outside the group, enriching the environment so that it affords more diverse perception. For example, artists and designers often decorate their studios with a variety of external objects, with the goal of stimulating creative ideas.

Computer agents are one potential tool for supplying external stimuli. Agents that find materials related to a given context (such as local search results that consider both keywords and location on mobile devices) are a promising tool for supporting informal idea generation. The design space for these agents is large: what drives the retrieval of stimuli, which algorithms are used, and how and when the stimuli are presented are all decisions that might impact the utility of these agents.

We propose that showing pictures that are retrieved automatically in real-time, based on the ongoing content of a brainstorming conversation, is one design useful for stimulating ideas [31]. This tactic has several benefits. First, using conversations as the driving force and search engines for retrieval may free the system from the need for domain-specific knowledge models, widening the scope of tasks to which they might apply. These conversations also contain rich semantic information that suggests the topics people are attending to at a given time. The agent can use this knowledge to manipulate the alignment of the stimuli with the conversations, providing fine-grained control over the degree of coherence or diversity of selected stimuli to the ongoing conversation and allowing the system to be tuned to specific theoretical concerns and/or task needs.

Second, using pictures as extra stimuli may be more effective than language in stimulating idea generation. Pictures provide a natural way to present multiple ideas at once. For example, a picture of a car may introduce other topics through the visual context, such as the color of the car, the street view and the traffic pattern etc. A linguistic statement, on the other hand, typically has a narrower conceptual scope (e.g., "there is a car"). Due to individual and cultural differences in perception [22], a picture may be interpreted in multiple ways, which can be a valuable basis of conceptual diversity.

Third, picture support may also address social barriers. As stimuli prepared by agents, they may be perceived as external artifacts rather than as representations of personal opinions. Group members may thus be more comfortable commenting on and developing ideas from the pictures.

## THE CURRENT STUDY

In the current study, we examine the effects of internal (cultural differences) and external (picture support) sources of diversity in group brainstorming. For cultural diversity, we compare two-person groups with three types of cultural composition: all Americans (AA pairs), all Chinese (CC pairs), and intercultural groups consisting of one American and one Chinese (AC pairs).

For picture support, we compare two plausible strategies to a baseline of no support. One strategy, *Congruence*, uses standard information retrieval techniques, creating a query from words recently used in the conversation and returning the picture labeled with the most similar set of keywords. The other strategy, *Stimulus*, attempts to increase the diversity of the stimuli it provides to the group. Like *Congruence*, it finds a set of relevant pictures based on keyword similarity, but instead of just choosing the most relevant picture, it considers both the picture's similarity to the ongoing discussion and the diversity and rarity (infrequency) of the keywords the picture is labeled with. Based on the socio-cognitive model of brainstorming, we might expect the *Stimulus* method to supply more diversity, and thus be more useful for group brainstorming.

We consider "useful" in two dimensions. First, we evaluate a group's productivity, defined as the number of original ideas generated by the group. We also evaluate the breadth of a group's ideas, as measured by the average semantic similarity between all pairs of ideas generated by the group. Based on the work described above, we hypothesize how group cultural composition and picture support may influence productivity and breadth of concepts covered. In terms of productivity:

*H1*: Both types of picture support will enhance productivity compared to no support; further, the *Stimulus* method will lead to better productivity than the *Congruence method,* because external diversity improves stimulation.

*H2*: Intercultural groups (AC pairs) will have better productivity than intracultural groups (AA and CC) because cultural diversity contributes to productivity.

*H3*: The effect of picture support on productivity will be greater for intercultural groups (AC pairs) than intracultural groups (AA and CC). This is because individuals from different cultures may perceive and interpret pictures even more diversely [5][22], therefore pictures would trigger more ideas for intercultural groups.

In terms of breadth of concepts:

*H4:* The *Stimulus* method will result in greater breadth of concepts than other approaches because the *Stimulus* method emphasizes choosing pictures that contain multiple or rare topics, and thus ones that may semantically diverge from topics the group already explored.

*H5:* Intercultural groups (AC) will cover broader concepts than intracultural groups (AA and CC) because of conceptual diversity between cultures.

*H6:* The effect of the *Stimulus* method in broadening concepts is larger for intercultural groups (AC) than intracultural groups (AA and CC). Similar to the mechanism behind *H3*, because cultures differ in picture perception and interpretation, *Stimulus* support would be even more powerful for intercultural groups.

