Joint Action Coordination through Strategic Reduction of Variability

Cordula Vesper (vesperc@ceu.hu)

Department of Cognitive Science, Central European University Frankel Leó út 30-34, Budapest 1023, Hungary Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

Laura Schmitz (laschmit@uos.de)

Department of Cognitive Science, University of Osnabrück Albrechtstraße 28, 49076 Osnabrück, Germany

Natalie Sebanz (sebanzn@ceu.hu) Günther Knoblich (knoblichg@ceu.hu)

Department of Cognitive Science, Central European University Frankel Leó út 30-34, Budapest 1023, Hungary Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

Abstract

How do people coordinate actions with others? We tested the hypothesis that pairs of participants strategically reduce the variability of their action performance to achieve synchronicity in the absence of visual feedback about each other's actions. Consistent with this prediction, participants moved faster and less variably in a condition where they could not see their task partner's movements compared to a condition in which visual information was available. The accuracy of the resulting coordination was the same in both conditions. These findings are interpreted as evidence for general strategic adaptation in the service of real-time action coordination when only minimal perceptual information is available.

Keywords: Joint action; coordination strategy; cooperation; social cognition.

Introduction

Whenever people coordinate their actions with other people, they are engaged in a 'joint action' (Clark, 1996; Marsh et al., 2009; Sebanz, Bekkering, & Knoblich, 2006). Depending on the specific task and the presence or absence of an explicit joint action goal, different mechanisms and processes will make coordination of multiple people's actions possible. For instance, a couple might discuss through verbal or non-verbal communication who is responsible for preparing dinner and who will set the table (Clark & Wilkes-Gibbs, 1986). Or a group of friends might help push-start a car by using perceptual cues and haptic information to predict when everyone else will push (van der Wel, Knoblich, & Sebanz, 2011; Wilson & Knoblich, 2005). In yet other cases, coordination might arise without prior planning as when two strangers unintentionally walk in synchrony (van Ulzen et al., 2008).

While people in these and many other everyday examples make use of visual, auditory or haptic information to guide their joint efforts, this is not always possible. Sometimes coordination is required in contexts where only little or even nothing is known about the coordination partner and how or when the partner will perform a particular action. In these cases, all that might be represented is one's own action part ('ME'), the fact that someone will take care of another action part ('X') required to achieve the joint goal and the joint action goal ('ME+X') achieved by combining the individual action parts (Vesper et al., 2010). Thus, a precise representation about the partner's task might not be available. We claim that in these cases, coordination is supported by very general mechanisms and processes that are not required to the same extent if more information about a task partner is available. The present study addressed the mechanisms and processes allowing people to achieve coordination in this kind of minimal joint action situation.

More specifically, we investigated whether people who intend to coordinate their actions under real-time constraints and with no access to visual information about a task partner's actions adapt their own actions in a way that will make interpersonal coordination most likely. Such a coordination strategy (Vesper et al., 2010) reliably simplifies coordination in a general way, i.e. it is a modulation of one's own behavior that does not directly depend on how a task partner's particular action will unfold.

One example of strategic adaptation is to behave in a way that will make one's own actions predictable. When timing is not critical, this could involve relying on shared or conventional knowledge (Clark, 1996). For example, someone might decide to wait at the Brandenburg Gate to meet a friend in Berlin when they forgot to agree on a precise meeting point in advance (Schelling, 1960). Similarly, if each member of a group has to guess a number such that the sum of all numbers matches a randomly selected target number, providing consistent and therefore predictable guesses can be beneficial to achieve the desired group outcome (Roberts & Goldstone, 2011).

In situations in which actions need to be coordinated in real-time, making actions predictable can involve minimizing the variability of one's own performance. Recent empirical evidence for this claim is provided by a study in which pairs of participants performed a simple two alternative forced choice (2-AFC) reaction time task next to each other with the goal of synchronizing the timing of their response button presses (Vesper et al., 2011). An analysis of mean reaction times and the trial-by-trial variability of reaction times indicated that participants responded faster and with less variability in joint action compared to individual baseline performance. This in turn positively affected coordination such that pairs whose members responded fast and with little variability were on average better synchronized.

