

Pair Research: Matching People for Collaboration, Learning, and Productivity

Robert C. Miller¹, Haoqi Zhang², Eric Gilbert³, Elizabeth Gerber²

¹MIT CSAIL
Cambridge, MA 02139
rcm@mit.edu

²Northwestern University
Evanston, IL
{hqz,egerber}@northwestern.edu

³Georgia Tech College of Computing
Atlanta, GA
gilbert@cc.gatech.edu

ABSTRACT

To increase productivity, informal learning, and collaborations within and across research groups, we have been experimenting with a new kind of interaction that we call *pair research*, in which members are paired up weekly to work together on each other's projects. In this paper, we present a system for making pairings and present results from two deployments. Results show that members used pair research in a wide variety of ways including pair programming, user testing, brainstorming, and data collection and analysis. Pair research helped members get things done and share their expertise with others.

INTRODUCTION

Academic research is a hard endeavor. The path to discovery has many hurdles that can dampen enthusiasm for making progress. While collaboration and help-seeking can boost productivity and produce better research, few mechanisms in academic research directly promote and facilitate it. This is particularly troubling for young researchers, who may only receive feedback on their work in group meetings, work with only a few collaborators (one of whom is their advisor), and be more easily frustrated by setbacks, real or imagined.

We have been experimenting with a new kind of interaction within a research group that we call *pair research*, as a generalization of pair programming. Each week, group members pair up, guided by a matching algorithm. Each pair meets for a one- or two-hour session, of which half the time is spent working together on one person's project, and the other half working on the other person's project. The work might be any activity involved in computer science research, including pair programming, user testing and design critique, data collection and analysis, brainstorming and research discussion, and mentoring and advising. The following week, different pairs are formed, and the process repeats.

Inspired by pair programming, pair research seeks to increase collaboration, promote informal learning, and boost productivity within and across research groups. For collaboration,

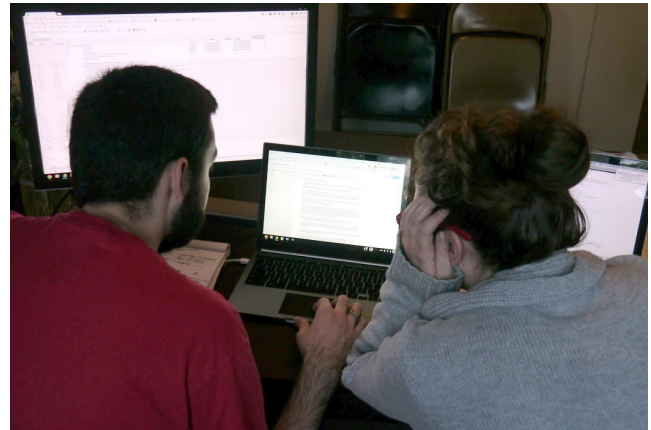


Figure 1. Pair research in action: two group members working together on each other's projects.

it offers people the chance to interact with many partners with diverse perspectives and expertise, and to get to know them better by working closely with them. Reciprocation [5, 12] is built into the process, so neither partner needs to feel in debt to the other afterwards. For learning, it may spread knowledge and skills within a group, and between groups if multiple groups participate in the same pairing system. For productivity, it helps people make concrete progress on their own projects, by drawing on other people's expertise, effort, and social pressure. We chose research groups as our context because—while convenient—they exemplify a kind of creative work where the benefits of pair programming have not been explored.

Pair research is a new socio-technical system consisting of both a social process, meeting to work as a pair, and a technical system that supports the process, collecting people's preferences and automatically matching them to help each other do work. We developed a spreadsheet prototype to manage pair research. Each week, the system automatically reminds group members to submit what they want help on and what they can help with, makes an optimal pairing given that information, and notifies members of their partners.