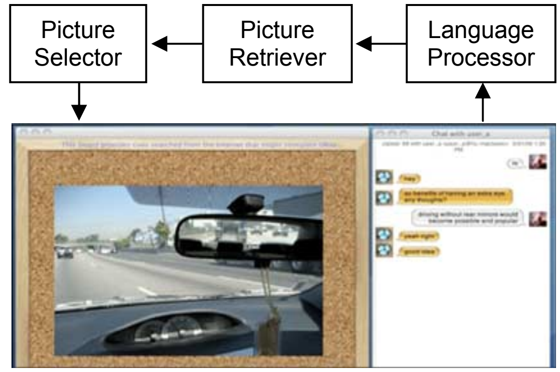


**Figure 2. The agent monitors the group conversation (right) and selects pictures to display to the group (left).**

## METHOD

### Experimental Design and Procedure

Two-person groups were asked to perform three similar brainstorming tasks with three types of picture support: *Congruence*, *Stimulus*, and *None* (i.e., no support)[1]. Three types of cultural groups were formed: two Americans (AA pairs), two Chinese (CC pairs), and one American and one Chinese (AC pairs). Cultural composition was a between-groups manipulation, while picture support and brainstorming tasks were within-subject manipulations. Their orders were counterbalanced using Latin squares.

Participants were brought to the laboratory and instructed about the brainstorming topics and provided with four conventional brainstorming rules [23]: (a) the more ideas the better; (b) the wilder the ideas the better; (c) combination and improvement of ideas are better; and (d) avoid evaluating others' ideas. Participants were instructed to brainstorm in groups, but they were not specifically informed about their partner's identity or background. Groups engaged in free conversations to brainstorm via a chatroom and had 15 minutes for each task. Between tasks, we switched which type of picture support they received.

### Participants

There were 54 participants (65% female) recruited from a large U.S. university and surrounding community. Of these, 29 were self-identified Americans living in the U.S. or Canada who had grown up in the U.S. or Canada and spoke English as their native language. The remaining 25 participants were self-identified Chinese speaking Chinese as their native language but who were fluent in English. Although they were all currently studying or working in the

[1] We chose showing no pictures rather than random pictures as the baseline for two reasons. First, no support resembles the current scenarios of teamwork, increasing the value of the work. Second, in an unpublished pilot study, showing random pictures did not appear to help brainstorming.

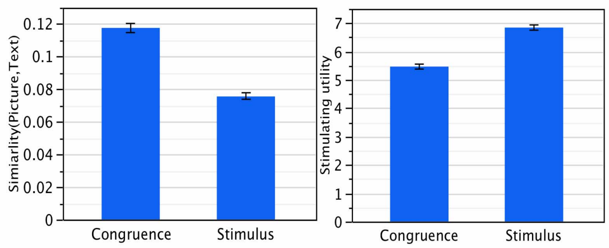


**Figure 3. Validating picture selection methods. Left: Similarity score (TF-IDF). Right: Stimulating utility score.**

U.S., the majority of them grew up in China, Hong Kong or Taiwan and had been in the U.S. for less than 2 years.

Participants were randomly assigned to brainstorming groups. The majority of the participants (83%) reported that they did not know their fellow group members prior to the study. There were a total of 27 groups formed (10 AAs, 9 ACs and 8 CCs).

### Tasks

Teams performed three brainstorming tasks of equivalent difficulty. The "extra thumb" and the "extra eye" questions, which have been used in many studies [9][30], ask participants to brainstorm about the benefits and difficulties for people having a hypothetical extra thumb or an extra eye at the back of their heads in the future. The "having wings" task is a newly designed task that asked participants to brainstorm about the benefits and difficulties for people having a pair of wings in the future.

### Picture Support

Picture support was implemented using Wang et al.'s IdeaExpander prototype, which monitors brainstorming and retrieves pictures to show based on the contents of ongoing conversations [31]. Figure 2 shows a screenshot and the high-level system architecture. Participants brainstorm in a chat window on the right, while the system displays pictures it chooses based on the conversation on the left. The system consists of three main components.

*Language processor.* IdeaExpander monitors the chat conversation to identify keywords currently being discussed. Because brainstorming conversations include both on-task and off-task remarks, a machine learning classifier trained by data from an earlier study [30] is used to determine whether a remark contains an idea or not. Accuracy of this binary classification is 80% (Cohen's Kappa=.61).

