Chapter 17

What Makes Collaborations across a Distance Succeed?: The Case of the Cognitive Science Community[1]

Christian Schunn, Kevin Crowley, and Takeshi Okada

Summary

Scientific collaborations are increasingly being conducted at a distance, despite the many factors that make collaboration at a distance difficult. We focus on the discipline of cognitive science because it is both young and highly interdisciplinary, thus potentially increasing both the rewards and difficulties of collaboration at a distance. Using questionnaire data from practicing cognitive scientists, we examined the impact of distance on who is likely to collaborate, the success of the collaboration, and the process of collaboration. There were few differences between those collaborating at a distance and locally. Surprisingly, collaborations at a distance were more successful than collaborations conducted locally. Distant collaborations depended crucially on frequent face-to-face contact. We conclude with a discussion of how these results change our understanding of collaboration in general and, in particular, of collaboration at a distance.

Modern scientific collaborations frequently are carried out by researchers who are geographically distributed, in part because necessary resources for the research are often distributed geographically, and in part because scientific collaborations continue after collaborators change academic or research institutions. In this chapter, we examine the nature and consequences of scientific collaborations that are collocated or distant. We ask whether they function differently than local collaborations. Understanding how such collaborations function might provide better insights in the nature of collaboration more generally.

There are some recently developing pragmatic reasons to study the differences between local and distant collaborations. With the development of the Internet and various computerized collaboration tools, it is becoming possible to achieve some of the features of face-to-face collaboration from a great distance. These products beg the question of what features of local collaborations are particularly important to duplicate. Certainly, not all features of these collaboration tools are worth the expense, and some features may actually hurt the collaboration process (Vera, Kvan, West, & Lai, 1998).

The goal of this chapter is to examine role of distance in the collaborative process within the domain of cognitive science. We ask if scientists working at a distance differ from those working locally, whether working across a distance changes the collaborative process, and what scientists do to make collaborations across a distance succeed and continue. To answer these questions, we present survey data collected from cognitive scientists at a recent conference.

Three Distinctions among Scientific Collaborations

The way that collaborations function is likely to depend upon the needs and relative skills of the collaborators. Three factors are likely to play an important role. First, there is the distinction between peer and apprenticeship collaborations. Peer collaborations are those collaborations

between two researchers at the same professional level (e.g., faculty collaborations, postdoctoral collaborations, graduate student collaborations, etc.). By contrast, apprenticeship collaborations are those collaborations between researchers at different professional levels (e.g., faculty-postdoctoral collaborations, faculty-graduate student collaborations, postdoctoral-graduate student collaborations, etc). Apprenticeship collaborations embody an inherent power difference and knowledge difference that can profoundly affect the process of collaboration. However, in many successful scientific apprenticeship collaborations, the collaborators treat one another in some sense as intellectual equals, who each have the authority to challenge each other's thinking; this phenomenon occurs even in otherwise hierarchical societies, such as Japan (Schunn, Crowley, & Okada, 1998).

Second, there is a distinction between intra- and interdisciplinary collaborations (Dunbar, 1995; Schunn et al., 1998; Schunn, Crowley, & Okada, In press). For example, Schunn et al. (1998; in press) (found important differences in the process, frustrations, and benefits of intradisciplinary and interdisciplinary collaborations. Moreover, different factors predicted which collaborations were successful..Dunbar (1995) found that lab groups in molecular biology were more successful when comprised of members with relatively different training backgrounds.

A third important distinction is local versus distant collaborations, where distant collaborations are those that require travel to the primary places of work of each of the collaborators. Distance has been shown to have a large impact on who is likely to collaborate (Allen, 1977; Kraut, Egido, & Galegher, 1990). One argument is that proximity begets frequent contact, which in turn begets work conversations, which in turn begets collaborative research projects (Kraut et al., 1990). Also many pragmatic factors (time and money) make collaborating across a distance difficult.

Questions about Collaboration at a Distance

Collaboration at a distance among scientists represents a serious challenge to current notions about collaboration. Conventional wisdom is that collaborations on complex projects like those typically found in science require face-to-face interaction. Many social, motivational, and cognitive factors are involved in collaboration on complex tasks that are negatively affected by distance (e.g., Kiesler & Cummings, 2001 [chapter 3]); one wonders why scientists would ever choose to collaborate at a distance and how they ever manage it successfully.

One hypothesis that we explore is that scientists who collaborate at a distance are somehow different than those who only collaborate locally. For example, perhaps collocation is not important for some disciplines or for collaborations that have been ongoing for a long time.

A second hypothesis is that there are no special distinguishing features of scientists who collaborate at a distance, but pragmatic features of the situation could motivate forming acollaboration at a distance. For example, students might seek to finish projects after they have graduated and moved away. Alternatively, collaborations started between colleagues at a given institution may continue when one colleague moves away, perhaps because of continued shared funding.

A third hypothesis is that collaborations at a distance can have real advantages that lead to their formation. At the simple physical resource level, scientists at different locations may have access to different equipment or data, an advantage that may cause them to form distant collaborations. More theoretically interesting is the possibility that there is something about the nature of collaborations that is helped by not being collocated. Most research on effects of distance on collaboration suggests that this is not the case (although see Sproull & Kiesler,

1991). Yet, as we shall demonstrate in this chapter, scientific collaborations at a distance can be distinctly successful.

Whatever the reason for collaboration at a distance, there is also the question of how distance changes the research process. Do scientists collaborate differently when working at a distance? The reduction of physical presence in distance collaborations is likely to affect how researchers exchange information, divide labor, monitor each others progress, deal with frustrations, etc.

