Building Better Design Teams: Enhancing Group Affinity to Aid Collaborative Design

This paper discusses ConvoCons, a novel system of conversational icons intended to encourage affinity between collaborators unobtrusively. Using a reification of Bonnie Nardi's framework for social connection and affinity, ConvoCons overlay an existing application and display varying media that can encourage collaborating partners to begin developing affinity through informal conversations. This research explores whether dyads working on a collaborative multitouch application with ConvoCons develop more affinity than dyads that do not while solving simple design problems and a freeform design task. Results indicate that after an average of 23.25 minutes, affinity, defined as a function of conversational and behavioral cues, was 40% higher ($p < 0.001$) in the ConvoCons group than in the control group. This research offers a framework for evaluating affinity within groups and a foundation for exploring software-based methods of improving the effectiveness of collaboration within design teams.

Introduction

Imagine a new employee joins a team of designers to create an interface for an application. The design team typically works in pairs, and her partner for this project is somebody she has never met. The designers share a multitouch device to place the GUI components and create the interface of their product. There is some awkwardness and formality as they work. Having never worked together, they are strangers trying to create a shared vision of a design. Suddenly, a message pops up on the screen; the new employee reads it—some sort of a question. She looks across the screen and her partner has one as well, so she reads hers aloud. The other em-

Design Computing and Cognition DCC'10. J.S Gero (ed), pp. xx-yy.

ployee reads hers, and they realize it is a riddle. They begin talking about the riddle and about their past work experience. They continue to work, and when a work-related question arises about a design idea, the partners now have no problem asking each other questions; the awkwardness has been removed.

The story above illustrates the social awkwardness that can occur when working with a new partner for the first time. Typically one has little idea of what to expect from a partner and unless there has been an introduction before, a person may find him- or herself hesitant to start a conversation as no affinity has been developed. When individuals work together for the first time they lack knowledge of one another's reputations and other relational elements typically useful for successful cooperation [3]. Strangers cooperating for the first time without a shared connection to facilitate introductions and establish common ground may at first struggle to establish a level of affinity needed for productive cooperation [8][23]. Individuals seek affinity as a means to fill a need for interpersonal relationships and established affinity is necessary for sustained cooperative relationships [15][34].

This research describes the evaluation of a user interface technique developed to more quickly build affinity and effective collaboration strategies among strangers by promoting incidental conversations. This system of conversation-starting icons, called ConvoCons, offers conversation starters to encourage an informal discourse between new partners that Nardi identified as a central component of group affinity [23]. An example screenshot of an application with overlaid ConvoCons can be seen in Figure 1; in this case, the ConvoCons are text within circles that are oriented towards two partners facing each other and collaborating on a tangram puzzle.

ConvoCons are part of an ongoing research effort to explore means of using interfaces to promote constructive collaborative strategies among groups of individuals using computers to facilitate their work, with a particular emphasis on collaboration involving creativity and design [24].

In this paper, we provide an analysis of ConvoCons used in a multitouch tangram application and their effectiveness in encouraging affinity between dyads, a pair of individuals treated as one unit. Tangrams is a Chinese puzzle game consisting of seven geometric shapes (five triangles, a square, and a parallelogram) that is used to create a variety of shapes both freeform (creative) as well as in a pre-specified pattern (problem solving). It is important to note that while ConvoCons appear concurrently with the tangram puzzles, they are semi-transparent and serve as passive interface elements (they do not recognize user input), thus users are free to attend to

or ignore the ConvoCons without adversely affecting their ability to get work done.

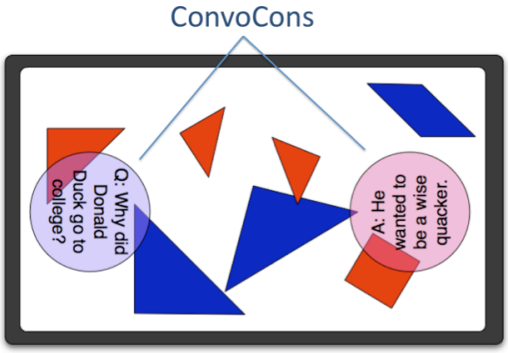

Fig. 1 A sample joke ConvoCon: the participant on the left would have a privileged view of the question while the one on the right would have a privileged view of the answer.

The first research question (Q1) is "Does the presence of ConvoCons lead to increased incidental conversations?" In order to answer this question, we defined incidental conversation to be dialogue unrelated to the tangram task at hand and looked at the amount of such dialogue between dyad members that worked on tangrams with ConvoCons vs. without Convocons. Our secondary research question, (Q2) is, "Do ConvoCons lead to increased affinity between participants?" For this research question, we operationalized a definition of affinity based on two components: conversational affinity and behavioral affinity. Total affinity is based on a percentage of the interactions that demonstrated affinity vs. those that did not.

