

Interpersonal Coordination of Perception and Memory in Real-Time Online Social Experiments

Alexandra Paxton

paxton.alexandra@gmail.com
Institute of Cognitive and Brain Sciences
Berkeley Institute for Data Science
University of California, Berkeley

Jordan W. Suchow

suchow@berkeley.edu
Social Science Matrix
University of California, Berkeley

Thomas J. H. Morgan

thomas.j.h.morgan@asu.edu
School of Human Evolution and Social Change
Arizona State University

Thomas L. Griffiths

tom_griffiths@berkeley.edu
Department of Psychology
University of California, Berkeley

Abstract

The subtle hum of interpersonal coordination that runs throughout social communication and interaction shows how individuals can subtly influence one another's behaviors, thoughts, and emotions over time. While the majority of research on coordination has been done on face-to-face interaction, recent advances in crowdsourcing have opened the opportunity to conduct large-scale, real-time social interaction experiments. We take advantage of these tools to explore the interpersonal coordination in a "minimally interactive context," distilling the richness of natural communication into a tightly controlled setting to explore how people become coupled in their perceptual and memory systems while performing a task together. Consistent with previous work on postural sway and gaze, we find that individuals become coupled to one another's cognitive processes without needing to be co-located or fully interactive with their partner; interestingly—although participants had no information about nor direct communication with their partner—we also find hints of the context-shaping effects that social forces can play in face-to-face interaction.

Keywords: interpersonal coordination; human communication; online experiments; social interaction

Introduction

Research on the phenomenon of *interpersonal coordination* focuses on the subtle ways in which our interactions with others directly affect our own behaviors, feelings, and thoughts. Interest in coordination (also known as interactive alignment, interpersonal synchrony, mimicry, and more; see Paxton, Dale, & Richardson, 2016) has surged over the last several decades as a framework for understanding how contact with others shapes our cognition and behavior, with much of it focusing on how we become more similar over time during task-oriented or friendly contexts.

A growing perspective in this area has taken inspiration from dynamical systems theory, conceptualizing interaction as a complex adaptive system from which coordination arises as an emergent phenomenon according to contextual pressures (Riley, Richardson, Shockley, & Ramenzoni, 2011). A fundamental principle of this dynamical systems perspective holds that coordination should not be static across contexts nor over time. Exploring how new contexts and contextual demands—like interpersonal conflict (Paxton & Dale, 2013), friendly competition (Tschacher, Rees, & Ramseyer,

2014), or specialized task demands (Fusaroli et al., 2012)—has become a central part of this perspective, laying out under what conditions coordination disappears, increases, or demonstrates complementary rather than synchronous in-phase patterns.

There is similar interest in comparing how coordination changes across different behavioral or cognitive systems. Under the dynamical systems perspective, a context's unique pressures, the resulting coordination dynamics, and their impact on the interaction may differ over time and across settings. For example, some of the earliest work in this subset of coordination research has found that—during task-related interaction—individuals tend to become more similar over time across a variety of metrics (Louwerse, Dale, Bard, & Jeuniaux, 2012) but that specific kinds of coordination can differentially help or hurt outcomes (Fusaroli et al., 2012).

Broadly, during tasks that are neutral (Shockley, Santana, & Fowler, 2003), cooperative (Louwerse et al., 2012), or competitive (but not conflict-driven; e.g., competitive games, Tschacher et al., 2014), previous work broadly suggests that individuals' behavior and cognition become more similar over time. A range of behavioral signals—both high-level (e.g., gesture; Louwerse et al., 2012) and low-level (e.g., postural sway; Shockley et al., 2003)—become synchronized, even when the interacting individuals are unable to see one another (Shockley et al., 2003) or are separated in time (Richardson & Dale, 2005).