Critically, the study showed that it was the reduction in variability that predicted how successful coordination was, as demonstrated by a correlation of variability and asynchrony that persisted when controlling for the potential effects of mean reaction time. Thus, the more predictable actions were, the more successful interpersonal coordination was. Given that performing tasks at higher speed tends to reduce temporal variability (Wagenmakers & Brown, 2007), participants most likely used speeding as a means to reduce their action variability. A second experiment demonstrated that speeding and predictability were only correlated with asynchrony when task partners intended to synchronize their button presses, but not in an experimental condition where the two people merely performed the task next to each other without a coordination goal. This suggests that the coordination strategy of making oneself predictable is used specifically to achieve intentional joint action coordination.

The aim of the present study was to extend these earlier findings (Vesper et al., 2011) by addressing three predictions following from the concept of a coordination strategy (Vesper et al., 2010) - generalizability, specificity, and independence. The first prediction was that the link between response speed, response variability, and asynchrony of task partners' actions would also be useful for coordination of more complex, temporally extended joint actions. Therefore, we instructed pairs of participants to each use a computer mouse to move a cursor on a screen from a start location towards a target with the joint goal of reaching the target at the same time (Figure 1). Thus, the task required two people to synchronize the endpoints of two-dimensional aiming movements. When they reached the target auditory feedback informed participants about their coordination accuracy.

The second prediction was that a coordination strategy will predominantly be used in situations in which no or only little information about a task partner's actions is available. In this 'minimal' case, all someone can do is to adapt his or her own actions in a general way to make coordination most likely. In contrast, when task or perceptual information is available, other mechanisms and processes will support coordination. For instance, co-actors can monitor (Malfait et al., 2009; Schuch & Tipper, 2007) or predict (Graf et al., 2007; Knoblich & Jordan, 2003) when and how another person will perform a particular action. Consequently, in many situations, perceptual information is beneficial for joint action coordination. As an example, when two people build a toy model together such that one person (the director) verbally instructs another person (the builder) which parts to assemble, coordination is more successful if the director can see what the builder is doing (Krych-Appelbaum et al., 2007). Similarly, two people who jointly search a shared workspace for a target object are more efficient in their search if they receive information about where each of them is currently looking at (Brennan et al., 2008). To test the specificity of coordination strategies, we compared an experimental condition in which co-actors did not receive visual information about each other (Other *Hidden*) with one in which they could see each other and each other's ongoing action performance (Other Visible). We hypothesized that a speeding and predictability strategy would predominantly be employed in the Other Hidden condition, whereas for Other Visible, we expected that the additional perceptual information would allow co-actors to use a different mechanism for coordination. This could involve monitoring and anticipating the partner's computer mouse movements. Therefore, we expected reaction times and movement variability to be smaller in Other Hidden compared to Other Visible. Given that perceptual information often positively influences coordination, we also hypothesized that asynchronies between co-actors' actions in Other Visible would be smaller, indicating better coordination accuracy when more information is available.

The third prediction was that a coordination strategy is used in a general way, independently of how the task partner actually performs an action. This means that the partner's particular action performance is not directly relevant for one's own strategic adaptation. In contrast, when other mechanisms such as monitoring and prediction are used, one's own action performance should be directly related to the task partner's action performance. One way to address this prediction is to compare the actually measured asynchronies between task partners' actions with asynchronies that are calculated after co-actors' reaction times have been shuffled and randomly matched. This method effectively treats the data as if each person's actions were not targeted towards a corresponding action of the co-actor because their actions now come from different trials. We hypothesized that this procedure would affect coordination in Other Hidden to a lesser extent than in Other Visible, indicating that co-actors in the former case adapt in a general way that is independent of the task partner's particular action performance, whereas in the latter case, co-actors make use of the given perceptual information and take into account how the partner's action unfolds on a trial-by-trial basis.

