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

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

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

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

Abstract

The quiet 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 studies face-to-face interaction, recent advances in crowdsourcing afford the opportunity to conduct large-scale, real-time social interaction experiments. We take advantage of these tools to explore 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 found 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 their partner and no means of direct communication, we also found hints that social forces can shape minimally interactive contexts, similar to effects observed 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 in 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 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)—change coordination dynamics 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, the unique pressures of a context, the resulting coordination dynamics, and the impact of those dynamics 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), 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 during interaction. This synchronization occurs 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 an expanded view of experimental paradigms: Even as we continue to embrace more complex naturalistic interactions (e.g., Paxton & Dale, 2013; Tschacher et al., 2014), to fully map the interaction space we must also develop experimental methods for analyzing "minimally interactive contexts" (Hale, Pan, & Hamilton, 2015)—

that is, situations in which our interactions with others are limited in behavioral channel, scope, or time.

Online experiment platforms and crowdsourcing can be powerful tools for creating both fully interactive and minimally interactive paradigms. By connecting people digitally, researchers can fully control the experimental experience—deciding how much social information partners will have about one another, establishing which communication channels can be used, and potentially crafting interactive studies for groups beyond the dyad. Crowdsourcing platforms such as Amazon Mechanical Turk (`http://www.mturk.com`) have been extensively used as a means to collect data on individuals (Buhrmester, Kwang, & Gosling, 2011). However, by developing real-time interactive paradigms for these platforms, researchers interested in social behavior can study experimentally situated social processes beyond the lab without compromising the richness and complexity of true interactive contexts.

Does Coordination Emerge in Extraordinarily Minimally Interactive Contexts?

Here, we build on previous findings that people become coordinated across behavioral channels even when they have very little access to each other. Previous work has tended to preserve elements of more typical human interaction—like speech and language—to examine how restricting interaction can influence coordination in other behavioral channels (e.g., gaze coordination or postural sway entrainment; Richardson & Dale, 2005; Shockley et al., 2003). However, understanding the emergence and role of coordination requires us to continue to manipulate social settings, carving out the limits of coordination to identify the processes and constraints that create it.

To do so, 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 remember the length of a line while under cognitive load. Specifically, the current study focuses on understanding how interacting individuals become entrained in perception and memory over time, becoming a "line estimation system" (cf. Dale, Richardson, & Kirkham, 2011). This allows us to continue mapping the course of coordination across cognitive and behavioral systems: Building on a robust tradition on transactional memory and collective cognitive systems (e.g., Tollefsen, Dale, & Paxton, 2013), we explicitly test whether low-level perception and memory processes become more similar through contact with others.

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 limited perceptual and social information about their partner. Next, we investigate whether any observed coordination effects could simply be an artifact of the joint learning context. Finally, we look to whether any social factors (such as rapport and affect, which play vital roles in face-to-face interaction; e.g., Tschacher et al., 2014) might influence these dynamics, despite the minimal context. We are interested to explore whether some social influences surface as emergent effects even though the game does not facilitate any explicit social behaviors.

Method

All research activities were completed in compliance with oversight from the 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 in the order they arrived to experiment. All participants were over 18 years of age and were fluent English speakers (self-reported); recruitment was restricted to participants within the U.S. with a 95% 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 their own mean accuracy over all trials (mean = $1.80; range: $0.00—$1.95). Participants were not aware of the value of their earned bonus until after completing the experiment.

Procedure

Data collection was 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`).

Each participant was individually recruited on Amazon Mechanical Turk to play a "Line Estimation Memory Game" (advertisement: "Test your memory skills!"; see Fig. 1 for experiment flow). Upon completing informed consent, participants were told (1) that they would be playing a game in which they would be required to remember and recreate line lengths; (2) that they would first complete their training trials individually and would then play with a partner; and (3) that they would receive a bonus based on their own accuracy on the final guess of each test trial. Participants were given no information about their partner other than being able to see the guess that their partner made during test trials.

On each trial (i.e., each new stimulus set), participants were shown 3 red lines, each of a different length, and were asked to remember their lengths.[2] The three lines were left-aligned within a 500x25px box and were displayed for 2 s, followed by a blank screen for 0.5 s. Participants were then provided with an empty 500x25px box and given 1 s to recreate the

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

[2] In a pilot study, participants performed at ceiling when given only 1 line to remember and recreate. Two more lines were added to increase the memory load.

Figure 1: Experiment flow

length of the target line (#1, #2, or #3). Participants made decisions by positioning their cursor over the box at their estimate of the rightmost extremity of the line and clicking.

During training, participants were given feedback in the form of the true length of the target line (as a grey bar above their own guess) for 2 s. This was accompanied by a message telling the participant that they had guessed correctly (i.e., within 4px of the true line length; "Your guess was correct!") or incorrectly ("Your guess was incorrect") or that they had not submitted a guess within the 1-s time limit ("You didn't respond in time").