The systematic testing of coordination across a variety of interaction contexts is vital to charting its dynamical landscape. This methodical exploration of different factors will eventually enable us to identify control parameters and key factors of initial conditions that shape how coordination emerges and how it impacts interaction outcomes. Doing so, however, requires researchers on interpersonal coordination to expand our view of experimental paradigms: Even as we continue to embrace more complex naturalistic interactions (e.g., Paxton & Dale, 2013; Tschacher et al., 2014), we must continue developing experimental methods for analyzing "minimally interactive contexts" (Hale, Pan, & Hamilton, 2015)—that is, situations in which our interactions with oth-

ers are limited in behavioral channel, scope, or time—to fully map the interaction space.

In the search for both fully interactive and minimally interactive paradigms, we believe that online experiment platforms and crowdsourcing can serve as powerful tools in experiment development and deployment. By connecting people digitally, researchers are able to more fully control the experimental experience, from deciding how much social information about their partner(s) will be available to establishing what communication channel(s) can be utilized to crafting interactive studies for groups beyond the dyad. We propose that researchers interested in social cognition and behavior take advantage of the rapid advances in crowdsourcing platforms to develop real-time experiment models.

Crowdsourcing for Interactive, Interpersonal Social Experiments

Crowdsourcing platforms like Amazon Mechanical Turk (<http://www.mturk.com>) have been extensively used as a means to collect data with relatively simple but robust experimental paradigms, like surveys (Buhrmester, Kwang, & Gosling, 2011) and mouse-tracking (Freeman, Dale, & Farmer, 2011). At first, work in this domain required researchers to use survey creation tools or to program bespoke experiments, but more recently, cognitive scientists have worked to create solutions to support the efficient creation of a wider range of experiments (e.g., Gureckis et al., 2016).

To date, many of these experiment platforms have focused on isolated individuals, making it difficult for researchers to study situated social processes. We believe that the next step in online experimentation, then, is to move to models of real-time social interaction, creating interactive or interdependent experimental paradigms that construct interacting dyads (or even larger networks of people) to understand social processes and phenomena. In doing so, real-time interactive paradigms deployed on crowdsourcing platforms can provide researchers interested in social behavior the opportunity to expand their experimental capabilities beyond the lab while not compromising on the richness and complexity of true interactive contexts.

A recurring concern for using online experiments lies in its participant population. Like all convenience samples, there can be questions about the degree to which the participants reflect the broader population dynamics—including the use of undergraduate students at Western universities as participants in return for course credit, who often do not reflect global demographics (Henrich, Heine, & Norenzayan, 2010). Considerations of sampling and population representativeness are vital for any study, and researchers should carefully consider their sampling choices at the outset of their work. For those interested in using online participants (especially from Amazon Mechanical Turk), recent surveys suggest that U.S.-based MTurk workers are more diverse in a variety of ways than typical college students but not entirely reflective of the general U.S. population (e.g., Buhrmester et al., 2011; Paolacci

& Chandler, 2014).

The Present Study

Here, we build on a robust tradition within interpersonal coordination that demonstrates that people become coordinated across a number of behavioral channels, even when they have very little access to one another. In this case, we focus on task performance within a minimally interactive context through a real-time cooperative online experiment—a nominal game that asks players to correctly perceive and (under cognitive load) remember a line. Specifically, the current study focuses on understanding how interacting individuals become entrained in perception and memory over time, becoming a sort of “line estimation system” (cf. Dale, Richardson, & Kirkham, 2011).

We approach the current study with three main research questions. First, we ask whether people become more coupled in their perceptual and memory systems over time, despite the limited perceptual and social information about their partner. In keeping with previous research, we expect to find that individuals will start to exhibit similar behavioral dynamics as they continue to play with one another.

Next, we investigate whether any observed coordination effects could simply be an artifact of the joint learning context, and we hypothesize that coordination dynamics will still be a significant influence on players’ cognitive and behavioral systems.