Method

Participants

Twenty-four students (14 women) participated in pairs. They were between 19 and 25 years old (mean 21.1 years) and right-handed. They gave prior informed consent and received monetary compensation for their participation.

Material and Apparatus

A "space mission" scene was created on two computer screens placed next to each other (Figure 1). The scene contained three elements presented on a dark blue background. First, close to the outer margin of each screen, a yellow "spaceship" was drawn (ca. 2.5 cm x 1.9 cm; position centrally on the vertical axis), indicating the starting position for each trial. Second, on the inner margin of each screen, a blue half circle was drawn on one of three possible locations, indicating a "planet" as the target (radius ca. 2.0 cm or 3.8 cm; position at 20 %, 50 % or 80 % from the upper screen margin). When both screens were visible (Other Visible), the two half circles together formed a complete "planet". Finally, centrally between "spaceship" and "planet", on one of five possible locations, an array of small differently-sized white dots was drawn to represent an "asteroid belt" (ca.1.9 cm x 9.3 cm; position at 20 %, 35 %, 50 %, 65 % or 80 % from upper screen margin). It served as a potential obstacle between start and target locations.



Figure 1: The "space mission" scene (example layouts). A) In *Other Hidden*, each participant only saw one half of the scene due to an occluder placed between the participants. B) In *Other Visible*, both participants saw the complete scene.

The stimuli were presented on two 17"-screens (resolution 1280 x 1024 pixel, refresh rate 60 Hz). In individual baselines and in *Other Hidden*, a black card board (70 x 100 cm) was set up between the two participants and between the two screens. The experiment was run on two Dell OptiPlex computers that were connected through a null-modem cable to allow online data exchange. For data

collection, two special gaming computer mice (Logitech G500) were used that were sampled at 100 Hz and that had automatic acceleration turned off. Matlab version 2012a was used for controlling the experiment and for data analysis.

Procedure

There were four experimental parts: the *Other Hidden* condition (Figure 1A), the *Other Visible* condition (Figure 1B) and two individual baselines. Each participant first performed four practice trials and then the first individual baseline, while the task partner waited in another room. After both participants had finished their first individual parts, they performed the two joint conditions together. The order of *Other Hidden* and *Other Visible* was counterbalanced. Finally, each participant separately performed another individual baseline. Each of the four parts consisted of six experimental blocks à 16 trials with short breaks in between. The overall duration of the experiment was about 1.5 hours.

At the beginning of a trial, the start location ("spaceship") was presented for 600 ms. Next, the target ("planet") and obstacle ("asteroid belt") appeared at a location that was randomly chosen from the possible locations. The frequency of target and obstacle locations and the target size were counterbalanced within each block. The relation of target and obstacle locations determined whether the direct path between start and target location was blocked by the obstacle or not. At the same time when target and obstacle appeared, the spaceship briefly flashed for 200 ms by showing flames at the rear engine. The purpose was to redirect participants' attention to the start location where a mouse cursor (a yellow circle) was now visible.

Participants were instructed to move the mouse cursor to the target without moving over the obstacle. A short feedback tone (100 ms) was played as soon as they moved into the target area, i.e. no button press was required. The feedback tones for the left-seated and the right-seated participants differed in frequency so that they could be distinguished (1100 Hz, 1320 Hz). Additionally, visual feedback about the accuracy of the trial was given: The planet turned red indicating negative task performance 1) if participants' movements were too slow (movement onset > 600 ms or reaction time > 1600 ms), 2) if they moved over the obstacle area, 3) if the task partner had made any of these mistakes or 4) if co-actors did not reach the target synchronously (absolute asynchrony $> 400 \text{ ms})^1$. In all other cases, trials were successful and the planet turned into a bright green. Participants then returned to the start position and the next trial started.