Background & Context

The system and research model for ConvoCons are an applied reification of Nardi's observations that affinity plays a central role in the process of creating and sustaining connections necessary for productive collaboration. While the majority of the dimensions of affinity that Nardi observed were physical in nature, we chose to focus on the aspect of incidental communication, i.e., conversations outside of productive work such as commenting on the weather. Nardi suggests that this informal discourse leads to connections that aid in collaboration critical to productive collaborative strategies [23]. We hypothesize that these affinity bonds, potentially promoted through discussing ConvoCons, lead to the critically important state of so-

cial cohesion [19]. By conducting this experiment with two people sharing a small (15.4") multitouch device, two of the other activities that promote affinity are added to the work context: human touch (the occasional brush of the hand) and a shared experience in a common space (where the common space is both physical and virtual) [23].

The importance of affinity for effective collaboration can also be seen in Schmid's discussion of affinity as the cornerstone to the development and use of social capital [30]. Specifically, ConvoCons seek to improve what Schmid refers to as positive affinity, which connects individuals to one another through a build up of social capital and, as a result, reduces the free rider problem. While this study does not examine the effect of affinity over time (it assumes the validity of Schmid's theory), the ConvoCon system is designed to accelerate this accumulation of affinity and decrease the time needed to build social capital. Kellogg and Erickson [18] suggest that social translucence, the idea that user activity needs to be apparent to other users, is a key to effective collaboration. ConvoCons are designed to increase social translucence in that by building affinity between partners, group members will be able to understand their partners' cues better in order to collaborate through turn taking and directing work with their partners. In addition, Convertino et al [8] suggest that in order for group members to successfully collaborate, they must develop converging measures, which is the idea that they have a common ground or shared representation of the task. Both of these concepts are components of the third dimension of affinity identified by Nardi, that of a shared experience within a shared space [23].

However, the creation of common ground has been shown to cause problems within groups when individuals focus on the elements they share and never move beyond that to share expert knowledge needed to solve a problem. Larson described this tendency in his study of doctors working on collaborative diagnosis where each doctor had been shown a different piece of the medical problem and a successful diagnosis could only occur when information was shared [21]. Analogously, earlier ConvoCons prototypes explored the use of a centralized, shared conversation starter, but designs that provided each participant with separate, privileged information (**Fig. 1**) were observed to be more likely to prompt participant conversation. Participants asked about one another's pieces of the question and answer [24].

The ConvoCons Approach

ConvoCons were developed based on an initial collaboration study that suggested that the ambiguity present in a somewhat confusing user inter-

face served as a means of creating converging measures through users' discussion of the unusual interface. However, given the cost of encouraging poor interface design to promote collaboration, the ConvoCon system was developed to serve the same role without the cost to general usability of the system [24]. Related research by Clear and Daniels has explored the use of icebreakers to encourage better collaboration techniques between distant learners [6]. In addition, Fisher and Tucker have used online games as a means of providing an out-of-classroom means for online students to gain affinity with one another [10]. However, unlike these previous approaches, ConvoCons are built into the tasks, and collaborators are free to attend to them or not; there is no structured ice-breaking mechanism or time required outside of the task.

Also, while ConvoCons are designed to encourage incidental conversations and affinity, the goal is not simply to connect people, but rather to encourage stronger collaborative working behavior. In our research we have often observed members of dyads who, although working on the same problem within the same virtual and physical space (on the same device), failed to acknowledge or utilize their partners, instead tackling the problems separately, avoiding interaction with their partners when possible. Previous research by Rogers looked into the use of shared displays to serve as an icebreaker to promote and track conversations within a social setting; in this research we use the shared display as a work tool (rather than an icebreaker) although Rogers' and our own work share the goal of using technology to bring people closer together [28]. Finally, while the underlying goal of Karahalios' social catalysts is similar [17], in calling for designers to consider interfaces as a means of promoting social connections, her work focuses primarily on aiding individuals in finding collaborators. In contrast, our work assumes that group members are already paired by a work assignment. ConvoCons are intended to ease new partners' transition into working with higher affinity on a shared goal.

ConvoCons System Architecture

The ConvoCons system is designed to be overlaid on any Java application and can be used with several simultaneous client applications, e.g. two or more people using ConvoCon-enabled applications at different sites. By allowing this level of adaptability of ConvoCons, we are able to test a variety of configurations to test their effectiveness in encouraging collaboration within a wide range of applications and environments. The touch interface is not required for ConvoCons; it is helpful in this context for affording an easy approach to simultaneous use of an application by two co-located users. The touch-based gesture recognition system is built using

Sparsh-UI, an open-source API for platform independent touch-based applications [25,32]. While this study used visual ConvoCons containing text, the ConvoCon architecture supports media including auditory signals, videos, and images.

Methods

Thirty-six participants were recruited from the psychology department participant pool and paired into 18 dyads (the dyad will be the unit of analysis). Each dyad was then randomly assigned to be in either the experimental group (ConvoCon enabled tangrams; n=9) or control group (plain tangrams, n=9). Participants in all dyads had no previous relationship beyond "seeing each other around" (n=1 dyad) although the majority (n=17) had never met before. Participants were instructed to arrive and wait at different entrances to the research lab to prevent interaction before the start of the study.