Finally, we look to whether any social factors might influence these dynamics, even in this minimal context. Social factors like rapport and affect play vital roles in face-to-face interaction during naturalistic experiments (e.g., Tschacher et al., 2014). Although the game does not facilitate any explicit social behaviors, we are interested to explore whether some social influences surface as emergent effects, given their importance to coordination research.

Method

All research activities were completed in compliance with oversight from Committee for the Protection of Human Subjects at the University of California, Berkeley.

Participants

Participants ($n = 148$) were individually recruited from Amazon Mechanical Turk to participate as dyads ($n = 74$). Participants were paired with one another according to the order in which they began the experiment. All participants were over 18 years of age and fluent English speakers (self-reported); participation was restricted to only recruit from participants located within the U.S. with a 95% HIT approval rate.¹

The experiment lasted an average of 11.69 minutes (range: 7.98–21.34 minutes). All participants were paid \$1.33 as base pay for finishing the experiment and earned a bonus of up to \$2 for the entire experiment based on mean accuracy

¹A measure of MTurk worker quality, capturing how often their work is rejected by a requester. A 95% HIT approval rate means that only 5% of all of their submitted HITs have been rejected.



Figure 1: Experiment flow

over all trials (mean = \$1.80; range: \$0.00—\$1.95). Participants were not aware of the value of their earned bonus until after the entire experiment was completed.

Procedure

All data collection procedures were run on Amazon Mechanical Turk (<http://mturk.com>) using the experiment platform Dallinger (v3.4.1; <http://github.com/dallinger/Dallinger>). Code for the experiment is available on GitHub (<http://github.com/thomasmorgan/joint-estimation-game>), and the resulting experiment data are available on the OSF repository for the project (<https://osf.io/8fu7x/>).

Each participant was individually recruited on Amazon Mechanical Turk to play a “Line Estimation Memory Game” (advertisement: “Test your memory skills!”; see Figure 1 for experiment flow). Upon completing informed consent, participants were told that they would be playing a game in which they would be required to remember and re-create line lengths. Participants were informed that they would be completing their training trials individually and would then begin playing with a partner. Participants were given no information about their partner other than the guess that their partner made; no information about the partner’s identity was shared.

In each trial, participants were shown 3 red lines, each of a different length, and were asked to remember all three of them.² The 3 stimulus lines were displayed for 2 seconds then removed, providing participants with a blank screen for 0.5 seconds. Participants were then told which line to re-create (#1, #2, or #3) and were then given 1 second to submit

their guess at how long the target line had been. To do so, participants were given a blank box and used their cursor to fill in the box with a blue line. All lines were presented within bounded boxes of 500 pixels (wide) by 25 pixels (high).

During training, participants were then shown the correct length of the target line (as a grey bar above their own guess) for 2 seconds. This was accompanied by a message telling the participant that they had guessed correctly (“Your guess was correct!”) or incorrectly (“Your guess was incorrect”) or that they had not submitted a guess within the 1-second time limit (“You didn’t respond in time”).

During testing, participants’ stimulus viewing, waiting, and recreation times remained the same as during training, but they no longer received information about whether their guess was correct or incorrect. Instead, after both participants had submitted their first guess, participants were shown their guess (in blue) above their partner’s guess (in green). Both participants were then asked whether they wanted to change their own guess or to keep the guess they had submitted. If either participant in the dyad indicated that they wanted to change their guess, that participant was then allowed to change their guess (again with a 1-second time limit) *while* still being able to view their partner’s guess. Participants who chose to keep their previously submitted guess was informed that their partner chose to submit a new guess and waited for the other participant to finish. At that point, participants were again allowed to change or keep their guess. This process continued until both participants chose to keep their guess.

Participants were informed that their final accuracy would only be calculated for their final guess. However, because they had no means to communicate with their partner about whether each would be accepting or changing their guesses, each participant could not have known whether their decision to keep the guess would have been their final guess for the

²A pilot version of this study showed that participants performed at ceiling when given only 1 line to remember and recreate. The additional 2 lines were added to strictly increase the memory load, as opposed to adding difficulty in other ways (e.g., creating a moving stimulus).

trial.