*Picture retriever.* IdeaExpander uses keywords drawn from remarks classified as containing ideas to retrieve candidate pictures to show. The study used a labeled picture database specific to tasks of the study. Previous coding of brainstorming logs resulted in a coding scheme containing 110 (thumb), 118 (eye) and 112 (wings) idea categories [30]. We collected 60 pictures for each task from Flickr and coded each picture with the applicable idea categories (Krippendorff's alpha=.5). We then labeled each picture with the tags it already had from Flickr and the words contained in the codebook descriptions of the idea categories. The agent matches conversational turns it classifies as containing ideas against the keywords in the database using TF-IDF in order to retrieve a *relevant set* of pictures. In the current study, the four most relevant pictures were retrieved each time.

*Picture selector.* IdeaExpander then selects pictures to display, using either the *Congruence* or the *Stimulus* method. As described earlier, *Congruence* emphasizes coherence between pictures and recent ideas, retrieving the picture from the relevant set with the highest TF-IDF score. For the *Stimulus* method, we defined a stimulating utility score that prefers pictures that contain multiple idea categories or categories that are less commonly discussed (i.e., rare ideas). We used the dataset from [30] to estimate the probability of generating each idea, and weighed each idea *i* as log(1/probability of idea *i*). The utility score for a picture is the sum of weighted scores of the ideas pertaining to it. The *Stimulus* method selects the picture from the relevant set with the highest utility score. In both methods, pictures that have already been shown are excluded.

Figure 3 shows the mean TF-IDF similarity between pictures and conversational contents, and the stimulating utility of pictures selected by each algorithm. As expected, pictures selected by *Congruence* were more similar to the ideas that triggered them than pictures selected by *Stimulus* ($t[1744]=-12.98$, $p<.0001$), while pictures selected by *Stimulus* had higher stimulating scores than pictures selected by *Congruence* ($t[1744]=10.84$, $p<.0001$).

For both versions, the agent updates the picture space (the board on the left of Figure 1) with a new picture—if available—every three seconds.

## MEASURES

We evaluate groups' outcomes with two measures, productivity and breadth of ideas. Productivity addresses *outcomes*: the quantity and originality of the ideas that are generated. Breadth of ideas addresses both outcomes—one goal of brainstorming is to generate a variety of ideas—and *process*, in the sense that ideas generated may stimulate further ideas during the conversation by activating a wider variety of concepts among group members.

### Productivity

To account for both quantity and originality aspects of idea generation, we coded the brainstorming data with a two-level strategy. At the first level, we asked coders to classify whether each conversational turn contained an idea or not. Turns codes as containing an idea were then coded as either duplicates (minor variations of an idea already contributed) or having originality (ideas not yet proposed by the group).

Two coders coded conversations from three randomly selected brainstorming groups (about 13% of the data) to assess reliability. Inter-coder agreement was satisfactory both at the first level (coding idea versus no-idea, Cohen's

Kappa= .95) and the second level (coding duplicate versus having originality, Cohen's Kappa= .80).

We used the number of turns coded as containing original ideas as our measure of productivity.

### Breadth of Concepts

Breadth of concepts was conceptualized as the average semantic distance between any two original ideas generated in a brainstorming session. Intuitively, semantic distance is how far apart the concepts expressed by the ideas would be in a semantic network where concepts are represented and organized as a graph. For example, "cow" and "sheep" would have a lower semantic distance than "cow" and "electron" in most people's semantic networks.

To operationalize this idea, we started from a semantic network based on a database of word association norms [18] that was generated by empirical studies that ask people to explicitly associate words. For example, an experimenter might ask participants to say the first three things that come to mind when they see the word "music". This procedure generates a set of word association frequencies that does not, however, capture all possible word associations. These hidden associations may be uncovered by applying the statistical procedure of singular value decomposition (SVD) to map sparse raw data into a multidimensional space that represents words as vectors of numerical features, similar to what latent semantic analysis does [11]. Using SVD on the word association frequencies results in a multi-dimensional word association space (WAS) [28].

Table 1 illustrates using some keywords ("bald", "glasses", "industry", "music", and "beauty") as queries to retrieve the words most related to them from the semantic space. We see both associations that one might make directly, such as "bald" and "scalp", but also associations uncovered by the SVD procedure such as "glasses" and "squint" that people would be unlikely to make directly.