Dyads were instructed to sit across the table from one another to allow co-located collaboration with the multitouch device, a Stantum SMK 15.4, placed length-wise between them (see **Fig. 2**) [4,29,31]. All dyads then followed the procedure below (see **Fig. 3**). Dyads were first given a brief description of the technology and told they would have five minutes to play with the interface and teach themselves how to use it. After dyads completed the five minutes of playtime, dyads were then given the first pattern to create with the tangram pieces (Task 1). Dyads worked on the pattern until completion and then were given the next pattern for a total of three patterns (Tasks 1-3). Upon completion of the patterns, the dyads were then given up to five minutes to create any new pattern of their choice (Free-form activity). All interactions were video recorded, and the software logged user inputs.

The dyads assigned to the ConvoCon group were exposed to ConvoCon riddles and jokes upon the first touch of the multitouch interface. One participant was given a privileged view of the riddle while the other was given a privileged view of the answer (see Figure 2). Each ConvoCon remained visible for thirty seconds, followed one minute later by another ConvoCon. ConvoCons did not affect the interaction with the tangrams application; they did not block access to the application, nor could users control them. There were a total of ten ConvoCons displayed to dyads over a 15-minute time period (see Figure 3). The ConvoCon group had a mean completion time of 23.25 minutes (SD=7.00) for the entire interaction: Play, Tasks 1-3, and Freeform. This timing resulted in most participants completing part

of Task 2 as well as all of Task 3 and the freeform task without ConvoCons. Turning ConvoCons off midway was intended to allow the researchers to observe whether or not the effects of ConvoCons would be sustained throughout the working session.

Fig. 2 The multitouch device between two participants.

Finally, we administered an exit survey to all participants based on a modified version of the survey Convertino developed to assess the similar concept of common ground development [7]. Using this survey, consisting primarily of five-point Likert scale questions, we compared the control and experimental groups to determine how well the respective members felt their dyad worked together and the agreement that was reached within the dyad. Dyads in the ConvoCon group also participated in a brief, unstructured interview in order to obtain their input on the ConvoCons experience to obtain formative feedback that would aid in refining the system.

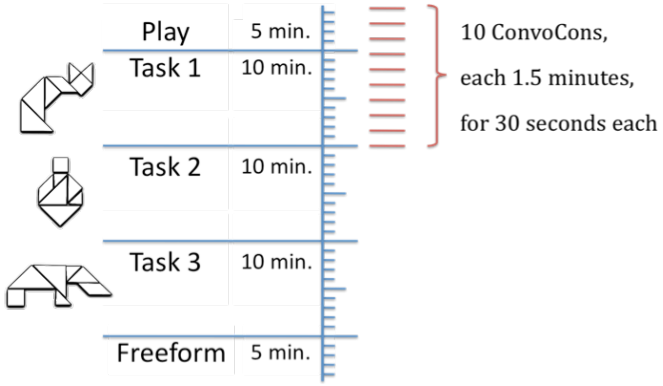

Fig. 3 Sample timeline of the procedure (task times varied based on completion time), the 10 minutes per task is just an example, some took less time and some took more.

Our choice of using visual versions of ConvoCons was taken due to research indicating that visual background noise has less of an adverse effect on performance compared to auditory background noise [9]. This was important since ConvoCons are not directly related to the task at hand in this experiment, they may be viewed as background noise at times of particularly difficult work. The post-task survey and interviews were designed to evaluate that issue. The choice of using text-based ConvoCons for the initial study was made because reading text requires some cognitive load and, with a partner present, is often spoken allowing us to determine whether participants were attending to the information more readily than if we had used simple images that require less cognitive processing and are not naturally spoken.

Through an iterative process we discussed in an earlier paper, we settled on riddles and jokes rather than news headlines, trivia, or facts about tangram puzzles. Participants lacked the contextual information to discuss news headlines and trivia, while facts about tangram puzzles afforded little discussion [24]. The jokes (obtained from children's collections) were chosen as a way of lightening the mood, which Goffman's study of role distances has shown to be an effective means of initiating new members into a group and allowing senior members a break from the stress of the roles they play [11]. Riddles alternated with jokes in the same order for all ConvoCon dyads. There was no observable difference between dyads' attentiveness to riddle or joke-based ConvoCons, although in informal observation, the joke-based ConvoCons did appear to be more effective in creating affinity bonds within a dyad, particularly in stimulating discussion after the disappearance of the ConvoCon The riddle-based ConvoCons often provoked more discussion while they were still present, however, as participants sometimes tried to solve them before reading the answer or would comment on how the answer "made sense."

Framework for Measuring Affinity

To measure whether the ConvoCons system increases affinity, a measurable definition is required. Nardi [23] defines affinity as a "feeling of connection between people." The issue of empirically measuring affinity is similar to the problem Goudy observed with "rapport" where there are multiple, sometimes conflicting, definitions and limited clearly defined metrics for measurement [12]. With this problem in mind, we adapted Nardi's definition and framework and narrowed it in the context of our multitouch environment to the "convergence of thoughts, actions, or ideas" and made the following operationalized assumptions for measurement purposes within a multitouch collaborative context.

Seven tag categories were established partially based on observations of over forty dyads performing collaborative work on multitouch devices, and partially based on common notions of affinity, such as socially appropriate conversational distance [13,27].