For clarity, we will refer to each new stimulus set as a *trial* and to each submitted line length estimate within each trial as a *guess*. This means that some participants may have submitted multiple guesses per trial. The last submitted estimate—the one by which trial-level accuracy is calculated—will be referred to as the *final guess*.

All dyads completed 10 training trials (alone) and 15 test trials (with their partner). All training and test stimuli were randomly generated for each dyad, but both participants within the dyad were given the same stimuli. After participating, each individual participant was asked to complete a series of questionnaires about the game on a series of 1-10 Likert-style scales, including the perceived difficulty of the task, how engaged they were in the task, and questions about their own and their partner's cooperativeness and trustworthiness.

Measures

Error Error was measured as a ratio relative to the total possible error on a given target stimulus trial. That is, rather than taking a given guess's error relative to the total line length, error was calculated as the maximum *possible* error. For example, if the target stimulus was 60 units long, participants could either under-estimate the line length by 60, or over-estimate it by 40. As such, the maximum possible error for that trial would be 60, and the participant's error would be calculated relative to that maximum possible error.

Training Improvement To provide a baseline for participants' improvement during training, we calculated the slope of participants' normalized error over all training trials.

Coordination To measure how participants' perceptual and memory systems became more similar over time, we calculated the cross-correlation coefficient of participants' guess errors across trials (Paxton & Dale, 2013), within a window of ± 5 guesses. Although cross-correlation produces information about leading and following behavior, we have no *a priori* expectations about which of the two participants would emerge as a leader (given they have no information about their partner nor any assigned roles), so our first-pass analyses ignore any directionality by averaging across the each incremental lags (i.e., leading/following in both participants' directions).

Results

All analyses were performed in R (R Core Team, 2016). Linear mixed-effects models were performed using the `lme4` package (Bates, Machler, Bolker, & Walker, 2015) using the maximal random slope structure for each random intercept to achieve model convergence (Barr, Levy, Scheepers, & Tily, 2013). All main and interaction terms were centered and standardized prior to entry in the model, allowing the model estimates to be interpretable as effect sizes (Keith, 2005).

Model 1: Coordinated Error over Time

Our first model tested whether individuals' ratings became more similar over time while accounting for training improvement. We constructed a linear mixed-effects model predicting cross-correlation coefficients of normalized error with absolute lag and training improvement, using maximal random effects structures for experiment and ratings of difficulty after the experiment.

As predicted, we find that dyads were significantly and strongly coupled in their error ratings ($\beta = -0.43$, $p < 0.0001$), with no effect of training improvement ($\beta = -0.06$, $p = 0.36$). In other words, we find that players were more likely to produce lines with similar errors at the same time, even across repeated guesses within a single trial, regardless of how well-adapted they were to the task during training.

Model 2: Learning and Coordinative Processes

However, during the experiment, both participants are learning to play the game at the same time that they are influencing one another. To ensure that the similarity observed in our first model was not simply an artifact of both participants improving individually, we next tested the relation between participants' (1) adaptation to their partner's perceptual estimation and memory and (2) own performance changed over time. If participants were adapting along both avenues, we might find evidence of these dual processes through differences in their rates of adaptation over time. We calculated the former using normalized error ("true error") and the latter as the absolute difference in the participants' guesses ("partner error").