With the WAS, it is straightforward to assess the strength of association between two concepts by computing the cosine similarity between the sets of words used to express each concept. Cosine similarity for most cases ranges between 0 and 1, where 1 represents perfect association, so we convert it to a distance metric, where higher scores represent more

**Table 1. Lists of words strongly associated to the queries (words in bold) retrieved from the semantic space.**

| **bald** | **glasses** | **industry** | **music** | **beauty** |
|---|---|---|---|---|
| scalp | lens | employment | band | pretty |
| headband | contacts | career | instrument | beautiful |
| shampoo | blur | business | woodwind | attractive |
| beard | vision | occupation | oboe | gorgeous |
| comb | sight | corporation | flute | ugly |
| brush | squint | agency | viola | handsome |
| strand | eyelash | task | trombone | cute |
| mustache | eye | factory | tuba | model |
| forehead | blind | duty | guitar | elegant |

diversity, by taking its inverse. And to generate a metric that represents the breadth of the whole set of ideas, we compute the *average* similarity between each pair of original ideas generated by a given group, and then invert it to a distance measure. We then take log-transformation for normality of distribution.

*Breadth-of-concepts* = log(1/*avg. cosine of all idea pairs*)

## RESULTS

The main units of analysis were groups, because the hypotheses primarily concerned how different types of picture support influenced the outcomes of different cultural groups. We used mixed model ANOVAs to account for possible interdependencies caused by repeated measures or social influences within groups [14]. The type of mixed models adjusts the estimation of variance and typically provides more conservative results.

The basic model for analyzing group outcomes treated brainstorming trial and group as random variables. Brainstorming trial was nested within group. Group cultural composition, picture support, and the interaction between the two variables were included as fixed effects. Brainstorming topic was used as a covariate. Note that in mixed models, it is standard to estimate the degrees of freedom associated with the denominators by using Satterthwaite's approximation. Non-integer degree of freedom results may occur (see [15] for details). To estimate effect sizes, we computed Cohen's $d$ from the sample means and standard deviations [6]. For computing effect sizes of picture support, correction was applied to account for within-subject correlations [4][17].

### Talkativeness

Group brainstorming is a task relying on using language to verbalize ideas. Cultural differences in conversational behaviors [27][30] and linguistic fluency (e.g., speaking a second language) may thus be confounded with brainstorming performance. To examine this possibility, we looked at whether cultural groups differed in talkativeness using a linear model of the form outlined earlier and the number of words typed by pairs as the dependent measure.

Group cultural composition had a nonsignificant effect on talkativeness ($F[2,22.9]=2.18$, $p=.13$). Because this effect approached significance, we examined differences between groups further using post-hoc t-tests. AA groups typed more words than CC groups ($t[22.8]=2.06$, $p<.05$), but there were no significant differences between AA and AC or between CC and AC pairs. These results weakly suggest that cultural composition of a group may influence overall talkativeness. We thus included this factor as a covariate in our later analyses. Note that picture support did not affect talkativeness. This suggests that if picture support were effective, its mechanism is unlikely to involve promoting talkativeness.

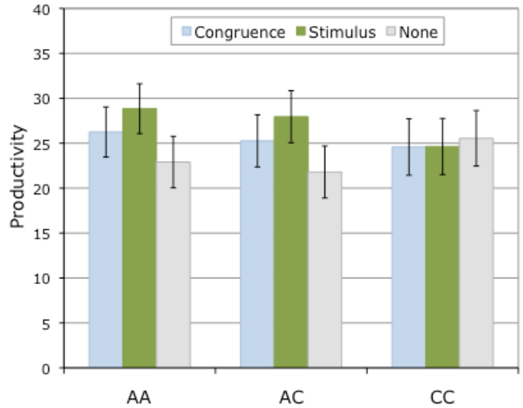


**Figure 4. Adjusted productivity by group composition and picture support. Means and standard errors were estimated by the linear mixed model.**

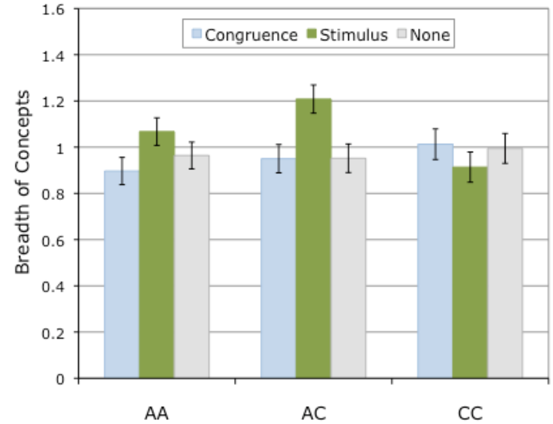


**Figure 5. Breadth of concepts by group composition and picture support. Means and standard errors were estimated by the linear mixed model.**

### Productivity

To test H1, H2 and H3, we constructed a linear mixed model to evaluate the effects of cultural groups and picture support on group productivity (number of original ideas by groups). We included talkativeness as a covariate in the analysis. Figure 4 shows adjusted productivity scores after accounting for the influence of talkativeness.