This paper makes the following contributions: (1) the design space for pair research, drawing from literature, existing systems, and early prototyping experience; (2) a pair research system which collects needs and preferences and matches members automatically; and (3) the results of two deployments, one with roughly 10 people over 2 months, the other with almost 30 people over 6 months, still ongoing. The de-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CSCW'14, February 15 - 19 2014, Baltimore, MD, USA
Copyright is held by authors. Publication rights licensed to ACM.
ACM 978-1-4503-2540-0/14/02\$15.00.
<http://dx.doi.org/10.1145/2531602.2531703>

ployments show that members used pair research in a wide variety of ways, and that it helped them start and complete tasks, share expertise, and learn about other group members.

RELATED WORK

Pair research is inspired in part by pair programming [2], in which two programmers work at the same workstation, with one writing code while the other watches and reviews the code as it is typed. Pair programming has been shown in some studies to be more efficient than solo programming, producing higher-quality code. The effect of pair programming on learning is less clear. One controlled study of introductory computer science showed no benefit for exam scores, but an improvement in attitude toward the course [20]. The two partners in pair programming are typically working on the same development project, sharing a common overall goal. Pair research differs in that its two partners are usually working on different projects.

In pair programming, pairs typically spend a full iteration of the software project working together, which may last for weeks. Belshee has argued for *promiscuous pair programming* instead [3]. Experimenting with pair-switching at various intervals from every 30 minutes to every 3 days, he found the highest task completion rate for 90-minute switching intervals. At this rate, pairs typically consisted of one person already familiar with the task, and one person who was still learning. Frequent switching seemed to propagate knowledge quickly around the development team. By switching partners every week, pair research similarly seeks to spread knowledge and skills quickly within a group.

Pair research is intended to increase informal workplace learning, or *over-the shoulder learning (OTSL)* [18]. One of the benefits of OTSL over more formal learning approaches is shared context. The helper understands the task the learner is trying to do and the learner's goals for doing it. Twidale also argues that encouraging OTSL may require not only tool support but also organizational changes that assign positive value to giving help to colleagues. Pair research as a practice may help to induce these changes.

Also related is apprentice learning [4]. Berlin and Jeffries followed several graduate student interns and their mentors in a computer science research laboratory for several months, a similar setting and duration as the results reported in this paper. One finding was that incidental learning, triggered by specific incidents in problem-solving, was a major benefit of interactions between apprentices and mentors. But apprentices also had to use strategies to minimize their use of the mentor's time. By matching people who can best help one another and rotating pairs frequently, pair research seeks to provide some of the benefits of apprenticeship learning without reliance on a particular mentor.

DESIGN SPACE

In designing the pair research process, we faced a number of design decisions. This section discusses these decisions and some alternatives. Some decisions were based on prior work, and others learned from experience of early design iterations.

How many people should work together? We use pairs by extension of pair programming, and for simplicity. We briefly experimented with 3-person teams as a way to resolve the problem of an odd-numbered pool. The small teams were harder to schedule, took longer to work together, and could not always manage a fair three-way exchange of effort. This finding is consistent with decades of small group research [9]. The current process simply leaves an odd person unpaired for the week. Larger groups have occasionally been suggested. For example, one week a member noticed several people asking for help with learning Node.js, so a study group formed independently of the pair research process.

How should people be matched? Initial experiments used randomized matching, which often led to people tailoring the pair research session to their partner's expertise rather than their needs. To support people receiving help from those who are most able to help, our current prototype solicits preferences from members and computes an optimal pairing given those preferences.

Should preferences be expressed about activities or people? An *activity* is the particular work to be done or kind of help needed. The people are *consumers* who need help and *providers* who give help. Considering other systems in the people-matching design space, online dating sites typically focus on people, but some (e.g. HowAboutWe.com) allow members to propose dating activities and have other members rate that activity. In contrast, marketplaces typically focus on the products or services for sale, but some (e.g. AngiesList) rate the service providers. Our system focuses on activities rather than people, in order to make pair research primarily about work progress rather than social interaction. Members nevertheless used preferences in complex ways, to specify interest in both the activity and the other person.

Who should specify the activities? An activity might be proposed by the consumer, e.g. "I need help with R", or by the provider, like "I can help anybody with design sketching." We chose to focus on consumer-specified activities in order to make productivity a primary goal. A project to-do item then becomes a natural activity. If we had wanted to prioritize learning instead, we might use provider-specified activities, as does Skillshare.com, or for that matter, traditional classroom education. Although consumer-specified help requests were the norm, we also saw provider-specified activities, mainly by one group member who preferred not to ask for help but offered it instead.