Peshikin has suggested that qualitative methods are the most useful means of observing social interactions [26]. To quantify the affinity that we observed within dyads, we used an approach based on Anfara et al. in their discussion of making qualitative data gathering techniques transparent [1]. Videos of dyads were tagged with codes according to the behavior observed. The coding tags were derived from the below seven categories of affinity (see **Table 1**).

Coding for Affinity

Using the video recordings of the hand movements and voices of dyads, we divided the videos up between each task (Play, Task 1, Task 2, Task 3, and Freeform). The researchers then classified each five-second block of the video based on two overall constructs: the type of **behavior** (9 codes) and type of **conversation** (16 different codes). See Table 1. For conversations, the codes were then grouped into four larger categories: ConvoCon-related (e.g. reading ConvoCon text to each other, discussing ConvoCons, trying to solve ConvoCon riddles, etc.), *ConvoCon-indirect* (laughter within 1 minute of ConvoCon appearance and non-work talk within 1 minute of ConvoCon appearance), *non-ConvoCon affinity* (talking about year in school, major, directing partner since giving and taking direction requires a level of comfort one another, etc.), and *low/no affinity conversations about task* (e.g. getting unstuck, teaching their partner how to perform a system action, etc.) These four conversation categories were further grouped into affinity-related and low/no affinity. Participants' behaviors were coded as *affinity-related* (e.g. close proximity of hands, turn taking where one places and the other adjusts, etc.) or *low/no affinity* (e.g. hand avoidance, independent work where one partner is working on one section of the pattern while the other is working on another without any shared vision).

Each five-second block of video received one tag related to dyad behavior and one related to dyad conversation. For each task, three proportional affinity scores were calculated: overall affinity, affinity conversation, and behavioral affinity. The overall affinity score is equal to all blocks with affinity codes divided by all blocks that exhibited some conversation or behavior. Affinity conversation equals the number of blocks with conversation affinity divided by the number of total blocks whether or not conversation or behavior was present. Similarly, affinity work equals the

number of blocks with behavioral affinity divided by the number of total blocks. See Table 2 for a sample calculation. The affinity scores were then compared for each task between the experimental and control group through a Student's T-Test.

Table 1: The tags used to code the videos

- Types of Conversation
  - Affinity – Directly Tied to ConvoCons
    - Riddle Solving
    - Both Reading
    - Laughing (ConvoCon)
    - Talking (ConvoCon)
  - Affinity – Indirectly Tied to ConvoCons
    - Talking (within 1 min. of ConvoCon)
    - Laughing (within 1min. of ConvoCon)
  - Affinity – non-ConvoCon (not tied to ConvoCons)
    - Playful Conversation
    - Conversation About Partner
    - Planning Solution (not fixing)
    - Discussing Freeform
    - Directing Partner
    - Affirmation, gratitude, etc.
  - Low/no Affinity
    - Getting 'unstuck'
    - Teaching
    - Other Talking
    - Work related (w/i 1 min. of ConvoCon)
- Types of Behavior
  - Low/no Affinity
    - Independent
    - Turn Taking (independent)
    - Avoidance (hands)
    - Grabbing (taking pieces from other's 'personal space')
  - Affinity Related
    - Turn Taking (one places, other adjusts)
    - Directing-Following
    - Close Proximity (hands)
    - Shared Plan
    - Building--adding on to other's (Creative/freeform only)
- No Talking or Action

A total of 5,149 blocks were given conversational and behavioral codes by a single coder, one of the researchers. In order to ensure the coding method was valid, two dyad videos (a total of 674 blocks) were randomly selected and the category codes for behavior and conversation were compared between the researcher and a second coder using percent agreement and Cohen's Kappa (calculated in SPSS version 18). The second coder was an undergraduate who was not informed about the purpose of the experiment. She was trained for approximately an hour. The second coder was also provided with a 1-3 sentence description of each code but was not provided a specific video example of the code. She then completed one practice video and received feedback from the researcher after each task had been coded, which were only checked to ensure that she understood the process—particularly that each block should have only one conversation and one behavioral code. After completing the practice video she then tagged the two videos used for the calculation of interrater reliability. For the behavior category codes across both videos, there was a 90.0% agreement between coders with a Cohen's Kappa of 0.612. For the conversational category codes across both videos there was a 90.7% agreement with a Cohen's Kappa of *k*-0.708. Both of these Kappa scores fall into the range of scores that Landis and Koch referred to as "substantial agreement" [20].

Table 2. To illustrate the calculation of three proportional affinity scores based on tag codes of 5-second video blocks: for a 4-block task (20 seconds), affinity conversation is 25% (1=affinity, 0=low/no affinity, blank=no talking). Affinity work (behavioral) is 75%. The overall affinity is 67% (all affinity blocks / all blocks that exhibited some conversation or behavior).

Conversation	1			0
Behavior	0	1	1	1
	5 sec	10 sec	20 sec	25 sec

Results

As this work was an initial study intended to verify the feasibility of ConvoCons as an interface technique to encourage increased affinity in collaborative work, we have only analyzed the data using basic statistical methods and have not identified any other variables of interest at this time. In reading these graphs it should be noted that ConvoCons typically stopped appearing between the end of task 1 and the middle of task 2. In addition, the puzzles used for each task, presented in a consistent order to

all participants, were intended to progress from simplest to solve to hardest to solve. All Student T-Tests were conducted with an α=0.05.