In support of H1, there was a main effect of picture support ($F[2,44.2]=8.04$, $p<.001$). Post-hoc t-tests showed that both the *Stimulus* method and the *Congruence* method led to better productivity than no picture support (*Stimulus* versus *None*: $t[43.3]=4.01$, $p<.0001$, Cohen's $d=.29$; *Congruence* versus *None*: $t[45.5]=1.99$, $p<.05$, Cohen's $d=.08$). There was also a trend for the *Stimulus* method to provide better productivity support than the *Congruence* method ($t[44.4]=1.88$, $p=.07$, Cohen's $d=.20$).

Hypothesis H2 was not supported by the analysis. Group cultural composition did not have a main effect on productivity ($F[2, 24.7]=.05$, *n.s.*). AC pairs did not produce more ideas than intracultural groups.

To test H3, we focused on how different cultural groups performed when picture support was available and not available. There was a significant interaction between group cultural composition and picture support on productivity ($F[4, 40.25]=3.37$, $p<.05$). Intercultural groups generated more ideas when using either type of picture support than when using no support (AC & *Stimulus* vs. AC & *None*: $t[38.7]=4.00$, $p<.0005$, Cohen's $d=1.01$; AC & *Congruence* vs. AC & *None*: $t[40.5]=2.20$, $p<.05$, Cohen's $d=.28$). AA pairs also generated more ideas when using either picture support than no support (AA & *Stimulus* vs. AA & *None*: $t[43.7]=3.88$, $p<.0005$, Cohen's $d=.42$; AA & *Congruence* versus AA & *None*: $t[46.3]=2.14$, $p<.05$, Cohen's $d=.09$). There was no effect of picture support on the productivity of CC pairs.

In support of H3, as indicated by effect sizes (Cohen's $d$), the effect of picture support on enhancing productivity was greater for intercultural groups than for intracultural groups. However, we did not detect differences across cultural groups for any picture support conditions. Intercultural groups had the greatest improvement when picture support was available, but still did not produce more ideas than intracultural groups.

### Breadth of Concepts

To test H4, H5, and H6, we used a linear mixed model with the breadth of concepts measure as the dependent variable. There was a moderate correlation between breadth of concepts and number of original ideas ($r=.27$). Although the correlation was not high, in order to ensure the results of concept breadth were independent of productivity, we included number of original ideas as a covariate in our model. Figure 5 shows the means of breadth of concepts estimated by the statistical model.

Picture support had a main effect on breadth of concepts ($F[2,45.9]=4.90$, $p<.01$). In support of H4, using the *Stimulus* method of picture selection resulted in broader concept coverage than using no pictures ($t[45.6]=2.47$, $p<.05$, Cohen's $d=.36$), and also broader than using the *Congruence* method ($t[46.6]=2.89$, $p<.01$, Cohen's $d=.62$). The *Congruence* method, in contrast, did not help to increase breadth of concepts (*Congruence* vs. *None*, $t[45.6]=.44$, *n.s.*).

H5 was not supported by the analysis. Group cultural composition did not have a main effect on breadth of concepts ($F[2, 23.8]=.54$, *n.s.*). AC pairs did not cover broader concepts than intracultural groups (AA or CC).

The interaction effect between group cultural composition and picture support was significant ($F[4,44.5]=4.92$, $p<.005$). Intercultural groups had broader concept coverage when using the *Stimulus* method than no support (AC & *Stimulus* vs. AC & *None*: $t[46.2]=3.94$, $p<.0005$, Cohen's $d=1.16$), and than using the *Congruence* method (AC & *Stimulus* vs. AC & *Congruence*: $t[46.2]=3.98$, $p<.0005$, Cohen's $d=1.65$). AA pairs, when using the *Stimulus*