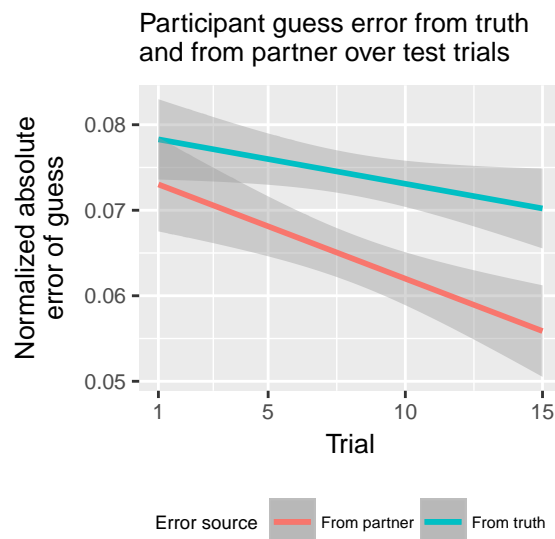


Figure 2: Difference over time in coordinative and learning processes, or the change in guess deviation from truth (in blue) and from their partner's guess (in red) over the game.

To answer this question, we built a linear mixed-effects model predicting cumulative guess number with difference between "true error" and "partner error" and participants' training improvement, with the maximal random effects

structure for participant and difficulty. As expected, we did find that participants' rates of adaptation to their partner significantly differed from their rates of adaptation to the game ($\beta = -0.06$, $p < 0.008$; see Figure 2). In other words, this model suggests that individuals are simultaneously engaging in both learning *and* coordinative processes during the game, becoming attenuated to the learning task while also synchronizing to one another's cognitive processes.

No other predictors reached statistical significance (all $ps > 0.25$).

Model 3: Social Factors in Minimally Interacting Contexts

To explore the role that social judgements can play even in minimally interactive contexts, we finally turned to consider how trust might impact coordination. For this model, we captured a third measure of coordination: the participant's willingness to change their guess.

We created a linear-mixed effects model predicting the total number of times a participant changed their guess during a given trial according to their trust rating of their partner and the difference in participants' first guess on that trial, while controlling for trial number and how much they improved during their own training. Again, the model included the maximal random effects structure for participant and difficulty.

We found that greater trust in their partner predicted a small but statistically significant *increase* in the number of iterations of guesses within a trial ($\beta = 0.07$, $p < 0.042$), although we found no difference in the number of guesses based solely on the differences in the partners' first guesses ($\beta = 0.01$, $p = 0.65$).

Interestingly, we found that participants with greater training improvement took more guesses on test trials when they did better in their training trials ($\beta = 0.05$, $p < 0.028$). We also saw an effect of trial ($\beta = -0.05$, $p < 0.02$), indicating that people changed their guesses fewer times in each trial as the game wore on.

Discussion

Inspired by established lines of research on interpersonal coordination, we here explored how minimally interactive contexts can shape the emergence of interpersonal dynamics. Using an online experiment developed to provide participants with only a single means of communication—namely, their partner's ability to correctly perceive and recall the length of a target line—we found evidence of coordination of cognitive systems despite minimal social information and context.

As expected, we found evidence for low-level perceptual and memory coordination between players throughout the game. Congruent with findings about postural sway (Shockley et al., 2003) and gaze (Richardson & Dale, 2005), the present work suggests that some behavioral and cognitive processes can become coordinated even when separated by time and space, given some access to the relevant process in another person and a task-based interactive context to which

that process is essential. Moreover, learning and game performance of the individual players were unable to completely account for the players' similarity to one another.

Finally, like other work across the coordination research area, we found that participants' coordination behaviors were shaped by subtle social judgements. Within coordination research, rapport has long been upheld as one of the important predictors of coordination (Hove & Risen, 2009). Similarly, we find that players' decisions to change their guesses were influenced by trust in their partner.

Taken together, our findings contribute to the ongoing efforts to understand the form, function, and emergence of coordination. We were specifically interested in pursuing three important questions around interpersonal coordination of perception and memory: whether it emerges during minimally interactive contexts, whether it can be distinguished from other contemporaneous behavioral and cognitive processes, and whether it is influenced by social judgements. In addition to finding support for these hypotheses, we also hope to have provided an example of the utility of real-time interactive experiments and crowdsourcing platforms to investigate core principles of interpersonal coordination and human interaction at a larger scale without relinquishing experimental control.