Exit Survey

The control group (n=18 participants; 9 dyads) had a mean age of 20 (SD=2.09) with 12 males and 6 females. All but one participant in the control group indicated that they had used a multitouch device (such as an iPhone) at least once with a mean response on a 5-point Likert scale of 3.0 ("A Few Hours") and a median response of 2 ("Tried it Once). One participated self-reported as "Life of a Party" on a 5-point Likert scale of sociability, with a mean rating of 2.89 and a median score of 2 ("Prefer tight groups").

The experimental group (n=18 participants; 9 dyads) had a mean age of 21 (SD=3.21) with 9 males and 9 females. Five participants in the experimental group indicated they had never used any form of multitouch device, with a mean score of 2.53 and a median score of 1 ("None") and 3 ("A Few Hours"). There were two medians for multitouch use because both categories had an equal number of responses. One participated self-reported as "Life of a Party" on a 5-point Likert scale of sociability, with a mean rating of 3.00 and a median score of 2 ("Prefer tight groups").

Analysis of survey results was conducted both through an analysis of individuals within each group as well as by grouping data into dyads where agreement at an appropriate level was scored a "1" and disagreement with one partner providing a score of "neutral" or lower scored as "0". There were no statistically significant differences between groups on the questions intended to assess participant's feelings of affinity toward their partners.

Completion Time – Log Data

Including the play time and the freeform task, there was no significant difference between the experimental group (mean=23.25 minutes, SD=7) and the control group (mean=23.25 minutes, SD=6.5). While groups were told they had five minutes to "play" with the system and learn the controls, the ConvoCon group was more likely to utilize the full play time while the control group often reached a point where both partners would awkwardly stare at their feet, the screen, or away from each other before asking to move on to the puzzle, resulting in a statistically significant difference between the groups (p=0.008) time taking during the Play task. Since one concern we had was that ConvoCons and the incidental conversations might distract groups from the work at hand, we calculated the mean completion time just for the three puzzles to look at just the effects of Convo-

Cons on work efficiency. There was no significant difference in time spent on the three puzzle tasks between the experimental group (mean=14.75 minutes; SD=4.75) and the control group (mean=16.25 minutes, SD=6.5).

Quantitative Evaluation of Video Data

Q1 of this study, whether ConvoCons produce more incidental conversations, is addressed by our score for conversational affinity with the means and standard deviations seen in Table 3. It should be noted that for frequency of incidental conversations, we did not count all conversational labels that we classified as signs of affinity—only the tags that were not related to work were counted (e.g. "playful conversations" and "talking about partner"). There was a significant difference between the frequency of incidental conversations between groups for the playtime, task 1, task 2, and task 3 (p=0.001, p=0.002, p=0.01, and p=0.021 respectively). However, there was not a significant difference between groups in the frequency of incidental conversations during the freeform task (p=0.11). Overall, there was a significant difference between groups (p<0.001) thus supporting the idea that ConvoCons increase the frequency of incidental conversations.

Table 3 Means and standard deviations of incidental conversations over all tasks indicate a higher frequency in the experimental (ConvoCons) group. Asterisks denote significant differences.

	Play*	**Task 1***	**Task 2***	**Task 3***	**Freeform**	**Overall***
ConvoCons	15.78	8.22	5.44	7.67	2.67	7.96
SD	7.68	6.96	4.33	9.82	3.00	2.64
Control	3.11	0.44	1.22	0.33	1.22	1.27
SD	7.20	0.88	2.39	0.71	1.72	2.35

Table 4 corresponds directly with a portion of Q2 of this study: whether or not the use of ConvoCons leads to increased affinity. As expected from the literature on ice breakers that providing a shared framework for conversation allowed participants to begin incidental conversations at an early stage, ConvoCons resulted in a 20% increase in conversational affinity during the Play task which was statistically significant at p=0.006. Furthermore, a statistically significant difference between the experimental and control groups were maintained throughout the duration of the study (task 1, p=0.025; task 2, p<0.001; task 3, p=0.004; and freeform, p<0.001). While the researchers expected a significant increase in conversational affinity for the freeform task in both groups because of the task's unstructured context, the control group only saw a 4% increase compared to the experimental groups 26% increase in conversational affinity. This differ-

ence came from the control group discussing the freeform task less, with conversation centering on the general shape that would be made and very little planning and coordination of the task compared with the experimental group. In fact, it was not uncommon to observe one individual in the control group take control of the freeform pattern rather than sharing the work and design with his or her partner. Taking the mean across all tasks (not seen in the graph), there was a significant difference ($p<0.001$) between the experimental group with a mean of 32.4% affinity (SD=10.2%) and the control group with a mean of 10.6% affinity (SD=9.3%).

Table 4 ConvoCons serve as an early conversation starter for groups and the conversational affinity increases steadily. Circles are the ConvoCon group and squares represent the control group.