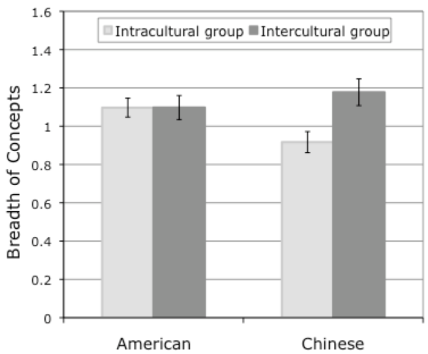


**Figure 6. When using the Stimulus picture selection method, breadth of concepts by individuals' cultural backgrounds and types of cultural groups in which they worked.**

method, had marginally broader concept coverage than no support (AA & *Stimulus* vs. AA & *None*: $t[44.8]=1.7$, $p<.1$, Cohen's $d=.48$), and broader coverage than using the *Congruence* method (AA & *Stimulus* vs. AA & *Congruence*: $t[46.5]=2.8$, $p<.01$, Cohen's $d=.98$). There was no difference in breadth of concepts for CC pairs across picture support conditions.

In support of H6, the *Stimulus* method had the greatest effect on broadening concept coverage when cultural diversity was available in the groups (i.e., intercultural group). What is especially noteworthy is the comparison between cultural groups. When the *Stimulus* method was used, AC pairs had greater breadth of concepts than both types of intracutural groups (AC & *Stimulus* vs. CC & *Stimulus*: $t[47.7]=3.28$, $p<.005$, Cohen's $d=.53$; AC & *Stimulus* vs. AA & *Stimulus*: $t[47.5]=1.66$, $p<.1$, Cohen's $d=1.26$). There was no difference between intercultural and intracultural groups for the other picture support conditions.

#### *Agency in Broadening Concept Coverage*

With the *Stimulus* method, intercultural collaboration resulted in greater breadth of concepts than intracultural work. This raises questions as to whether individuals with different cultural backgrounds working in the intercultural group *both* contributed to concept coverage. Did the *Stimulus* method influence individuals of both cultures?

To address this issue, we computed breadth of concepts at the individual level. The measure assessed the average semantic distance between any ideas generated by an individual on a brainstorming topic. To analyze, we used linear mixed modeling with individual-level breadth of concepts as the dependent variable. Individual cultural background (American or Chinese), types of cultural groups (intercultural or intracultural group), picture support, and interactions of the three variables were set as fixed effects. Brainstorming topic and number of original ideas generated by individuals were included as covariates. We first conducted an outlier analysis to exclude data with studentized residuals exceeding ±2 [15]. Ten out of 162 observations were excluded through this procedure. The revised linear model was reasonable in fit ($R^2=.72$).

Figure 6 shows how individual culture and group cultural composition interacted when the Stimulus method was used (mean breadth of concepts estimated by the linear model). Americans and Chinese had similar breadth of concepts when working in *Stimulus*-supported intercultural groups (AC pairs, the dark bars in Figure 6) ($F[1,100.9]<1$, *n.s.*). Under this specific condition, individuals from the two cultures proposed ideas with comparable breadth in concepts, and therefore, both appeared to take active agency to increase concept coverage.

As an interesting contrast, the *Stimulus* method did not appear to help Chinese individuals when they worked with other Chinese (i.e., CC pairs). As Figure 6 shows, the difference between Chinese working in intercultural groups vs. intracultural groups was significant ($F[1,92.1]=8.55$, $p<.005$, Cohen's $d=1.12$). Americans, on the other hand, did not change depending on the cultural background of their partners ($F[1,80.9]<1$, *n.s.*). Under the *Stimulus* condition, Chinese also appeared to be more adaptable to the cultural contexts in which they worked than Americans.

## DISCUSSION

In general, conversationally retrieved pictures emphasizing the quality of stimulation enhanced both the originality and the breadth of ideas generated. Pictures emphasizing contextual coherence supported productivity to a lesser extent than the *Stimulus* method, and did not facilitate breadth. Because the *Congruence* method selected pictures that were most related to the ongoing conversation, the pictures may not have been conceptually new or stimulating. Therefore, the failure to support the breadth of concepts may not be surprising. Overall, the general pattern of results with respect to picture support methods is consistent with the socio-cognitive view of brainstorming positing that conceptual diversity is crucial.

Another goal of our study was to understand how an external source of diversity, picture support, interacted with an internal source of diversity, cultural differences. Cultural diversity did not help brainstorming when there was no appropriate support. The *Congruence* method only helped AC pairs generate more ideas than the baseline of *None*, but did not make intercultural groups more productive or conceptually broader than intracultural groups. The *Stimulus* method, on the other hand, helped AC pairs cover broader concepts than both AA and CC pairs. It appears that this theoretically motivated design helped pairs leverage their cultural differences, thereby improving brainstorming outcomes.