Participants were told to think of the task as a space contest that requires securing planets from an alien nation by landing on a planet before them. According to this background story, in some areas of the universe (individual baselines), this could be achieved alone, whereas in other

¹ During individual baselines, only the first two criteria generated negative feedback.

areas of the universe (*Other Hidden*, *Other Visible*), they would have to arrive at the planet at the same time as the task partner in order to win. Thus, task instructions explicitly mentioned that participants should be as fast and as accurate as possible, while strongly focusing on arriving at the same time. Co-actors were not allowed to talk.

Results

For the purpose of the present paper, we only report analyses of mean reaction times (RT; measured as the time from the start signal until the target was reached), standard deviation of reaction times (STD) and absolute asynchrony between participants' reaction times (ASYNC)². These dependent variables were acquired by averaging over all trial types within a condition, i.e. we did not differentiate between different target and obstacle locations or target sizes. All trials in which participants' own RT was slower than 1600 ms or in which they moved over the obstacle area were excluded from further analyses (1.1 % in *Other Hidden*, 0.7 % in *Other Visible*).



Figure 2: Results. A) Mean RT. B) Mean trial-by-trial variability measured as STD. The dotted lines show individual baseline performance before (upper line) and after joint action (lower line). Error bars display within-subject confidence intervals (Loftus & Masson, 1994).

We first tested whether participants made use of a coordination strategy predominantly in the case where they did not receive visual information about the task partner's action. Confirming this hypothesis, RTs in *Other Hidden* were significantly faster, F(1,23) = 14.01, p < .01 (Figure 2A) and less variable, F(1,23) = 5.36, p < .05 (Figure 2B) than in *Other Visible*. Moreover, as described in more detail below, coordination between co-actors was equally good in the two conditions.

To investigate the hypothesized relation of RT, STD and ASYNC, we performed zero-order and partial correlations.

For *Other Hidden*, these analyses indicated that both RT and STD significantly influenced ASYNC such that shorter and less variable RTs led to better coordination between coactors (for exact results, see Figure 3A). Crucially, however, when controlling for RT in a partial correlation, STD still predicted ASYNC, whereas when controlling for STD, the relation between RT and ASYNC did not persist. Thus, as predicted, participants' response variability was critical in determining how well coordinated co-actors were when no online perceptual information about the task partner's actions was available. In contrast, in *Other Visible*, RT and STD did not predict ASYNC, although RT and STD were correlated (for exact results, see Figure 3B).



Figure 3: Zero-order and partial (in parentheses) correlations for A) *Other Hidden* and B) *Other Visible*. The thick arrow in A indicates that the relation between STD and ASYNC still holds when controlling for the influence of RT with a partial correlation. * p < .05, ** p < .01, *** p < .001.

Finally, we tested the hypothesis that the speeding and predictability strategy in Other Hidden is general in the sense that it depends only to a certain extent on how the task partner actually performed his or her actions. In contrast, coactors in Other Visible should take the other person's actual movements into account by monitoring and predicting the other's action. Based on this reasoning, we hypothesized that the coordination outcome from a general strategy should depend less on a trial-by-trial match of task partners' actions, whereas when using perceptual information this should be relevant. To test this prediction, we compared the originally measured asynchronies with asynchronies in which the specific trial-by-trial relation between co-actors' actions was destroyed by randomly shuffling the order of trials from one member in each pair (separately for the different trial types) and re-calculating the asynchrony between the two persons' response times.

In line with our hypothesis, a comparison of original and shuffled asynchronies in the two conditions indicated an unequal effect of the shuffling: Although asynchronies increased in both conditions, shuffling co-actors' trial order had a significantly stronger effect for *Other Visible* than for *Other Hidden*. This was demonstrated statistically by an interaction of the factors Condition (*Other Hidden, Other*

² For every pair and condition, half the trials were used to calculate ASYNC for one person and the remaining trials for calculating ASYNC for the other person (randomly distributed). This allowed us to perform all analyses with the full degrees of freedom.

Visible) and Trial Order (original, shuffled), F(1,23) = 10.59, p < .01 (Figure 4). There was also a main effect of Trial Order, F(1,23) = 16.88, p < .001, but no significant effect of Condition, F(1,23) = .01, p > .9. Thus, co-actors reached the same level of coordination performance in the two conditions.