A final question that we examine is whether scientists need to behave differently in collaborations at a distance. That is, the factors that determine whether the collaboration is successful may or may not differ as a function of distance. If one believes that distance does have many negative influences on the process of the collaboration (as the research literature suggests), then one should expect that collaboration at a distance requires special methods for keeping the collaboration going (e.g., different organizational features, different supporting technologies, etc.). Distance also might place differential weighting on the role of social and cognitive factors in collaboration success. For example, local collaborations may depend more heavily on good social interactions and similar research styles. Distant collaborations may depend more heavily on good organization and scheduling. Alternatively, distance may exacerbate conflicting styles.

The Case of Cognitive Science

The field of cognitive science is an interesting domain in which to examine collaborations, and in particular, distant collaborations. First, cognitive science is focused on understanding intelligent behavior from a variety of perspectives, and is thus inherently interdisciplinary; it draws primarily on research from cognitive psychology and artificial intelligence, as well as on linguistics, philosophy, neuroscience, education, and anthropology (Collins, 1977; Hardcastle,

1996; Simon, 1980; Simon & Kaplan, 1989; Schunn et al., 1998; Schunn et al., in press). This interdisciplinarity forces researchers to find collaborators from other disciplines. For example, computer scientists interested in modeling human performance often work with cognitive or developmental psychologists to collect data with which to test their models.

Second, cognitive science is a relatively young field, only 25 to 45 years old, depending on whether one uses the first larger-scale community beginnings in the 1970s or the first academic research projects in the discipline as the official start date in the 1950s (Schunn et al., 1998; Simon & Kaplan, 1989). As a young field, comparatively few researchers were trained in cognitive science per se and relatively few institutions have large numbers of cognitive scientists in the home disciplines whose perspective or skills might be necessary for a particular research project. Thus, researchers often seek collaborators at distant universities and institutes, even if they might otherwise prefer to work with researchers in their own institution. However, if researchers initiate distant collaborations in cognitive science, they do so primarily because of their need for intellectual resources than rather than physical resources. Unlike particle physics or astronomy in which data is collected and analyzed across distributed locations because there are only a handful of facilities in the world where the appropriate studies can be conducted, the tools of the trade in cognitive science are widely available. Twenty years ago, a lab with an eye-tracker or a fast computer may have drawn cognitive scientists from all over the world, but this equipment is no longer exotic and expensive. Currently, some neuropsychological research within cognitive science requires specialized facilities such as fMRI (functional Magnetic Resonance Imaging), but, compared with linear accelerators or a space-based telescope, these machines are cheap and comparatively accessible in medical centers through the world.

Historical Data on Collaboration in Cognitive Science

An interplay of interdisciplinarity and distant collaborations was part of cognitive science from its beginning as a field. One of the earliest and most famous collaborations in cognitive science was the collaboration between Allen Newell and Herbert Simon. Through much of the early work, Newell was working at RAND in California and Simon was at the Carnegie Institute of Technology in Pittsburgh (now Carnegie Mellon University). McCorduck (1979) writes:

> Their collaboration also included J. C. Shaw, a senior programmer at RAND, though Simon and Cliff Shaw seldom saw or spoke to each other. Newell carried out the middleman's role, mostly by long-distance telephone between Pittsburgh and Santa Monica. “I thought he was terribly daring, running up those incredible $200-per-month phone bills,” Simon laughs now. “But, then, Al really taught me how to think big about money. (p. 139)

Thus, we see that significant effort and some technology was required to keep the distant collaboration going. McCorduck notes another interesting fact about the role of distance in this collaboration:

> Cliff is a very taciturn guy [Newell says]. One of my dominant recollections is going in and talking with him about some of these problems, and going through a whole session and he wouldn't say a single word, and getting up and leaving. This was when I hardly knew him. It's probably the case that the whole scientific enterprise with the three of us would never have worked out if we were all sitting in one place. Cliff found this way of working, with me located miles away, to be just about the right level of controlled interaction for him to flower. And so I operated both by letter and by telephone - by two and three hour-long conversations a week through this whole period - so in fact the three of us never got together, almost. (p. 144)

Thus, we see distance playing dual roles in this famous collaboration in cognitive science. On the one hand, distance collaborations were a barrier to communication that had to be overcome through effort and expense. On the other hand, distance proved to have some positive consequences for the process of the collaboration.

A Study of Collaboration in Cognitive Science

We used a correlational, retrospective methodology asking scientists to reflect on a particular collaboration that had just produced an intermediate product, a conference publication. We emailed authors from a particular an yearly meeting of the Cognitive Science Society, and asked them a series of questions about the structure and process of their paper. The original goals of the study were to understand the process of collaboration in science and to examine the difference between intra- and interdisciplinary collaborations (Schunn et al., 1998; Schunn et al., in press). For this chapter, we have compared local and distant collaborations.

Method

The Annual Meeting of the Cognitive Science Society was first convened in 1979. In 1983, it became peer reviewed. The paper selection criteria have become stricter each year since then. In the past several years, approximately 35% of submitted papers were accepted. Over the years, the conference grew steadily in size, stabilizing in the early 1990s at approximately 140 papers and posters (including member abstract posters) and 500-600 attendees. The modal number of authors is 2, followed by 1, and then 3 authors. Less than five percent of papers have more than 3 authors. Throughout its history, the conference has been a central location for cognitive scientists to meet, network, and exchange the latest findings from their ongoing research.

All authors of multi-authored papers and posters to be presented at the 17th annual conference in Pittsburgh, Pennsylvania were contacted two months prior to the conference by

email. Permission to conduct the interview studies was obtained from the conference organizers in advance of contacting the participants. Ninety-six multi-author papers and posters had been accepted, 94 for which we had valid email addresses, consisting of 222 total authors. Seventy-five (34%) questionnaires were completed and returned, representing at least one author from 56 (60%) of the papers.

Three years later, we sent email to the authors of the 1995 conference papers asking them whether the work reported at that conference had been published elsewhere in the form of a book chapter or journal article. It is possible that the authors and reviewers were not able to immediately evaluate the success of interdisciplinary workr. We received responses from over 82% of our original set of respondents (n = 34 local, n = 12 distant).