Future Directions

Some of the questions left open by the present study may provide interesting avenues for future work, both for better understanding some of the effects identified here and for extending them into novel territory.

First, we found that participants were more likely to change their guesses during the game when playing with a partner who they rated trustworthy *after* the game. One plausible explanation for this behavior may be that participants were more willing to change their own guess to be closer to the guess provided by a partner whose perceptual and memory capabilities they trusted. Another (not necessarily exclusive) possibility is that feelings of trust are built throughout the game: Since the ratings were made at the end of the experiment, it may be that the partner's behavior throughout the interaction—*aside* from accuracy (e.g., duration of the interaction)—builds to create that feeling of trust. Future work should explore the degree to which these explanations may (individually or together) explain the emergent social phenomena observed even in minimally interactive contexts.

Second, we found that participants became coordinated even on a relatively short timescale, with the average dyad taking 11.69 minutes to complete the game. Future work could examine hysteresis effects in the individual participant's perceptual and memory systems by pairing two participants from different dyads with one another to see how long it takes before they begin to form new coordination patterns. With most research focusing on the initial creation of coordination, understanding the flexibility of these dissipative interpersonal coordinative structures would provide new insight into this phenomenon.

Finally, given the simplicity of this experimental design and the broad participant pools available through crowdsourcing sites, this work could be expanded to examine other important questions of scale in interpersonal coordination. The majority of research on interpersonal coordination has tended to focus on dyadic interaction, as we did here. While dyadic interaction is both conceptually and statistically more tractable than group dynamics, many real-world social settings include more than two people; therefore, any coherent theory around interaction must be able to account for these settings. It is especially important for coordination researchers from the dynamical systems perspective, as there may be crucial control parameters or vastly different state spaces when considering dyadic dynamics versus larger interactions. Using crowdsourcing and real-time social experiments, researchers are able to control the interaction space much more tightly, enabling the targeted focus on specific processes across a massive potential participant population.

Acknowledgements

Our thanks go to Nelle Varoquaux and to the Dallinger development team for their assistance in debugging the experiment. This work was funded in part by DARPA Cooperative Agreement D17AC00004. This project was funded in part by a Moore-Sloan Data Science Environments Fellowship to AP, thanks to the Gordon and Betty Moore Foundation through Grant GBMF3834 and the Alfred P. Sloan Foundation through Grant 2013-10-27 to the University of California, Berkeley.

References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure in mixed-effects models: Keep it maximal. *Journal of Memory and Language*, 63(3), 255–278.
- Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data?, 6, 3–5.
- Dale, R., Richardson, D. C., & Kirkham, N. K. (2011). *How two people become a tangram recognition system*. Aarhus University, Aarhus, Denmark.
- Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, 2, 59.
- Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tlyn, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, 23(8), 931–939.
- Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., ... Chan, P. (2016). PsiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, 48(3), 829–842.
- Hale, J., Pan, X., & Hamilton, A. F. de C. (2015). Using interactive virtual characters in social neuroscience. In *Virtual reality (vr), 2015 ieee* (pp. 189–190). IEEE.
- Henrich, J., Heine, S., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466(7302), 29.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27(6), 949.
- Keith, T. (2005). *Multiple regression and beyond*. Boston, MA: Pearson.
- Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, 36(8), 1404–1426.
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23(3), 184–188.
- Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology*, 66(11).
- Paxton, A., Dale, R., & Richardson, D. (2016). Social coordination of verbal and nonverbal behaviors. In P. Passos, K. Davids, & C. J. Yi (Eds.), *Interpersonal coordination and performance in social systems* (pp. 259–273). Abington, Oxon; New York, NY: Routledge.
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Richardson, D., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045–1060.
- Riley, M. A., Richardson, M., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in Psychology*, 2, 38.
- Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 326.
- Tschacher, W., Rees, G. M., & Ramseier, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, 5, 1323.