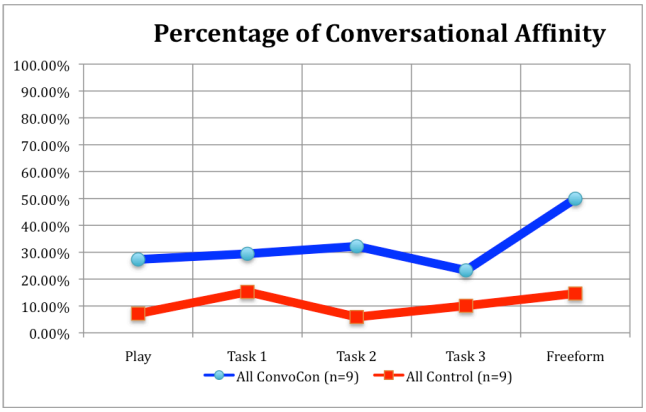

Table 5 also corresponds to Q2, whether or not the use of ConvoCons leads to increased affinity. To the researchers, this is the more important question since our ultimate goal for ConvoCons is to facilitate individuals' working with one another in a collaborative manner. As expected with groups that are working together for the first time, the level of behavioral affinity for both the experimental and control groups starts out with a non-significant difference (p=0.456). Once the first pattern is given and the individuals start trying to complete a shared puzzle, the proportion of behavioral affinity goes up for both the control and experimental group, although the increase for the experimental group is larger with a marginally significant difference between groups (p=0.069). Task 2 and task 3 see a minor, but steady increase in the percentage of behavioral affinity for the experimental group with a statistically significant difference compared to the control group (p=0.024 and p=0.013, respectively). In the final, freeform task the researchers expected both control and experimental groups to have a rapid increase in behavioral affinity as they work to realize a shared vi-

sion for a new pattern. However, while both groups did see a jump in behavioral affinity, the experimental group saw a much larger increase from 29.6% in task 3 to 60.6% compared with a 10.6% to 21.8% increase for the control group. This difference between the behavioral affinity and the control group for the freeform task was significant (p=0.002). Taking the mean score across all tasks results in a significant difference (p=0.004) between the experimental group with a mean of 30.6% affinity (SD=16.9%) and the control group with a mean of 12.3% affinity (SD=5.9%).

Table 5 Initially no difference in behavioral affinity exists; however, when the dyads begin working on the tasks the ConvoCon group shows increased behavioral affinity resulting in over 40% more affinity in the Freeform task.

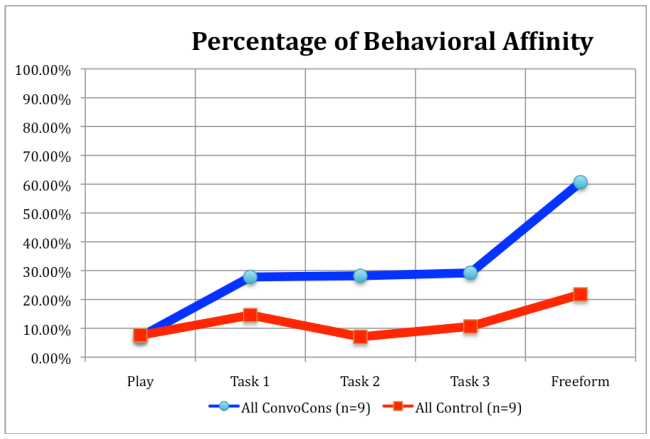

Table 6 The experimental group starts out with just under 20% higher affinity, due to the increased conversations occurring due to ConvoCons. Although a slight decrease occurs in the ConvoCon group after ConvoCons stop appearing in task 3, the difference between groups is 40% in the Freeform task.

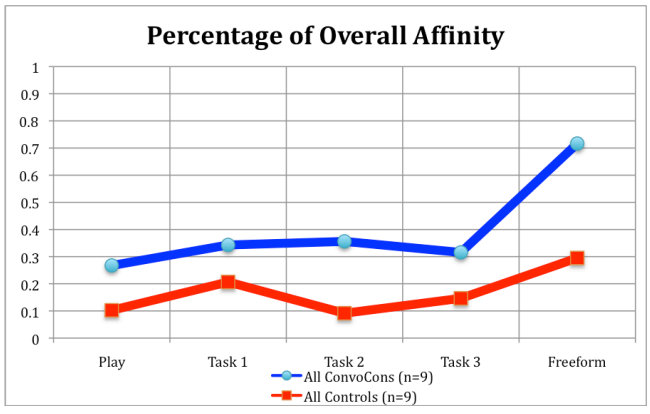

Table 6 displays the overall affinity across all tasks. Playtime, Task 2, task 3, and freeform all had statistically significant differences between the experimental and control groups (p=0.018, p=0.001, p=.002, and p<0.001 respectively). However, the difference between groups in task 1 was only marginally significant (p=0.068). The mean across all tasks was statistically significant (p<0.001) with the experimental group having a mean of 40.3% affinity (SD=10.9%) compared to a mean of 16.8% affinity (SD=7.8%) for the control group.

Exit Interviews

During the unstructured interviews of the nine ConvoCon dyads, four dyads indicated no signs of affinity here qualitatively assessed as agreement and the extent of building off one another's answers—this is different from the affinity coding used above) as they answered questions. Out of the four dyads that exhibited no signs of affinity, two of them exhibited a disconnect regarding their feelings toward ConvoCons where one person found them interesting/funny and the other had no opinion. Three of the nine dyads thought the ConvoCons were irritating or distracting and all three of these dyads expressed affinity or high affinity during the interviews. Only one of nine dyads had both members that enjoyed the ConvoCons and this dyad showed signs of high affinity.