Procedure

The multiple-choice questionnaire we developed was based upon exploratory interviews with cognitive scientists and previous questionnaire studies (Okada et al., 1995; Schunn et al., 1995). Three sections of the questionnaire are relevant to the current analyses. The first section assessed the primary backgrounds and professional status (faculty, graduate student, etc.) of the participant and the participant's collaborators. The second section asked participants to estimate how successful they thought the project had been and how likely they were to continue working with their collaborators. The third section asked participants to estimate how often communication had occurred within the collaboration, the means of communication (face-to-face meetings, email, etc.), the mesh or clash of the collaborators' background knowledge and intellectual styles, and the benefits and frustrations of the collaboration. All questions were in multiple-choice format, although additional comments and alternative, additional write-in options were possible for each question.

The unit of analysis for most of the analyses presented below is a paper (i.e., project) rather than a participant. Therefore, we used data from only one author per paper, the highestranking person (e.g., first author rather than second author), on the assumption that this person would have the most detailed knowledge of the project. When responses were received from multiple authors of the same paper, their responses were generally quite consistent. Moreover, changing the unit of analysis from paper to respondent did not change any of the results obtained in our analyses.

From the affiliations listed on the papers and posters, we categorized each collaboration as local or distant. Local collaborations (n = 40) were all those collaborations for which the authors were from the same department or different departments within the same institution. Distant collaborations were all those collaborations for which authors were either at different institutions within the same city (n = 2), in different cities within the same country (n = 7) or in different countries (n = 6), producing a total of n = 15 distant collaborations. When there were more than two authors, we used the locations of the first two authors. However, in data from third and fourth authors they always had the same local or distant designation as that of the first two authors. This coding scheme does not necessarily reflect where the authors were while the bulk of the work was completed. However, given that this is a conference submission and that much important work happens at the writing phase of a project, this scheme is likely to be an accurate reflection of where collaborators were for at least part of the project.

For continuous to continuous variable analyses, we used Fisher's r-to-z and multiple linear regressions. For nominal to continuous variable analyses, we used ANOVAs and MANOVAs. For continuous to nominal and nominal to nominal variable analyses, we used logistic regression.

Results

We first examined whether there were any differences in the types of collaborators and collaborations that occurred across distance. If there were differences in collaborator or collaboration type, then we would need to establish in our analyses whether or not differences in the process between local and distant collaborations were due to these other structural differences.

Table 17-1 about here

Table 17-1 shows the profiles of local and distant collaborations: the mean number of authors, the percentage of papers with any collaborators of different training backgrounds (using categories at the grain size of computer science, psychology, philosophy, linguistics, etc), the percentage of papers with all collaborators of equal professional status (at the grain size of undergraduate student, graduate student, postdoctoral, faculty member), the percentage of papers with faculty members as collaborators (based on the status level of the author responding to the questionnaire), the percentage of papers with psychology as the primary training background of one of the collaborators, and the percentage of papers in which the collaboration was less than or equal to one year old.. A multiple logistic regression predicting collaboration distance (using the six factors listed in Table 17-1 as potential predictors) found only one significant predictor and one marginally significant predictor. Distant collaborations involved marginally significantly fewer psychology collaborations ($\chi^2(55) = 2.3$, $\underline{p} < .15$) and significantly more recent collaborations ($\chi^2(55) = 4.3$, $\underline{p} < .05$). Both types of collaborations had very similar numbers of authors in the collaborations a similar frequency of multidisciplinary collaborations. Finally, distant collaborations had slightly (but not significantly) higher proportions of peer collaborations and faculty collaborations.

The relatively high proportion of recently begun collaborations among the distant collaborations suggests that the majority of these collaborations were not primarily local collaborations in which one of the authors had moved recently. Probably, the majority of these collaborations had truly been done at a distance. The relatively high proportion of short-term collaborations among the distant collaborations is also consistent with the notion that distant collaborations are difficult and may not be continued as long as local collaborations are continued.

Because there were two significant differences between local and distant collaborations (i.e., whether the first author was trained as a psychologist and collaboration length), subsequent analyses included those two factors as control variables .

Success of Collaboration

Before turning to more detailed analyses of structure and process, we will look at some bottom-line measures. Were distant collaborations less successful, reflecting their inherent greater difficulty levels? Figure 17-1 presents a variety of success measures for local and distant collaborations. According to the collaborators, the local and distant collaborations were just as likely to be rated "very successful." However, to our surprise, distant collaborations were more likely to result in a publication three years later. One possible explanation is that perhaps only very good ideas warrant starting a distant collaborations; if sufficient progress is made to produce a conference paper, the odds are better that it will be published as a full paper later. Future research will have to be conducted to examine this possibility further.

Figure 17-1 about here

Note that rated success and probability of publication is not the same thing. Many factors enter into whether research is considered successful, and publication is only one of them.

Similarly, many factors enter into whether a project is published (e.g., interest of the research community, writing ability of the researchers, importance of publications for ones career, etc.). In the case of interdisciplinary work, work that the researchers consider very novel and interesting may turn out to be too novel for the community and hence not publishable. Thus, it is not entirely surprising that we may fail to find an association between rated success and the probability of publication.

Another measure of the success of a collaboration is whether it is continued. Despite the higher probability of publishing the work three years later, respondents in distant collaborations were slightly less likely than respondents in local collaborations to think they would continue the collaboration (and also slightly less likely to actually continue the collaboration three years later). These (small) differences might be attributed to the greater difficulty associated with distant collaborations.