How public should the matching process be? Labor markets and dating sites usually keep preferences and matchings private. We chose instead to make them visible to all members of the system, in order to make it easier to learn about the skills and interests of others.

How long should a pair work? We want each person to walk away from pair research with the feeling that they have accomplished something in the time they worked together. But we also want to keep the time commitment low to encourage participation and to avoid fatigue. Based on the Pomodoro technique [6], our normal minimum is 30 minutes for each

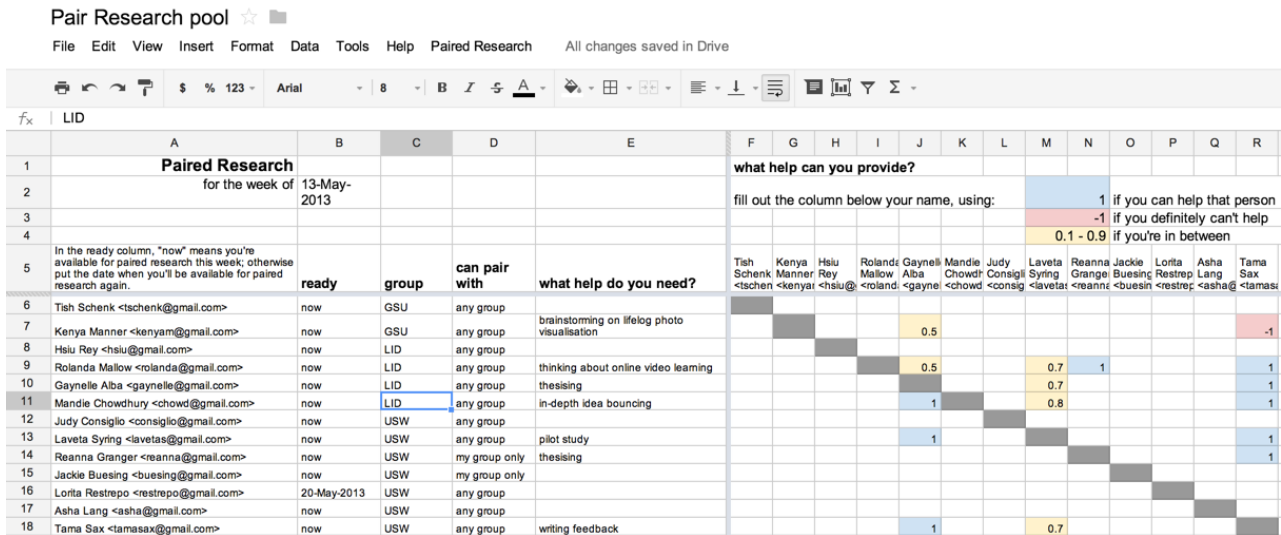


Figure 2. Prototype interface for managing the pair research pool. Each week, members describe what they need help with, and enter preference scores on how well they can help others. Members can also remove themselves from the pool temporarily.

partner. This includes a few minutes to talk informally and off-topic, which helps establish the trust and personal connection critical for learning. With each partner taking a turn, the normal result is a one-hour session. In practice, the duration is negotiated by each pair based on the goals they set for themselves, so some pairs meet for longer than an hour.

How often should pairs change? A secondary objective to matching people with those who are best able to help is to spread knowledge and skills. To promote this objective, we bias the matching to encourage new collaborations and discourage repeating recent pairings.

Should pair meetings be fixed-time or independently-scheduled? We initially prototyped the process at a regularly-scheduled group meeting time, but members requested more flexibility, so pairs now decide on their own when to meet. This flexibility also allowed people from other time zones to join the pool without having to negotiate a common meeting time for the whole pool. In practice, however, flexibility can lead to scheduling failures, discussed more later in this paper.

What pool of people should be involved? A large number of members in the pool provides a more diverse set of expertise, but may be uneven in quality. Our two deployments were centered on existing work groups, and varied in size between 10 and 30 members. Both deployments also included members with a range of seniority, from undergraduates to faculty.