All dyads, regardless of their feelings toward ConvoCons, indicated that after some time they began ignoring the ConvoCons and focusing more on the tasks. This was expected and part of our reasoning behind stopping ConvoCons after a period of time had elapsed. Further studies may indicate what the ideal threshold is for displaying ConvoCons long enough to facilitate a dyad's conversation but stopping the ConvoCons before the dyads decide to ignore them.

Four of the nine dyads indicated that they felt the ConvoCons were somehow related to the task. Of these four dyads, two showed no signs of affinity within the interview and two of them showed some signs of mild affinity. In addition, five dyads expressed feelings of pressure to complete the puzzles as part of the reason they began ignoring the ConvoCons—this was despite the fact that the groups were told that completion time was unimportant and that they would be given as much time as they needed or desired to complete the puzzles.

Finally, three of the nine dyads specifically mentioned that they felt the ConvoCons had an influence in getting them to begin conversations. However, one of these three dyads indicated no affinity during the interview.

Discussion

The 9% drop in conversational affinity for the experimental group from task 2 and 3 (mirrored in the overall affinity) is likely due to ConvoCons no longer appearing after task 2. Despite this drop, the freeform task saw the experimental group increasing conversational affinity to 45% higher than that of the control group.

One thing we noted was the apparent unreliability of the survey data—while both video coders saw groups where a single individual performed almost all work, both participants reported in the survey that they worked equally and the final product equally represented one another's goals. This result was surprising because previous studies that have examined similar concepts of rapport and development of common ground have relied almost exclusively on survey data [5,8]. This disparity between the survey data and the empirical observations suggests a need to explore new methods for assessing group work that has been echoed by other researchers who study improving group work [2].

Given the power of authority and the tendency to conform to assigned roles demonstrated by Milgram and Zimbardo [22,35], participants in early pilot studies in which the experimenter conducted a brief explicit training session may have been strongly focused on the puzzle tasks by 1) hearing our experimenter conduct training on the tasks. The unstructured playtime activity in this research was designed to lessen these influences. However, participants may have still been unrealistically focused on the tasks knowing that they would receive a departmental research credit for participating in the study.

The lack of difference in completion time between groups provides support that ConvoCons do not increase the time dyads require to complete work even though they produce more incidental conversations. These results are encouraging because they indicate that it is possible to use a software interface to increase conversational and behavioral affinity without adversely affecting the efficiency of work.

Conclusion

To return to our initial research question of "Does the presence of ConvoCons lead to increased incidental conversations," the data presented in this study suggests that they do, in fact, promote incidental conversations. Furthermore, the increase in incidental conversations does not appear to come at the cost of efficiency as measured by completion time.

At this point, our research focuses on affinity creation and does not look at the length of affinity bonds created nor does it explore whether or not affinity creation through our system promotes cooperation in a competitive environment; it simply seeks to explore a low-cost method of promoting affinity within a co-located dyad where neither partner has previous knowledge of the other. Future research will seek to answer these larger issues, but in this research we are seeking to establish a foundational framework for the design of interfaces to encourage specific collaborative behavior within working groups.

Regarding the secondary research question, "Do ConvoCons lead to increased affinity between participants," based on this study it appears that the incidental conversations promoted by ConvoCons are effective in producing a greater level of behavioral affinity, reifying Nardi and Whittaker's framework for affinity as a central element to collaboration. These results may also suggest that Schmid's theory of the role of affinity within the buildup of social capital for the reduction of free riders and increased motivation may be realized through ConvoCons, although further studies would be needed to explore free riders within larger groups rather than simple dyads. This effect may be further enhanced through the use of privileged, as opposed to shared, information within the ConvoCon display.

Acknowledgements

We thank Prasad Ramanahally and Jay Roltgen for their assistance in creating the drivers and software used for this study. In addition, we thank Dr. Joanne Marshall for her advice and help in exploring qualitative data. This research was performed with support provided by the U.S. Air Force Research Lab.

References

1. Anfara, V., Brown, K., and Mangione, T. (2002). Qualitative Analysis on Stage: Making the Research Process More Public. Educational Researcher, 31(7), pp. 28-38.
2. Bolia, R., Nelson, W., Summers, S., Arnold, R., Atkinson, J., Taylor, R., Cottrell, R., & Crooks, C. (2006). Collaborative Decision Making in Network-Centric Military Operations. Proc. of the Human Factors and Ergonomics Society.