Benefits of Collaboration

There were 14 different benefits of collaboration that respondents regularly listed. Overall, the frequency with which each benefit was selected was similar across local and distant collaborations (r =.81). A MANOVA revealed that twelve of the fourteen showed no difference at all (in decreasing frequency of mention): division of labor, different ideas, stimulating, increases enjoyment, challenging, motivating, different styles, different resources, increases the speed of research, similar ideas, helps monitor progress, and supports the research. Two factors were affected by collaboration distance, and these effects did not interact with first author discipline or collaboration length, the control variables. First, distant collaborators were less likely to mention having similar styles as a benefit of the collaboration (0% vs. 25%) in distant and local collaborations respectively; F (1,52) = 4.5, MSE=.14, p <.05). Second, distant

collaborations were more likely to mention having interaction and discussion as a benefit of collaboration (100% vs. 72% for distant and local collaborations, respectively; F (1,52) = 5.1, MSE=.15, p <.05). One might imagine that style similarity is not as relevant when the collaborator is far away. Yet, when the collaborator does interact, the interaction and discussions are always important or else their would be little incentive to participate in the collaboration.

Partially as a consistency check, the differences and similarities between local and distant collaborations were examined at a more fine-grained level, separating out the various levels of distance between collaborators. In all but one case, the fine-grained analyses produced similar results, suggesting that our boundary between local and distant collaborators was appropriately chosen. The one exception was for international collaborations vs. intercity collaborations: none of the intercity collaborators listed having different resources as a benefit, whereas two-thirds of the international collaborators listed having different resources as a benefit. Given how widely funding practices, subject pool types, and computational resources differ across countries, this difference is not entirely surprising, and is consistent with our own experience in international collaboration.

Frustrations of Collaboration

When asked about frustrations associated with the collaboration, the most common response (just under 40%) was to say there were no frustrations. This response may reflect a reluctance to send any negative comments in email about the collaboration. Seven frustrations were mentioned consistently, and their relative frequency was similar across local and distant collaborations (r =.73). The frustrations that were mentioned in both types of collaborations were (in decreasing frequency of mention): communication problems, different ideas, slow research process, different styles, personality differences, assignment of credit problems, and motivation problems.

A MANOVA revealed that none of the frustrations were mentioned significantly more often in distant collaborations; on average, distance accounted for only 2% of the variance in self-reported frustrations. Only one difference approached significance: having different styles was mentioned only in the local collaborations as a frustration.

Communication

Communication attributes of the collaborations were similar across local and distant collaborations. Respondents were just as likely to say they had very similar ideas, had a collaborator who frequently provided alternative hypotheses, had a collaborator with a similar style, felt they had an equal status in the collaboration, used email as a primary collaboration method, and communicated at least several times a week (see Figure 17-2).

Figure 17-2 about here

Two significant differences distinguished the communication in distant and local collaborations. The first was the use of regular research group meetings. Unsurprisingly, local collaborations were much more likely to use research group meetings as a primary communication method, F (1,53) = 5.8, MSE=.23, p <.02. The second difference, equally unsurprising, was the frequency of face–to-face meetings, F (1,53) = 4.4, MSE=2.4, p <.05. A larger proportion of the distant collaborators met infrequently. Neither of these effects of distance interacted with collaboration length and first author training background, the two factors partially confounded with distance.

Distribution of Work

Respondents were asked to rate what percent of the work reported in the conference paper was done by them, both overall and on different components of the project: selecting the research questions, selecting a design, providing the resources for the work, collecting the data, analyzing

the data, and writing the paper. On all of these dimensions, there was generally an asymmetry in the distribution of work across the authors; first authors tended to do more of the work. The one exception was providing resources for the work, which was slightly more often associated with second and third authors. Note, that whereas the sum of these percentages across authors for a given paper on a given dimension tended to be greater than 100%, there was generally good consistency in how the workload was perceived to have been distributed.

Did this authorship-work-asymmetry differ by collaboration type? Generally speaking, the answer is "no," with one exception. For data collection, there was a much smaller asymmetry of first and second authors in distant collaborations. In local collaborations, first authors claimed to have done 80% of the data collection and second authors only 14%. By contrast, in distant collaborations, first authors claimed to have done only 49% of the data collection and second authors 28%. This effect did not interact with collaboration length and first author training background. It appears that one of the uses of distant collaborations was data collection across multiple locations. When data is collected at one site, there is less need for multiple collaborators to contribute to data collection than when data are collected at multiple sites. The other activities of research are much less location-specific by nature, and thus it is not surprising the other dimensions did not show an effect of collaboration type.

Predicting Success

What factors are important in predicting the success of a collaboration and do those factors differ between local and distant collaborations? Here we focus on the rated success and estimated probability of continuing to work together because we have the largest number of responses for those variables. We examined which of the 8 process features of the collaboration were predicted local and distant collaborations separately. In predicting rated success for local collaborations,

two factors proved important: whether the primary communication method was the research group meeting (r =.45, p <..05) and the similarity of initial ideas (r =.50, p <.05). These predictors hold for all the varieties of local collaborations that could be distinguished (i.e., there was no interaction with type of collaboration in a multiple regression): psychologists and computer scientists, short-term collaborations and long-term collaborations, and for faculty collaborations and graduate student collaborations.

By contrast what predicted the rated success of distant collaborations was the frequency of face-to-face meetings (r =.60, p <.05). To verify that these correlation patterns were indeed different across local versus distant collaboration types, we examined the strength of the correlation with each of these four predictors in the other collaboration type. The correlation strengths were all almost zero in the opposite collaboration type. Thus, the predictors are indeed quite different. Table 17-2 summarizes the results of these regression analyses.

Table 17-2 about here

Our analyses do not say that the factors that were not statistically predictive are not important for successful collaborations. We believe, instead, that, in this community there are generally sufficient levels of the other factors to support successful performance. Further, we have restriction of range because all of these collaborations had obtained a measure of success by having papers accepted at the conference.

What do the correlation patterns imply? First, it appears that the research group meeting is a very important part of the research process for collaborations that can support them (i.e., local collaborations), but not for collaborations that cannot support them (i.e., distant collaborations). Second, the role of similar ideas is perhaps to avoid important conflict about the details and directions of the project. In distant collaborations, more of the work is distributed

(e.g., data collection) and having different initial ideas may be less a source of conflict (even though the distribution of similarity of ideas was the same for local and distant collaborations). Third, it appears that what is most crucial for the success of long distance collaborations is the feature they are most lacking: face-to-face contact.