USER INTERFACE

We developed a prototype user interface to manage the process, implemented as a collaborative spreadsheet (Figure 2). The heart of the user interface is a preference matrix for person matching. Each week, group members specify the help they need that week (e.g. “debugging Django”, “feedback on my writing”, “need a pilot user for my prototype”). Other members then fill out the preference matrix according to how able and interested they are to provide that help, where 1 is maximum interest, -1 is maximum disinterest, and 0 or blank is neutral. The preference matrix may be sparse, because

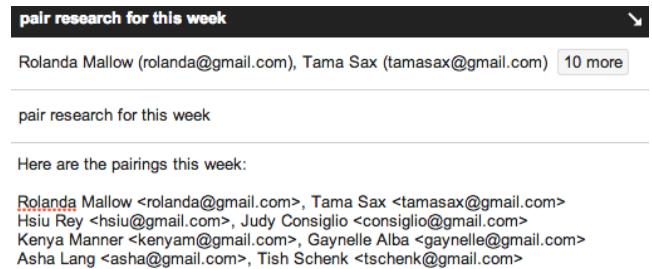


Figure 3. The pairing is computed and emailed automatically.

some members may not fill it out. Members are matched with others even if they do not fill out the matrix.

The preference ratings are used to bias the pair assignment for the week, so that people with high mutual preference are very likely to be paired, people with negative mutual preference are never paired, and people with neutral preference receive a random partner. To generate pairings, the system uses collected preferences to construct a weighted graph. Nodes in the graph represent members in the pool, and the graph contains an edge between two members if and only if both members have nonnegative preferences. The weight of an edge between two members is the average of their mutual preferences, plus a bonus if the pair has not been matched recently. A small random perturbation is added to break ties, and to cause random assignment for members who did not provide any preferences. The system then finds a *maximum weighted matching* on the constructed graph. Matching problems like this one are well-studied in the literature [14], and have applications in matching students to school [1], residents to hospitals [15], and donors to patients [16].

The system sends weekly email notifications to remind members to submit their needs and preferences. Members enter their needs in the spreadsheet on Saturday, then return to the spreadsheet again on Sunday to look at what others need and fill out the preference matrix. The pairs are assigned on Sunday night and emailed automatically. Figure 3 shows an example of a typical matching.

| | Group A | Group B |
|---------------|---------------------------------|--------------------------|
| <i>N</i> | 29 | 10 |
| gender | 18M, 11W | 7M, 3W |
| age | 20–40 | 20–33 |
| role | 6U, 16G, 5P, 2F | 9G, 1F |
| time period | 26 weeks (Nov '12 – May '13) | 7 weeks (Jan-Feb '13) |
| pairings | 244 | 31 |
| help requests | 207 (100%) | 46 (100%) |
| programming | 83 (40%) | 4 (9%) |
| user study | 41 (20%) | 7 (15%) |
| writing | 36 (17%) | 15 (33%) |
| brainstorming | 18 (9%) | 8 (17%) |
| data analysis | 7 (3%) | 8 (17%) |
| offering help | 18 (9%) | 2 (4%) |
| other | 4 (2%) | 2 (4%) |

Figure 4. Results of two deployments of pair research. Roles are (U)ndergraduate, (G)rad student, (P)ostdoc, and (F)aculty.

New columns were added to the spreadsheet as the process evolved. The *ready* column allows a member to temporarily drop out of the pool, for example while traveling or going away for holidays, but requires a date on which they will return, so that the reminder emails can automatically resume. The *group* and *can pair with* columns were added when the pool broadened to multiple research groups, so that a member can keep their pairing local if they choose. Undergraduates, in particular, expressed a reluctance to engage with members of remote research groups, perhaps because they do not (yet) feel ready to make a long-term investment in getting to know people in the wider research field.