### Role of Cultural Accommodation

One interesting observation is the similarity of performance patterns between AC and AA pairs (see Figures 4 and 5). Picture support appeared to have similar influences on individuals working in AC and AA pairs, such as enhanced productivity and breadth of concepts when brainstorming with the Stimulus picture support.

We suspect that Chinese participants' adaptation of communication behaviors may be responsible for the similarity between AC and AA pairs. Prior work on communication accommodation suggests that people may change their communication styles to accommodate partners from different cultures [13][30][32]. Studies have also shown that Chinese are more likely to adapt their behaviors to partners from another culture than are Americans [30][32]. This may explain the similarity of AC and AA pairs in the current study, as well as the differences between AC pairs and CC pairs.

The individual level analysis we conducted provides empirical support for the accommodation account by showing that Chinese individuals' brainstorming outcomes depended on the cultural backgrounds of their partners under the *Stimulus* condition. This suggests that studying the interactions between technologies and cultural factors such as accommodation merits future work.

### Second Language Use

CC pairs were not influenced by either type of picture support. Conversationally retrieved pictures did not enhance CC pairs' brainstorming outcomes in comparison to the baseline of showing no pictures. One possible account is the insufficiency of *verbalized* conceptual diversity, because the Chinese brainstormed in a second language (e.g., English). Because our mechanism of picture support requires verbal input to trigger, fluent expressions of rich concepts are crucial to picture retrieval and to sustain group brainstorming. Because both group members in CC pairs were using English as a second language, possibly not all ideas they thought about could be expressed fluently enough to trigger picture support.

As a future work, we are planning to integrate picture support and machine translation (MT) to enable group members of intercultural groups to speak their own native languages. Speaking in the native language may make it easier to express diverse and rare concepts, and thus may better foster verbalized conceptual diversity. By incorporating MT, we will study issues around language fluency, and evaluate whether the integration of pictures and MT may effectively support multilingual teamwork.

### Implications for Design

The combination of the two sources of diversity, cultural differences and conversationally retrieved pictures, speaks to two general design questions pertaining to culture and collaborative work: Whether the technology functions universally across cultures, and whether cultural differences may be used as a strategy to *support* certain work. These questions imply different stances with respect to the relationship between culture and technology.

The first question is essentially taking an evaluation stance, concerning whether the effects of a technology holds when moving to a different cultural context. Seeing cultural differences in technology use or task performance is typically interpreted as requiring specialization of design to ensure better culture-technology fit. Cultural differences are thus a target to be designed for, or around.

The second question takes the perspective that cultural differences are valuable resources that may become *part of* design. Cultural differences may introduce systematic diversity along many dimensions, such as language, social orientations, concepts, cognitive styles and life customs. In a group setting, interpersonal diversity may serve as a driving force to trigger positive group dynamics, such as promoting adaptation of behaviors so a desired effect can be attained (cf. the adaptation of Chinese in Figure 6), or increasing breadth of knowledge to attain more powerful collective intelligence (cf. the greater breadth of concepts covered by intercultural groups when receiving appropriate support in Figure 5). Cultural differences, in this view, become a design component, and may actually be the key to enabling certain technologies, such as enhancing group creativity through the combination of cultural differences and conversationally retrieved pictures.

We consider both views concerning the roles of cultural differences in design valuable. It is important for CSCW to design for cultural differences and to make domain-general intercultural collaboration easier to manage. It is also useful to consider the *utility* of cultural differences and incorporate it in design, such as naturally and systematically introducing diversity and dissent [19] that stimulate thinking and reflection beyond what a homogeneous cultural context can afford.

## CONCLUSION

In this study, we examined the effects of two sources of conceptual diversity, cultural differences and conversationally retrieved pictures, in collaborative brainstorming. We confirmed that showing pictures selected by emphasizing their stimulating utility, such as containing multiple or rare topics, helped to enhance brainstorming productivity and concept coverage. Multicultural composition as an internal source of diversity required using the stimulation-emphasis picture selection method as support to convert hidden cultural diversity to breadth of concepts, a valuable brainstorming outcome. We presented a view that, with appropriate support, cultural differences may introduce beneficial diversity that enhances task outcomes. It would be worthwhile to study the phenomenon under real-world distributed teamwork conditions, in which people are engaged in more meaningful work goals and situated in their own cultural contexts.