For predicting the estimated probability of continuing to work together, one factor proved predictive in distant collaborations: whether the collaboration was felt to involve an equal status relationship ($\underline{r}$ =.52). By contrast, no factors predicted that outcome in local collaborations (e.g., $\underline{r}$ =.07 for equal status). Schunn et al. (1998) found that interdisciplinary collaborations (but not intradisciplinary collaborations) also required an equal status relationship for the collaborators to want to continue the collaboration. Together these results suggest that collaborations that require extra effort to maintain, such as distant or interdisciplinary collaborations, require the mutual respect that comes with being treated with equal status in order to want to continue the collaboration. By contrast, easier collaborations (local and intradisciplinary collaborations) are seem not as sensitive to these social perceptions.

Discussion

Our investigation of scientific collaborations in a new, interdisciplinary field like cognitive science has provided several insights into the similarities and differences between the structure and process of local and distant collaborations. At their core, the collaboration types are quite similar. That is, it does not appear that there are large differences in who decides to engage in distant collaborations, with the possible exception that psychologists were slightly more likely to do so than researchers from other disciplines. Moreover, there were many similarities in the frequencies of frustrations and benefits, and in the types of processes that are used to coordinate research.

Distance Changing the Process

We did, however, find that distance did change some aspects of the research process. Not surprisingly, distant collaborations involved less use of research group meetings and fewer face-to-face meetings. Distant collaborations also involved a more even sharing of data collection activities across the coauthors, suggesting perhaps that one of the advantages of distant collaborations in this setting is access to a wider variety of data sources. Most importantly, different factors were important for successful collaborations at a distance. Local collaborations depended upon heavy use of research group meetings and similarity of initial ideas. By contrast, distant collaborations depended upon having frequent face-to-face meetings.

In a recent study, Susan Epstein (in press) interviewed 21 cognitive scientists who had engaged in successful interdisciplinary collaborations (as measured by funding and publications). A few of her findings are of particular relevance here. First, frequent meetings were important. For example, one researcher said that it was important to have students and collaborators present their work to each other regularly. Second, it appeared that the frequency of meetings was especially important at the beginning of a collaboration. For example, many long-term interdisciplinary collaborations start out with a regular weekly meeting. Third, the pragmatics of distant collaborations was an important and difficult hurdle to overcome. Frequent meetings were difficult or impossible. Visiting collaborators was expensive in both money and time, and involved significant physical and emotion stress. Similar observations were drawn from interviews of Japanese cognitive scientists who had engaged in successful interdisciplinary collaborations (Okada et al., 1995). The findings of these several studies suggest that regular meetings have both important advantages and disadvantages in distant collaborations.

Distance as a Benefit

Much of the research on distance work and collaborations across a distance has focused either on the negative consequences of distance, e.g., on trust between collaborators (Cramton, 2001 [chapter 8]), shared mental models (De Meyer, 1991; Olson et al., 2001 [chapter 5]), shared practice styles (Mark., 2001 [chapter 11]), or methods for remediating those negative consequences, e.g., through improved collaboration software (Holtham, 1991; Olson & Olson, 1991; Hollingshead, Fulk, & Monge, 2001 [chapter 14]). One of the most surprising findings of our investigations was that there are some positive benefits associated with distance *per se*. This difference may in part be a selection artifact. That is, researchers may be pickier about projects they work on across a distance as compared with those they work on locally. However, at some level, because we selected only accepted conference papers for a conference with a high rejection rate, the selection artifact should have been reduced.

We argue that there are positive consequences associated with distance. As the example of the Newell, Simon, and Shaw collaboration showed, not all people work well in the same physical space. Azmitia and Crowley (2001) argue that collaboration is best thought of as a rhythm of relatively social and relatively individual moments. It may be that physical proximity is beneficial for the relatively social components and damaging for the relatively individualistic components of collaboration.

Alternatively, it may be that collaboration across a distance provides access to wider influx of ideas. Researchers at different institutions will attend different talks and are more likely to have non-overlapping research groups. These differences can provide access to a broader set of expertise and ideas. This access to new ideas also has been examined by Walsh and Maloney (2001 [chapter 18]) in their study of scientists. In our own research, which has been primarily a collaboration across a distance, this access to different groups and ideas has been very

influential. At the level of theoretical frameworks, it has produced an infusion of both information processing and sociocultural approaches into our work. At the level of research methodologies, it has provided input from numerous research groups into developing instruments (e.g., items in our survey questionnaire). At the level of contextual situativity and practical implications, it has provided models of academic research and work practices across nations (the US and Japanese academic systems) and across fields (applied cognitive science, cognitive psychology, developmental psychology, and education).

Recently, we have partially replicated our finding of slightly greater success associated with distant collaborations. For the 2000 Annual Meeting of the Cognitive Science Society, all submissions (prior to acceptance) were categorized according to number of authors, location of authors (distant versus local), and discipline of authors (n = 163). The dependent measure was whether the submissions were accepted at the conference. Overall, only 36% of submissions were accepted. Although the difference was not statistically significant, 40% of local collaborations versus 48% of distant collaborations were accepted. Note that this analysis excludes single author submissions, which had a substantially lower acceptance rate—indeed, number of authors proved the best predictor whether a paper would be accepted, suggesting that collaborations are an important part of scientific progress in cognitive science.

Models of Collaboration

What possible models of collaboration processes might underlie the results found here? Two general types of accounts should be distinguished: an individual differences account and a process account. On the one hand, it is possible that different types of people engage in distant collaborations, and this produces a different set of factors which are important, not because the collaboration is being conducted at a distance, but because the people themselves are different.