The pair research spreadsheet is a low-fidelity prototype [19] of what may become a full-fledged web application. In general, we have found that collaborative spreadsheets and documents are excellent for this purpose, similar to paper prototyping [13] but for collaborative sociotechnical systems that might otherwise require programming to experiment with. Even though our spreadsheet now has a programmed backend that matches people and sends emails automatically, we could have run these parts by hand using a simple but nonoptimal greedy matching algorithm executed by a human wizard. Low-fidelity spreadsheets and docs permit early experimentation with potentially-risky social parts of a sociotechnical system, without investing development effort in the technical parts before it can be justified.

DEPLOYMENT STUDY

Pair research has been deployed in two university research groups, A and B, each run by one of the authors. Group A initially had 15 members at a single school, and the pool grew to include 29 people, including collaborators from five other schools. Group B had 10 members at a single school. Figure 4 shows statistics about the two deployments, including demographics of group members.

Method

We collected and analyzed three kinds of data about the deployments. First, to learn about the pair assignment process, we collected the spreadsheets containing help requests, preference ratings, and the resulting pairs for each week. The help requests were coded by one researcher using the categories shown in Figure 4.

Second, in order to learn about how pairs actually worked together, we sent a short weekly survey to group A for the last 9 weeks of its deployment. Group B did not receive the survey, because its deployment had already ended. The survey asked for how long pairs spent working together, what they actually worked on, the usefulness of the pairing session on a scale from 1 to 5, and open-ended comments on positive and negative aspects of the week’s experience. The survey produced responses for 55 of the 244 group A pairings, from 18 distinct members.

Third, to gather some experience from group B in the absence of survey data, we conducted in-person or video interviews with two members of group B and a remote member of group A, covering the same questions as the survey and also asking about collaboration outside of pair research and about opportunities for informal learning.

Results

Members requested specific help in the spreadsheet only some of the time. In group A, the median person asked for specific help only 20% of the time (5 weeks out of 26), and the most frequent asked for help 70% of the time (18 weeks out of 26). The system assigns partners even without a specific help request, so group A members regarded it as optional. Group B had far more specific help requests, because the spreadsheet was filled out synchronously in meetings to encourage participation. The median group B member made a specific help request 70% of the time (5 weeks out of 7).

Figure 4 shows the kinds of help requested, which varied between the groups. The most common Group A request was programming and debugging help (40%), while the most common request in B was for writing (33%). Both groups showed significant requests for user study design and testing (20% for A, 15% for B). We also saw instances where the specific-help field of the spreadsheet was used to make an offer of help instead of a request. One group A member was responsible for all 18 of these offers, suggesting that he preferred to give help than get it.

Pairing was promiscuous. Group A members had a mean 16.8 pairings over 26 weeks, with a mean 12.0 different partners, so 71% of the time they were working with a partner they had never had before. For contrast, most of group A’s members would normally work closely with only 1-2 other group members, counting the adviser. Pair research increased their number of close contacts by an order of magnitude.

Turning to the survey data, which collected the outcomes of a sample of pairings in group A, the mean time spent on pair work was 1.3 hours per week. But pair work failed to happen 16% of the time (9 instances out of 55 pairings in the survey) because of scheduling difficulties between the partners, suggesting that the practice might benefit from a regular scheduled time, at least for co-located groups. When pairs succeeded in meeting, group A members rated the usefulness of the experience as a mean 4.5 on a scale of 1 to 5, though these ratings may be biased by awareness that other group members might see them.

Comments in both the survey and interviews expressed the value of having a fresh perspective, different from the person's usual set of collaborators: *"It brings together people who are interested in each others' work but wouldn't otherwise act on that interest."*

Many members said the process helped spur them to action. *"It helped me do something that I was ambivalent about." "It is a 'forcing function' ... to work on things I have been putting off." "I [said] 'this week I need help with this and committed to do it."*

Others suggested that more preparation would have been useful before the pair work session. *"This could have been improved had we specifically written down our goals prior to the start."* The time gap between filling out the spreadsheet and meeting to work together also came up as a problem: *"I often find it restricting to 'commit' to working on a certain subject ...up to a week before I actually work on it." "Lately I have not known [what I want to do] far enough in advance."*