During testing, participants no longer received feedback their accuracy. Instead, after both participants had submitted their first guess, they were shown their guess placed above their partner's guess (see Fig. 1) and were asked whether they wanted to change their own guess. Each participant could individually change their own guess (again with a 1-s time limit) while seeing their partner's previous guess; participants were not informed of their partner's decision (to keep or change their guess) until after both participants had answered (and, if needed, changed their guess). Each trial ended when both participants were satisfied with their guess.

Participants were informed that their final accuracy bonus would only be calculated using 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 trial. As a result, our statistical models use all guesses, not just final guesses (see next section for more detail).

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. Stimuli were drawn from a uniform distribution between 1% and 100% (inclusive) of the total possible line length; this could have, by chance, resulted in some relatively easier stimulus sets for some dyads, which should be mitigated by our sample size. After participating, each participant completed 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 and Model Specifications

For clarity, we present the measures and model specifications together. Each measure used in one of our three models model is defined and written in bold the first time it is presented in this section.

Model 1 Specifications: Do Partners' Perception and Memory Couple in Minimally Interacting Contexts? Our first model tested our hypothesis that individuals' ratings would became more similar over time. To do that, we first calculated each participant's **error** for *each guess* of *each trial*. 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. We chose to use normalized error—rather than absolute error—as a measure of performance that natively controlled for the "possible wrongness" associated with any given line.

We then quantified perceptual and memory **coordination** (or how similar participants' perceptual and memory systems became over time) as the cross-correlation coefficient of participants' error. Cross-correlation—a common measure of coordination (Paxton & Dale, 2013)—was calculated using all guesses across all trials 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). Our first-pass analyses therefore ignore any directionality by incorporating **absolute lag**, averaging across the correlation value for each absolute lag (i.e., leading/following in both participants' directions).

To provide a baseline measure of **training improvement**, we calculated the slope of each participant's normalized error over all training trials. To account for individual differences in self-assessed task difficulty, we used ordinal **ratings of difficulty** that each participant gave after the task.

Our first model was a linear mixed-effects model predicting coordination of normalized error with absolute lag and training improvement as fixed effects, using dyad and difficulty ratings as random effects.

Model 2 Specifications: Can We Identify Signatures of Learning *and* Coordinative Processes? During the experiment, both participants are not simply influencing one another (as tested in Model 1)—but are also simultaneously learning to play the game. To ensure any similarity found by Model 1 would not be simply an artifact of both participants improving individually, we 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 could find evidence of these dual processes through differences in their rates of adaptation over time.

To do this, we used the normalized error values (described above) to derive two measures. To answer the latter point, we used each participant's normalized error for *each guess* in *each trial* as their **true error**—in other words, how much the participant differed from the stimulus. To answer the former, we calculated the absolute difference between both participant's true error to obtain the **partner error** for *each guess* in *each trial*—or how much the participant's guess different from their partner's.

Because we are interested in understanding this process dynamically, we captured participants' progress over time by creating a **cumulative guess counter**, serving as a form of abstracted time spent engaging with one another and the experiment. While the measure of coordination in Model 1 presented a time-abstracted measure of coordination across the entire experiment, this model provides a snapshot of coordination in real time, measuring the learning and coordinative processes from guess to guess.

For Model 2, we built a linear mixed-effects model predicting the cumulative guess counter with each participant's true error, partner error, the interaction term between the two, and training improvement (described above) as fixed effects. We also included random effects for participant and difficulty. (We did not include dyad as a random effect in this model because the variance in the guesses occurred at the participant level.)

Model 3 Specifications: Do Social Factors Impact Coordination in Minimally Interacting Contexts? To explore the role that social judgements can play even in minimally interactive contexts, our final model considered how trust might impact coordination. For this model, we captured a third measure of coordination: the participant's willingness to change their guess, which was captured by the **total number of guesses** that each participant submitted in each trial.

Because participants individually chose whether to keep their previous guess or submit a new one while being able to see their partner's guess, we could expect that participants who trust their partner more would be more likely to change their guess—especially if there were large **absolute differences between the partner's first guesses**. Trust was measured as each participant's self-reported Likert-style **rating of their trust in their partner** ("How much do you feel you trusted your partner's opinion during the experiment?").

Model 3 was a linear-mixed effects model predicting the total number of guesses in a trial with fixed effects for their trust in their partner and for the difference in participants' first guess on that trial, while controlling for trial number and how much they improved during their own training. Model 3 also included participant and difficulty as random effects.

Model Implementation

All models were built as linear mixed-effects models in R (R Core Team, 2016) with 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). While we do not have space in the current paper to provide precise model specifications, we have made our code (`http://www.github.com/a-paxton/perception-memory-coordination`) and data (`https://osf.io/8fu7x/`) fully and freely available for others.

Results

Model 1 Results: Coordinated Error over Time

As predicted, we found that dyads were significantly and strongly coupled in their error ratings (β=-0.43, p<0.0001), with no effect of training improvement (β=-0.06, p=0.36). In other words, 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 Results: Learning *and* Coordination

We found that participants' rates of adaptation to their partner significantly differed from their rates of adaptation to the game (β=-0.06, p<0.008; see Fig. 2). In other words, this model revealed signatures of simultaneous learning *and* coordinative processes during the game: Players became attuned to the learning task while coordinating with one another's cognitive processes.