3. Bolton, G., Katok, E., & Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. Journal of Public Economics, 89, 1457-1468.
4. Buxton, W., Hill, R., & Rowley, P. (1985). Issues and techniques in touch-sensitive tablet input. Proc. of SIGGRAPH '85, 215-224.
5. Carey, J., Hamilton, D, & Shanklin, G. (1986). Development of an Instrument to Measure Rapport Between College Roommates. Journal of College Student Personnel, 27(3), 269-273.
6. Clear, T. & Daniels, M. (2001). A Cyber-Icebreaker for an Effective Virtual Group? Proc. of the Conference on Innovation and Technology in Computer Science Education.
7. Convertino, G., Ganoe, C., Schafer, W., Yost, B., & Carroll, M. (2005). A multipleview approach to support common ground in distributed and synchronous geo-collaboration. Proc. CMV 05. 121-132.
8. Convertino, G., Mentis, H., Rosson, M., Carroll, J., Slavkovic, A., & Ganoe, C. (2008). Articulating common ground in cooperative work: content and process. Proc. of CHI '08, 1637-1646.
9. Ephrem, A., Brungart, D., & Parisi, J. (2006). Effects of Background Noise on Communication in a Collaborative Team Environment. Proc. of the Human Factors and Ergonomics Society.
10. Fisher, M. & Tucker, D. (2003). Games Online: Social Icebreakers That Orient Students To Synchronous Protocol and Team Formation. Journal of Educational Technology Systems, 32(4), pp. 419-428.
11. Goffman, E. (1961). Role-Distance. In Life as Theater: A dramaturgical sourcebook (1975), D. Brisset and C. Edgley (eds.). Aldine Publishing Company: Chicago, IL. 123-132.
12. Goudy, W. Potter, R. (1976). Interview Rapport: Demise of a Concept. The Public Opinion Quarterly, 39(4), 529-543.
13. Hall, Edward T. (1966). The Hidden Dimension. Anchor Books.
14. Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. TOCHI.
15. Honeycutt, J., & Patterson, J. (1997). Affinity Strategies in Relationships: The role of gender and imagined interactions in maintaining college roommates. Personal Relationships, 4, 35-46.
16. Kafai, Y. & Resnick, M. (1996). Constructionism in Practice, 1-8.
17. Karahalios, K. (2004). Social Catalysts: enhancing communication in mediated spaces. Doctoral dissertation, 24 June 2004.
18. Kellogg, W., & Erickson, T. (2002). Social Translucence, Collective Awareness, and the Emergence of Place. Proc. of CSCW '02.
19. King, J., & Star, S. (1990). Conceptual Foundations for the Development of Organizational Decision Support Systems. Proc. of the Hawaii International Conference on Systems Science, 143-151.
20. Landis, J.R. and Koch, G. G. (1977) "The measurement of observer agreement for categorical data" in Biometrics. Vol. 33, pp. 159--174
21. Larson, JR Jr., Cristensen, C., Franz, TM, & Abbott, AS. (1998). Diagnosing Groups: The pooling, management, and impact of shared and unshared case

information in team-based medical decision making. Journal of Personality and Social Psychology, 1998 Jul; 75(1): 93-108.
22. Milgram, S. (1974). Obedience to Authority. Londong, UK: Tavistock Publication.
23. Nardi, B. (2005). Beyond Bandwidth: Dimension of Connection in Interpersonal Communication. Proc. of CSCW 2005, 91-129.
24. (2009).
Proc. of HCI International 2009.
25. Ramanahally, P., Gilbert, S., Anagnost, C., Niedzielski, T., & Velázquez, D. (2009). Creating a Collaborative Multitouch Computer Aided Design Program. Proc. of WinVR '09.
26. Pehsikin, A. (1993). The Goodness of Qualitative Research. Educational Researcher, 22(2), pp. 23-29.
27. Richmond, V. and McCroskey, J. (1995). Immediacy - Nonverbal Behavior in Interpersonal Relations. Boston: Allyn & Bacon.
28. Rogers, Y. & Brignull, H. (2002). Subtle ice-breaking: encouraging socializing and interaction around a large public display. In Proceedings of Public, community and situated displays: Design, use and interaction around shared information displays, CSCW'02 Workshop, New Orleans, Louisiana, USA, 16th November.
29. Russell, D., & Sue, A. (2006). Secrets to Success and Fatal Flaws: The Design of Large-Display Groupware. IEEE Computer Graphics and Applications, 33-41.
30. Schmid, A. (2000). Affinity as social capital: its role in development. Jounral of Socio-Economics, 29(2), 159-171.
31. Scott, S., Grant, K., & Mandryk, R. (2003). System guidelines for co-located, collaborative work on a tabletop display. Proc. of the Eighth European Conference on Computer-Supported Cooperative Work.
32. Sparsh UI (2009). Retrieved September 2, 2009 from: http://code.google.com/p/sparsh-ui/
33. Tuddenham, P., & Robinson, P. (2007). Distributed Tabletops: Supporting Remote and Mixed-Presence Tabletop Collaboration. Proc. of IEEE International Workshop on Horizontal Interactive Human- Computer Systems '07, 19-26.
34. Whittaker, S. (2003). Theories and Models in Mediated Communication. In A. Graesser (ed.): The Handbook of Discourse Processes. Cambridge: Hillsdale, NJ: Lawrence Erlbaum.
35. Zimbardo, P., Maslach, C., & Haney, C. (2000). Reflections on the Stanford Prison Experiment: Genesis, Transformation, Consequences. In Obedience to Authority: Current Perspectives on the Milgram Paradigm, T. Blass (eds.) Mahwah, NJ: Lawrence Erlbaum Associates, 193-238.