Figure 4: Asynchrony measured (original) and re-calculated after randomly matching different trials from the members within a pair (shuffled). Error bars display within-subject confidence intervals (Loftus & Masson, 1994).

Discussion

The aim of the present study was to systematically test predictions following from the hypothesis that people strategically increase the speed and predictability of their actions to achieve real-time coordination with another person given only minimal perceptual information (Vesper et al., 2010, 2011). Pairs of participants performed mouse movements towards a target displayed on a computer screen. The joint goal was to reach the target at the same time as the task partner. Short feedback tones when arriving at the target informed participants about the accuracy of their joint coordination.

The present results confirm our hypothesis that co-actors strategically reduced the variability of their movements through speeding (Wagenmakers & Brown, 2007) whereby action variability contributed directly to the coordination outcome. In particular, when controlling for the impact of reaction times, response variability still predicted asynchrony, whereas when controlling for the impact of response variability, there was no longer a correlation between reaction time and asynchrony. This not only replicates earlier findings (Vesper et al., 2011) but also demonstrates that this coordination strategy can support coordination in complex, temporally extended actions.

Furthermore, the present study shows that the predictability strategy is predominantly used in situations in which little or no information about the task partner is available. To test this hypothesis, we compared an experimental condition in which co-actors could not see each other (*Other Hidden*) with one in which visual information was available (*Other Visible*). Consistent with our predictions, participants' movements were significantly faster and less variable without visual information and the relation between reaction time, variability, and coordination accuracy was present only in the *Other Hidden* condition. This confirms that coordination strategies are specific such that they are predominantly employed when other mechanisms like monitoring and predicting another's actions cannot be used.

A third hypothesis was that the employment of strategic behavior modulations would not depend on how the task partner actually performs his or her actions. Therefore, we compared two types of asynchronies: One that we had actually measured (original) and one that we calculated after shuffling the order of task partners' reaction times and randomly matching them again (shuffled). This resulted in a measure of how much each person's action was related to the task partner's actual action performance. Confirming our hypothesis, shuffling the trial order affected coordination significantly less in *Other Hidden* compared to *Other Visible*. Thus, when making oneself predictable one's own actions do not or only to a small extent depend on how exactly the task partner performs his or her action.

A possibly surprising result of the current study is that coactors were on average equally well-coordinated in the two joint conditions. Although perceptual information is often beneficial for joint action (e.g. Brennan et al., 2008; Knoblich & Jordan, 2003; Krych-Appelbaum et al., 2007), this suggests that using a coordination strategy can compensate for a lack of perceptual information. Moreover, given that participants' actions were overall faster and less variable when no perceptual information was available, one might even argue that action performance was better without visual feedback. This is consistent with other evidence that having 'redundant' information potentially impairs joint action coordination. Specifically, when two people who jointly search a workspace not only receive visual information about each other's looking behavior, but can also talk to each other about the task, their search is considerably less efficient than when verbal communication is restricted (Brennan et al., 2008). As an alternative, the present findings might indicate that taking away perceptual information requires co-actors to put in extra effort in order to achieve the same degree of coordination.

How then did co-actors approach the task in the *Other Visible* condition? Although investigating this in detail is beyond the scope of the present paper, participants most likely used the available visual information to guide their actions either reactively (monitoring the task partner's action, then acting oneself) or predictively (anticipating when the task partner will reach the target and acting in accordance with this prediction). Further experiments could distinguish these two cases, e.g., by measuring participants' eye movements to determine at what time during the interaction they track their task partner's movements.

The present study has implications beyond human joint action. For instance, an important research topic in the

cognitive sciences is how to implement real-time interaction of a robot and a human user. To that end, mechanisms observed in human social interaction are currently being transferred to robot platforms, including natural-language discourse (Salem et al., 2010), action prediction (Bicho et al., 2011; Dindo, Zambuto, & Pezzulo, 2011) and continuous movement synchronization (Mörtl et al., 2012). Considering also general strategic behavioral adaptions for human-robot interaction can be beneficial for this endeavor. First, human users might employ a strategy such as making oneself predictable also when interacting with a robot so that coordination would improve if the robot used the same strategy. Second, human users might expect the robot to adapt its movements strategically so that robots that do so would appear more 'human-like' and thereby are more easily accepted as an interaction partner.