## ACKNOWLEDGEMENTS

We thank Meaghan Frank and Leslie Setlock for their assistance. The material is based on work supported by the National Science Foundation under Grants IIS-0803482 and IIS-0845351. Photos shown are licensed under Creative Commons.

## REFERENCES


1. Amabile, T. (1996). *Creativity in context*. Boulder, CO: Westview.
2. Anderson, J. R. (1983). A spreading activation theory of memory. *J. Verbal Learning and Verbal Behavior*, *22*, 261-293.
3. Brown, V., Tumeo, M., Larey, T. S., & Paulus, P. B. (1998). Modeling cognitive interactions during group brainstorming. *Small Group Research*, *29*, 495-526.
4. Cepeda, N. J. (2008). Effect Size Calculator. http://www.cognitiveflexibility.org/effectsize/.
5. Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proc. Nat'l Acad. Sci.*, *102*, 12629- 12633.
6. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Ed.. NJ: Erlbaum.
7. Diamant, E. I., Fussell, S. R. & Lo, F-L. (2008). "Where did we turn wrong?" Unpacking the effects of culture and technology on attributions of team performance. *Proc. of CSCW 2008* (pp. 383-391). NY: ACM Press.
8. Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: toward the solution of a riddle. *J. Personality and Social Psychology*, *53*, 497-509.
9. Dugosh, K. L., Paulus, P. B., Roland, E. J., & Yang, H-C. (2000). Cognitive stimulation in brainstorming. *J. Personality and Social Psychology*, *79*, 722-735.
10. Fischer, G. (2005). Distance and diversity: Sources for social creativity. *Proc. Creativity and Cognition.*
11. Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes*, *25*, 285-307.
12. Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Nunamaker, J. F. (1992). Electronic brainstorming and group size. *The Academy of Management Journal*, *35*, 350-369.
13. Giles, H., Coupland, J. & Coupland, N. (1991). *Contexts of Accommodation*. NY: Cambridge University Press.
14. Kenny, D. A., Mannetti, L., Pierro, A., & Livi, S. (2002). The statistical analysis of data from small groups. *J. Personality and Social Psych.*, *83*, 126-137.
15. Littell, R., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. SAS Inst.
16. Markus, H. R. & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224-253.
17. Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105-125.
18. Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida Word Association, Rhyme and Fragment Norms. http://w3.usf.edu/FreeAssociation.
19. Nemeth, C. J., & Oriston, M. (2007). Creative idea generation: Harmony versus stimulation. *European Journal of Social Psychology*, *37*, 524-535.
20. Nijstad, B. A., Stroebe, W., & Lodewijkx, F. M. (2002). Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology*, *38*, 535-544.
21. Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: a cognitive model of idea generation in groups. *Pers. and Social Psych. Review*, *10*, 186-213.
22. Nisbett, R. E. & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences*, *100*, 11163-11170.
23. Osborn, A. (1957). *Applied Imagination*. NY: Scribner.
24. Paulus, P. B., & Brown, V. R. (2007). Toward more creative and innovative group idea generation: a cognitive-social-motivational perspective of brain-storming. *Soc. Personality Psych .Compass*, *1*, 248-265.
25. Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*, 741-754.
26. Raaijmakers, J.G.W., & Shiffrin, R.M. (1981). Search of associative memory. *Psych. Review*, *88*, 93-134.
27. Setlock, L D., Fussell, S. R., & Neuwirth. C. (2004). Taking it out of context: collaborating within and across cultures in face-to-face settings and via instant messaging. *Proc. CSCW 2004* (pp. 604-613).
28. Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications: Festschrift in honor of Lyle Bourne, Walter Kintsch, and Thomas Landauer* (pp. 237-249).
29. Triandis, H. C. (1995). *Individualism and Collectivism*. Boulder, CO: Westview.
30. Wang, H-C., Fussell, S. R., & Setlock, L. D. (2009). Cultural difference and adaptation of communication styles in computer-mediated group brainstorming. *Proc. CHI 2009* (pp. 669-678). NY: ACM Press.
31. Wang, H-C., Cosley, D., & Fussell, S. R. (2010). Idea Expander: Supporting group brainstorming with conversationally triggered visual thinking stimuli. *Proc. CSCW 2010* (pp. 103-106). NY: ACM Press.
32. Wang, H-C., & Fussell, S. R. (2010). Groups in groups: Conversational similarity in online multicultural multi-party brainstorming. *Proc. CSCW 2010* (pp. 351-360).