Pair research was frequently cited as a way to learn about other group members and their work. A new member of group A expressed this very strongly: *'Our group's paired research is the best thing that happened to me since I came - [getting] to know individual's work one by one, face-to-face.'* Members of group B found this value in the spreadsheet alone, perhaps because they filled it out more than group A: *"Even though we're in the same lab, I don't know what specific tasks they're working on." "It's a good way to know who's doing what."*

Some members learned new tools or practices from working closely with their partners. *"I switched to Sublime because of pair research." "When I worked with [my partner] she decided to drop what she was planning to do for paired research and learn about bootstrap (we had talked about it when I showed her my experiment during paired research)."*

The fair exchange of the process was important to several participants. *"[I like] that we both got something out of it." "She helped me, but I feel a little bit guilty because I couldn't help (but it was her fault because she wasn't ready)." "I wish we had not spent as much time on my stuff and had spent more equal time on his."*

Finally, a number of comments showed that the process provided emotional and social support. *"I was able to help [him] on his thesis when he was stressed." "I felt energized and excited about my work after every pair research."*

DISCUSSION

The goals of the pair research system were to improve productivity, informal learning, and collaboration. For productivity, the results suggest that members feel motivated by the process, and do succeed in making progress on their work. For collaboration, members interact with far more people than they normally would, which increases opportunities for further collaboration. For learning, we saw some mentions of specific tools, skills, and work practices, and also learning about other group members and their work and their abilities, which again may lay the groundwork for future collaboration.

It is instructive to look at the differences between the group A and group B deployments, because each had different successes and failures that shed light on how the system could be improved. Group B used the spreadsheet far more effectively than group A because they didn't just rely on email reminders over Saturday and Sunday, but filled out the spreadsheet together at a Friday group meeting. Another contrast was the length of deployment. Group A continued using pair research throughout the spring semester, but group B stopped partly because the group leader is a bottleneck in the current system, having to approve the pairing before emailing it out to the group. On one occasion, the pairing wasn't announced until Thursday, leaving only two days left in the week to actually work. We have since changed the system to send the final pairing automatically.

Even the pair work session may need to happen more automatically. Both group A and group B struggled with fitting pair research into already-busy schedules. Group A experienced 16% failures due to scheduling, and group B stopped after 7 weeks in part because the entire group got busy, not just the group leader. One way to address this would make the entire process happen in one synchronous session, in which members fill out the spreadsheet, immediately determine the pairing, and then immediately break up to work in pairs.

Prior to deployment, it may be important to understand how pair research integrates in the organizational context. Group B may have stopped after 7 weeks in part because the team's reward structure may have recognized and reinforced individual excellence as opposed to team collaboration and learning [8]. Group B's membership may have been unclear [11] or not stable enough to learn and adopt a new collaborative practice [7]. Moreover, Group B's members could have relied on outside resources or support or felt that they were too similar to each other to support each other. Member personality types, such as openness and extraversion [10], and ability to communicate may have also influenced adoption. Assessing the organizational context prior is one way to address this concern.

One co-author's research group, inspired by pair research but not using the spreadsheet matching system, combined group meetings and one-on-one meetings into a single synchronous time block. The group meeting starts with brief updates and requests for ad-hoc meetings with other group members (which are recorded on a whiteboard), and then breaks up into those ad-hoc meetings, which continue until everybody's meeting needs are satisfied. Conducting pair research synchronously may be the right solution for a co-located group, though as groups grow and become geographically distributed, like group A, it becomes more challenging to bring the entire group together at the same time.

Limitations

To design and develop pair research, we deployed the system in two technical labs within universities characterized by high research activity. While the diversity and size of the sample was limited, like other systems designers [21] we found that small deployments are useful for developing a proof of concept and identifying social and technical factors to consider in

future designs. Further, we conducted observations throughout the deployment. The advantage to this research approach is the ability to collect in situ data, not just reflective data; the disadvantage is that bias is introduced through participant observation [17].

Future Work

Based on the initial outcomes, we will continue to deploy the system to better understand the social (i.e. size, gender) and technical (i.e. matching algorithm, capturing shared content) factors in academic research as well other domains such as public policy research in government and curriculum development in education. In addition, we understand that pair research likely does not come without costs. We will investigate whether confusion about the ownership of ideas occurs after meetings, whether the method prevents self-directed learning or novices from gaining sufficient expertise to advise others, or whether the method encourages social comparison.