On the other hand, it may be that the process of collaboration at a distance is itself different, and this by itself produces a different set of factors that are important. Our data tends to support the process account. There were some differences in background profiles across the local and distant collaborations. However, we examined the influence of those factors and found that they could not account for the differences in which factors were important across local and distant collaborations. We did also find some differences in the collaboration process that one might expect, and we have already discussed how those could plausibly lead to the strong influence of the obtained important factors. However, we have not ruled out all possible background differences. For example, we did not gather detailed information about personality types, nor amount of experience with collaborations generally.

What do our findings tell us about the nature of scientific collaborations generally? They certainly highlight the shared and co-constructed nature of collaborative work. Work is not simply divided among the collaborators and then continued, independently, in parallel. Instead, the success and continuity of the collaboration depends upon features of the relationship between the collaborators: that they meet often enough, in the right circumstances, and treat each other as equals. In other words, even if one is interested in the cognitive products of the collaboration, one must still pay careful attention to the social features in order to understand the process and success of collaboration. However, our findings also highlight the potential positive value of distance, and this surprising result must be explored further in future work.

Implications for Work at a Distance

Our findings suggest that there may be important advantages of collaboration at a distance and this has important practical implications for those researchers and practitioners who wish to understand work at a distance. First, our findings place in question the general push to create

new, expensive technologies that duplicate as many possible features of face-to-face collaboration at a distance (e.g., video-conferencing, shared virtual reality environments, etc.). Collaborations at a distance may have their own unique advantages that we may not want to remove through technological changes.

This realization brings us to the second important practical implication of our findings for distance collaboration. In order to decide when to bring in supporting technologies and which technologies to use, we need to know more about why distance might help the collaborative process. Our research has provided the first clues that it might help, but much further research needs to be conducted on why and under which circumstances it might help.

Methodological Caveats

There are many methods that one could use to study scientific collaboration processes, including experiments using undergraduates conducted in the lab (Okada & Simon, 1997; Azmitia & Crowley, 2001), observations of scientists working on a project (Dunbar, 1995, 1997; Trickett, Fu, Schunn, & Trafton, 2000; Trickett, Trafton, & Schunn, 2000), interviews of scientists about their collaborations (Epstein, In press; Okada, Crowley, Schunn, & Miwa, 1996; Schunn et al., 1998; Schunn et al., In press; Thagard, 1994), and historical analyses of famous collaborations (Okada et al., 1995). Each method has its advantages and disadvantages, particularly emphasizing different grain sizes of analysis. Each method contributes to our understanding of scientific collaboration. We were primarily interested in what factors motivated long distant collaborations and what factors contributed to the success of the collaboration. Therefore, we chose to use a methodology that gave long-term data rather than detailed short-term data. This also motivated our decision to collect data from practicing scientists rather than undergraduates in a lab experiment.

As does every methodology, our methodology has its disadvantages, which should be mentioned here as a caveat to our results. First, we relied heavily on self-report data. When possible, we compared the respondents' answers with publicly available information about them, and in all of those cases it was clear that the respondents were answering those questions accurately. However, there may have been some biases in their responses to some of the more subjective questions.

A second important caveat is that we had a relatively small number of distant collaborations in our sample. It is likely that this relatively small N prevented us from detecting the role of other important factors in the process and success of distant collaborations. At the same time, the advantage of small Ns is that one must focus on the factors with the largest effect size, a good strategy for the first research forays into a new area. Yet, the small N also prevented us from further exploring interactions with other distinctions in collaboration type. Within local collaborations, we were able to determine that our patterns held for peer and apprenticeship collaborations, student and faculty collaborations, and computer science and psychology collaborations. Interactions, however, may exist for distant collaborations. For example, an apprenticeship collaboration may require additional structure at a distance that a peer collaboration does not.

Third, it must be acknowledged that we had a potentially narrow range of successfulness in our sample since all projects had been accepted at the conference. It is likely that many other factors play an important role in the successful collaborations, but were present in sufficient quantity in all of our respondents' collaborations. So, we must qualify our results as being about what differentiates collaborations that have managed to produce at least one work product. However, focusing on collaborations that had all reached a similar threshold insured that there

was some level of equivalence in the phase of research across the various types of collaborations examined.

Fourth, our analyses are entirely correlational, and thus we do not yet know the true causal nature of relationships between our variables, and more importantly there were many correlations between our measured variables. Table 17-3 presents the intercorrelations between all the measures that were significantly associated with collaboration distance, directly, or behaved differently in the local and distant collaboration cases. Luckily, none of the correlations are so large that one should have large worries about mediated correlations.

Fifth, we have focused on one particular situation (a new, interdisciplinary field), and the generalizability to other situations must be explored further. In particular, the interdisciplinary nature of cognitive science, while perhaps forcing a greater level of distant collaborations, may perversely require even more face-to-face contact. Thus, the dependence upon face-to-face meeting frequency may not be as strong in other, more established, monodisciplines.

References

Allen, T. J. (1977). Managing the flow of technology. Cambridge, MA: MIT Press.

Azmitia, M.A. & Crowley, K. (in press). The rhythms of scientific thinking: A study of collaboration in an earthquake microworld. In K. Crowley, C. Schunn, & T. Okada (Eds.) Designing for science: Implications from everyday, classroom, and professional settings. Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, A. (1977). Why cognitive science. Cognitive Science, 1, 1-2.

Cramton, C.D. (2001). Attribution in distributed work groups. In P. Hinds & S. Kiesler, Distributed work (pp xx-xx). Cambridge, MA: MIT Press.

De Meyer, A. (1991). Tech talk: How managers are stimulating global R&D communication. Sloan Management Review, 32, 49-58.

Dunbar, K. (1995). How scientists really reason: Scientific reasoning in real-world laboratories. In R. J. Sternberg & J. E. Davidson (Eds.), The nature of insight (pp. 365-395). Cambridge, MA: MIT Press.