Aside from the main effect of partner adaptation, no other predictors reached statistical significance (all ps>0.25).

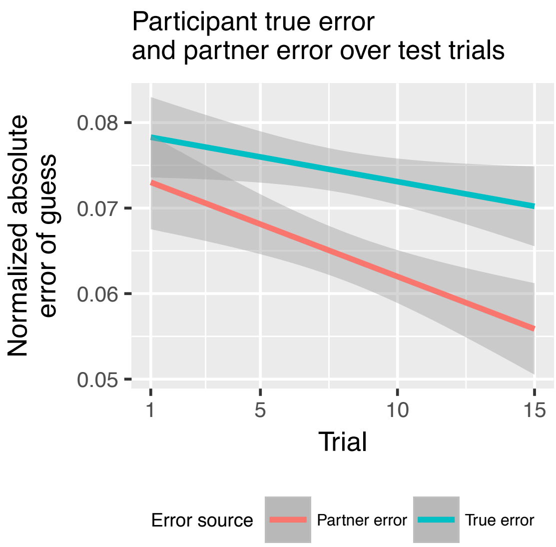

Figure 2: Difference over time in coordinative and learning processes, or the change in guess deviation from truth (in blue) and from their partner's guess (in red) over the game.

Model 3 Results: Social Signals in Minimal Contexts

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 (β=0.07, p<0.042), although we found no difference in the number of guesses based solely on the difference between the partners' first guesses (β=0.01, p=0.65). Ratings of partner trust were normally distributed around a mean of 5.96 (SD: 2.24; range: 1–10). In other words, although Models 1 and 2 showed participants improving and becoming more similar across trials, participants were more willing to concede that their partner's guess was correct when the participant trusted their partner—regardless of how similar or different the two partners' first guesses were on that trial.

Interestingly, we found that participants took more guesses on test trials when they improved more in their training trials (β=0.05, p<0.028). Assuming that those with the greatest training improvement were the most poorly performing initial players, this suggested that poorer-performing players were more likely to divide the cognitive labor of the task and follow the lead of their higher-performing partner.

We also saw an effect of trial (β=-0.05, p<0.02), indicating that people changed their guesses fewer times per trial as the game progressed. This could be an effect of learning (i.e., because both participants improved and became more similar from trial to trial), experiment fatigue (e.g., if participants simply wanted to end the game more quickly), or some combination of the two.

Discussion

Inspired by established lines of research on interpersonal coordination, we explored how minimally interactive contexts can shape the emergence of interpersonal dynamics. Using an online experiment that provided participants with only one channel of information about one another—their estimates—we found evidence of coordination of cognitive systems despite minimal social information and context.

As expected, we found 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 in 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. Individual learning and performance were unable to fully account for the players' similarity to one another.

Finally, like other work on coordination, we found that coordination was 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 found that players' decisions to change their guesses were influenced by their self-reported ratings of their partner's trustworthiness.

While a relatively simple paradigm, the present work contributes to the theoretical landscape of interpersonal coordination research in several ways. First, we continued to expand investigations of minimally interactive contexts, an underexplored avenue in an area that often relies on fully interactive paradigms; the present work extends our understanding of coordination by demonstrating that interacting individuals coordinate even when they have *only* a single channel of task-relevant sensory information available to them. Second, we explored the degree to which even low-level properties of memory and perception become entrained, despite this minimal information—explicitly probing questions of memory systems posed by transactive memory (Tollefsen et al., 2013). Finally, we applied questions of social impacts on coordination into lower-level behavioral and perceptual channels. Taken together, the present study found support for hypotheses that are natural but necessary extensions of a host of related previous work, providing explicit tests for ideas that are often implicitly accepted by coordination researchers.

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 subtle social judgements. In addition to these theoretical contributions, we also hope to have provided an example of the utility of 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 explored trustworthiness as a social construct during a task that allowed only minimal participation between participants. The trustworthiness measure was intentionally posed broadly, providing a signal of the latent social information that participants constructed when *only* their partner's task-related behavior was available to them. While we see their responses as a signal of very subtle social forces, we readily recognize that these ratings of trustworthiness may have been influenced by the individual's own confidence or ability—as many social judgements are often influenced by the assessor's own characteristics. Future work should expand on this to explore the interplay between individual and interpersonal assessments and should delve more deeply into understanding what might dynamics might be influencing these minimal social judgements.

Second, given the broad participant pools available through crowdsourcing, 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, but many real-world social settings include more than two people—settings which comprehensive theories of coordination must also capture. Crowdsourcing and real-time social experiments enable researchers 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 and a Moore-Sloan Data Science Environments Fellowship to AP (through the Gordon and Betty Moore Foundation [Grant GBMF3834] and the Alfred P. Sloan Foundation [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.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tyln, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, *23*(8), 931–939.

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.

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.

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.

Tollefsen, D. P., Dale, R., & Paxton, A. (2013). Alignment, transactive memory, and collective cognitive systems. *Review of Philosophy and Psychology*, *4*(1), 49–64.

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*, 1323.