Taken together, this study provides evidence that general strategic adaptations of one's own actions can effectively support coordination with other people in situations in which precise representations about the partner's task might not be available. The concept of a coordination strategy therefore complements other approaches towards joint action like those focusing on communication (Clark, 1996), action prediction (Wilson & Knoblich, 2005) or dynamic perception-action coupling (Marsh et al., 2009).

Acknowledgments

The research reported in this manuscript was supported by the ESF program EuroUnderstanding. We thank Veronika Herendy for her help in collecting the data.

References

- Bicho, E., Erlhagen, W., Louro, L., & Costa e Silva, E. C. (2011). Neuro-cognitive mechanisms of decision making in joint action: A human–robot interaction study. *Human Movement Science*, 30, 846-868.
- Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition*, 106, 1465–1477.
- Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Dindo, H., Zambuto, D., & Pezzulo, G. (2011). Motor simulation via coupled internal models using sequential Monte Carlo. *Proceedings of the International Joint Conference on Artificial Intelligence*, 2113-2119.
- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in realtime. *NeuroImage*, 36, T22–T32.
- Knoblich, G. & Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory,* and Cognition, 29(5), 1006–1016.
- Krych-Appelbaum, M., Law, J. B., Jones, D., Barnacz, A., Johnson, A., & Keenan, J. P. (2007). "I think I know what

you mean" - The role of theory of mind in collaborative communication. *Interaction Studies*, 8(2), 267–280.

- Loftus, G. & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin* and Review, 1, 476–490.
- Malfait, N., Valyear, K. F., Culham, J. C., Anton, J.-L., Brown, L. E., & Gribble, P. L. (2009). FMRI activation during observation of others' reach errors. *Journal of Cognitive Neuroscience*, 22(7), 1493–1503.
- Marsh, K. L., Johnston, L., Richardson, M. J., & Schmidt, R. C. (2009). Toward a radically embodied, embedded social psychology. *European Journal of Social Psychology*, 39, 1217–1225.
- Mörtl, A., Lorenz, T., Vlaskamp, B., Gusrialdi, A., Schubö, A., & Hirche, S. (2012). Modeling inter-human movement coordination: synchronization governs joint task dynamics. *Biological Cybernetics*, 106, 241-259.
- Roberts, M. E. & Goldstone, R. L. (2011). Adaptive group coordination and role differentiation. *PLoS ONE* 6(7), e22377.
- Salem, M., Kopp, S., Wachsmuth, I., & Joublin, F. (2010). Towards an integrated model of speech and gesture production for multi-modal robot behavior. *Proceedings* of the IEEE International Symposium on Robot and Human Interactive Communication, 649-654.
- Schelling, T. C. (1960). *The strategy of conflict*. Harvard University Press, Cambridge.
- Schuch, S., & Tipper, S. P. (2007). On observing another person's actions: influences of observed inhibition and errors. *Perception and Psychophysics*, 69(5), 828–837.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 71–76.
- van der Wel, R., Knoblich, G., & Sebanz, N. (2011). Let the force be with us: Dyads exploit haptic coupling for coordination. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1420–1431.
- van Ulzen, N. R., Lamoth, C. J. C., Daffertshofer, A., Semin, G. R., Beek, P. J. (2008). Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side. *Neuroscience Letters*, 432, 88–93.
- Vesper, C., Butterfill, S., Knoblich, G. & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23, 998-1003.
- Vesper, C., van der Wel, R., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211(3-4), 517-530.
- Wagenmakers, E. J. & Brown, S. (2007). On the linear relation between the mean and the standard deviation of a response time distribution. *Psychological Review*, *114*, 830–841.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3), 460–473.