Dunbar, K. (1997). How scientists think: On-line creativity and conceptual change in science. In T. B. Ward & S. M. Smith (Eds.), Creative thought: An investigation of conceptual structures and processes (pp. 461-493). Washington, DC, USA: American Psychological Association.

Epstein, S. L. (In press). Making interdisciplinary collaboration work. In S. J. Derry & M. A. Gernsbacher (Eds.), Problems and promises of interdisciplinary collaboration: Perspectives from cognitive science. Mahwah, NJ: Erlbaum.

Hardcastle, V. G. (1996). How to build a theory in cognitive science. Albany, NY: State University of New York Press.

Hollingshead, A. B., Fulk, J., & Monge, P. (in press). Fostering intranet knowledge sharing:An integration of transactive memory and public goods approaches. In P. Hinds and S. Kiesler. Distributed work (pp. xx-xx). Cambridge, MA: MIT Press.

Holtham, C. (1994). Groupware: Its past and future. In P. Lloyd (Ed.), Groupware in the 21st Century. Westport, CT: Praeger.

Kiesler, S. & Cummings, J. N. (2001). What do we know about proximity and distance in work groups? A legacy of research. In P. Hinds & S. Kiesler, Distributed work (pp xx-xx). Cambridge, MA: MIT Press

Kraut, R. E., Egido, C., & Galegher, J. (1990). Patterns of contact and communications in scientific research collaboration. In J. Galegher & R. E. Kraut & C. Egido (Eds.), Intellectual teamwork: Social and technological foundations of cooperative work. Hillsdale, NJ: Erlbaum.

Mark, G. (2001). Conventions for coordinating electronic distributed work: A longitudinal study of groupware use. In P. Hinds & S. Kiesler, Distributed work (pp. Xx-xx). Cambridge, MA: MIT Press.

Nardi, B. A., & Whittaker, S. (2001). The place of face-to-face communication in distributed work. In P. Hinds & S. Kiesler, Distributed work (pp. Xx-xx). Cambridge, MA: MIT Press.

Okada, T., Crowley, K., Schunn, C. D., & Miwa, K. (1996). Collaborative scientific research in Japanese cognitive science: Analyses of questionnaire survey data. Paper presented at the 1996 Meeting of the Japanese Cognitive Science Society. Advanced Telecommunications Research Institute International, Kyoto, Japan, June 20-22.

Okada, T., Schunn, C. D., Crowley, K., Oshima, J., Miwa, K., Aoki, T., & Ishida, Y. (1995). Collaborative scientific research: Analyses of historical and interview data, Paper presented at the 1995 Meeting of the Japanese Cognitive Science Society.

Okada, T., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. Cognitive Science, 2, 109-146.

Olson, G. M., & Olson, J. S. (1991). User-centered design of collaborative technology. Journal of Organizational Computing, 1, 61-83.

Olson, J. S., Teasley, S., Covi, L., & Olson, G. (2001). The (currently) unique advantages of collocated work. In P. Hinds & S. Kiesler, Distributed work (pp. Xx-xx). Cambridge, MA: MIT Press.

Schunn, C. D., Crowley, K., & Okada, T. (1998). The growth of multidisciplinarity in the Cognitive Science Society. Cognitive Science, 22(1), 107-130.

Schunn, C. D., Crowley, K., & Okada, T. (In press). Cognitive science: Interdisciplinarity now and then. In S. J. Derry & M. A. Gernsbacher (Eds.), Problems and promises of interdisciplinary collaboration: Perspectives from cognitive science. Mahwah, NJ: Erlbaum.

Schunn, C. D., Okada, T., & Crowley, K. (1995). Is cognitive science truly interdisciplinary?: The case of interdisciplinary collaborations. In J. D. Moore & J. F. Lehman (Eds.), Proceedings of the 17th Annual Conference of the Cognitive Science Society (pp. 100-105). Mahwah, NJ: Erlbaum.

Simon, H.A. (1980). Cognitive science: The newest science of the artificial. Cognitive Science, 4, 33-46.

Simon, H. A., & Kaplan, C. A. (1989). Foundations of cognitive science. In M. I. Posner (Ed.), Foundations of cognitive science. Cambridge, MA: MIT Press.

Sproull, L., & Kiesler, S. (1991). Connections : New ways of working in the networked organization. Cambridge, MA: MIT Press.

Thagard, P. (1994). Collaborative knowledge in Science. In A. Ram & K. Eiselt (Eds.), Proceedings of the 16th Annual Conference of the Cognitive Science Society (pp. 990-991). Hillsdale, NJ: Erlbaum.

Trickett, S. B., Fu, W.-T., Schunn, C. D., & Trafton, J. G. (2000). From dipsy-doodles to streaming motions: Changes in representation in the analysis of visual scientific data., Proceedings of the 22nd Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum.

Trickett, S. B., Trafton, J. G., & Schunn, C. D. (2000). Blobs, dipsy-doodles and other funky things: Framework anomalies in exploratory data analysis, Proceedings of the 22nd Annual Conference of the Cognitive Science Society. Mahwah, NJ: Erlbaum.

Vera, A. H., Kvan, T., West, R. L., & Lai, S. (1998). Expertise, Collaboration and Bandwidth Usability of Groupware, Proceedings of ACM CHI 98 Conference on Human Factors in Computing Systems (pp. 503-510). New York: ACM Press.

Walsh, J. P. & Maloney, N. G. (in press). Computer network use, collaboration structures, and productivity. In P. Hinds & S. Kiesler, Distributed work (pp. Xx-xx). Cambridge, MA: MIT Press.

Endnote

[1] Work on this project was funded by a grant from the Mitsubishi Bank Foundation. The first author would also like to thank the Department of Psychology at the University of Basel for providing time and support to write this chapter. We would also like to thank Jae Yun Moon, Sara Kiesler, and Pamela Hinds for helpful comments made on earlier versions of this chapter.