CONCLUSION

This paper presented pair research, a new socio-technical system that pairs members of a work group to work together on each other's projects each week. In two deployments, we found that the system motivates participants, helps them learn about their colleagues, and makes opportunities for further collaboration.

While we have been motivated by the academic research setting, the pairing system may find application to groups in other domains, particularly other kinds of hard knowledge work that involves diverse expertise. In particular, we find pair research to be effective as a form of transient, low-commitment collaboration, especially to overcome hurdles and promote progress on hard tasks by drawing on other people's expertise, effort, and social pressure. We look forward to expanding pair research to other research groups and experimenting in other work settings in future work.

ACKNOWLEDGMENTS

We gratefully acknowledge valuable help and suggestions from Katrina Panovich, Greg Vargas, Geza Kovacs, Hubert Pham, Elena Agapie, the anonymous reviewers, and all the wonderful people in the MIT UID group, the Georgia Tech comp.social lab, and the Northwestern Creative Action Lab. This work was supported in part by the National Science Foundation under SOCS-1111124. Any opinions, findings, conclusions, or recommendations in this paper are the authors' and do not necessarily reflect the views of the sponsors.

REFERENCES

1. Abdulkadirolu, A., Pathak, P. A., and Roth, A. E. The New York City high school match. *American Economic Review* 95, 2 (May 2005), 364–367.
2. Beck, K. *Extreme programming explained: embrace change*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2000.
3. Belshee, A. Promiscuous pairing and beginner's mind: Embrace inexperience. In *Proc. Agile Development Conference, ADC '05* (2005), 125–131.
4. Berlin, L., and Jeffries, R. Consultants and apprentices: observations about learning and collaborative problem solving. In *Proc CSCW* (1992), 130–137.
5. Cialdini, R. B. *Influence: the psychology of persuasion*. HarperCollins, 2009.
6. Cirillo, F. *The pomodoro technique*. Lulu.com, 2009.
7. Cohen, W. M., and Levinthal, D. A. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* (1990), 128–152.
8. Edmonson, A. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly* 44, 2 (1999), 350–383.
9. Levine, J. M., and Moreland, R. L. Progress in small group research. *Annual Review of Psychology* 41, 1 (1990), 585–634.
10. McCrae, R. R., and John, O. P. An introduction to the fivefactor model and its applications. *Journal of Personality* 60, 2 (1992), 175–215.
11. Mortensen, M., and Hinds, P. Fuzzy teams: Boundary disagreement in distributed and collocated teams. *Distributed Work* (2002), 284–308.
12. Regan, D. T. Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology* 7, 6 (1971), 627–639.
13. Rettig, M. Prototyping for tiny fingers. *Commun. ACM* 37, 4 (Apr. 1994), 21–27.
14. Roth, A., and Sotomayor, M. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Econometric Society Monographs. Cambridge University Press, 1992.
15. Roth, A. E. The evolution of the labor market for medical interns and residents: A case study in game theory. *Journal of Political Economy* 92, 6 (1984), 991–1016.
16. Roth, A. E., and et al. Kidney exchange. *Quarterly Journal of Economics* 119, 2 (May 2004), 457–488.
17. Spradley, J. P. *Participant Observation*. Holt, Rinehart and Winston, 1980.
18. Twidale, M. B. Over the shoulder learning: Supporting brief informal learning. *Comput. Supported Coop. Work* 14, 6 (Dec. 2005), 505–547.
19. Virzi, R. A., Sokolov, J. L., and Karis, D. Usability problem identification using both low- and high-fidelity prototypes. In *Proc. CHI* (1996), 236–243.
20. Williams, L., and et al. In support of pair programming in the introductory computer science course. In *OOPSLA* (2002), 197–212.
21. Zimmerman, J., Forlizzi, J., and Evenson, S. Research through design as a method for interaction design research in HCI. In *Proc. CHI*, ACM (New York, NY, USA, 2007), 493–502.