Table 17-1. Differences Between Local and Distant Collaborations.

	Local n = 40	Distant n = 15	
Mean (SD) # authors	2.4 (0.7)	2.3 (0.8)	
% Multidisciplinary	60%	67%	
% Peer collaborations	18%	33%	
% Faculty collaborators	42%	60%	
% Psychology as training background	58%	33%	
% Recent collaborations (less than or equal to 1 year)	23%	47%	*

Note. *p < .05

Table 17-2. Predicting Rated Success of Local and Distant Collaborations as a Function of Process Features of These Collaborations.

	Local Collaborations			Distant Collaborations		
Independent variables	Beta	SE	p	Beta	SE	p
Intercept	0.60	0.55		-0.37	2.82	
Similarity of initial ideas	0.26	0.17	*	-0.29	0.39	
Frequency alternative hypotheses were proposed	-0.15	0.13		-0.44	0.42	
Similarity of styles	0.0	0.13		0.42	0.43	
Perceived equal status	0.10	0.20		1.01	1.01	
Primary communication is email	-0.18	0.18		0.06	0.63	
Primary communication is research group meetings	0.42	0.19	*	0.09	0.57	
Communication frequency	0.11	0.13		-0.13	0.32	
Face-to-face frequency	-0.07	0.11		0.36	0.16	+
N	39			13		
R^2	.44			.70		

Note. +$p < .10$, *$p < .05$

Table 17-3. Correlations Among Items Differentiating Distant and Local Collaborations.

	1	2	3	4	5	6	7	8	9	10	11	12
1. Same location	1.0	-.22	-.30	.07	.14	.28	-.30	.12	-.19	-.12	.31	.28
2. 1st author psychology	-.22	1.0	.08	-.35	-.28	-.09	-.15	-.31	-.01	.24	.00	-.09
3. Collaboration length	-.30	.08	1.0	.00	-.15	-.33	-.12	-.15	-.02	.13	-.18	.06
4. Rated success	.07	-.35	.00	1.0	.09	.16	.37	.38	.09	-.19	.29	.32
5. Will continue collaborating	.14	-.28	-.15	.09	1.0	-.02	.04	.02	.09	-.04	.15	.12
6. Benefit same style	.28	-.09	-.33	.16	-.02	1.0	.24	.30	.13	-.22	.03	.17
7. Benefit interact & discuss	-.30	-.15	-.12	.37	.04	.24	1.0	.30	.18	.08	.20	.00
8. Similarity of init ideas	.12	-.31	-.15	.38	.02	.30	.30	1.0	.04	.02	.27	.02
9. Rated equal status	-.19	-.01	-.02	.09	.09	.13	.18	.04	1.0	-.24	-.12	-.11
10. Primary comm. email	-.12	.24	.13	-.19	-.04	-.22	.08	.02	-.24	1.0	.03	-.35
11. Primary comm r. group	.31	.00	-.18	.29	.15	.03	.20	.27	-.12	.03	1.0	.29
12. Face-to-face frequency	.28	-.09	.06	.32	.12	.17	.00	.02	-.11	-.35	.29	1.0

Note. - N varies slightly by item depending upon the response rate to each item, with $52 \leq n \leq 55$.

Figure 17-1. For local and distant collaborations, rated success levels (N=54), probability of publishing the work three years later (N=45), percentage estimated to be very likely to continue working together (N=53), and actual percentage working together three years later (N=45).

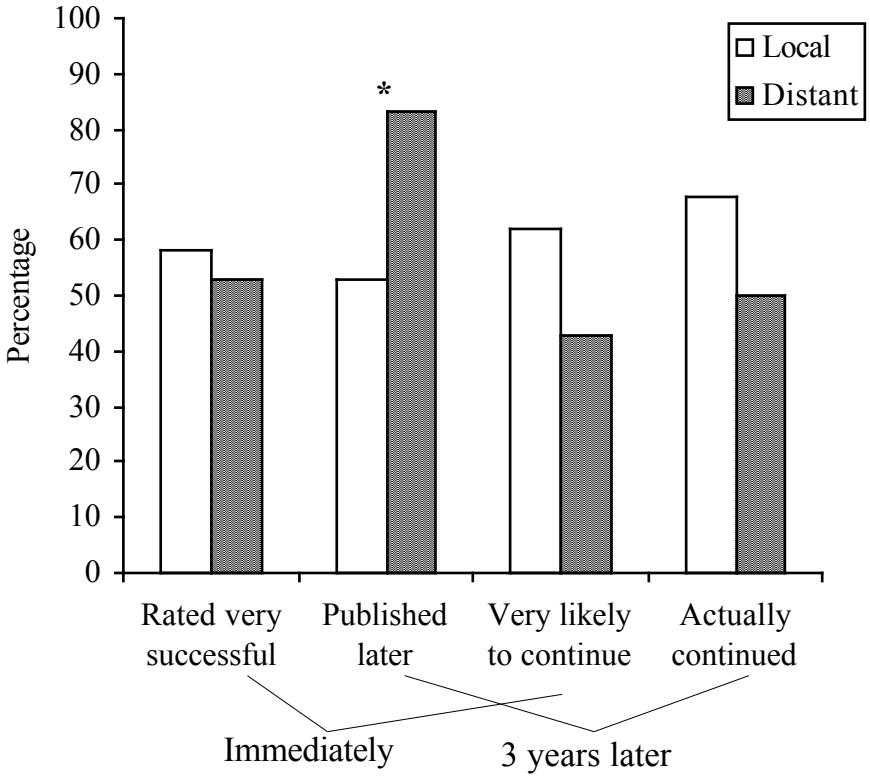

Figure 17-2. The percentage of collaborations with very similar initial ideas, collaborators who frequently presented alternative hypotheses, had a similar style, felt they had an equal status, used email as a primary communication method, used the research group meeting as a primary communication method, communicated at least several times a week, and met face-to-face less than once a month.