Examining Multiscale Movement Coordination in Collaborative Problem Solving

Travis J. Wiltshire (wiltshire@sdu.dk)
Centre for Human Interactivity, Department of Language & Communication, Campusvej 55
DK-5230 Odense M, Denmark

Sune Vork Steffensen (s.v.steffensen@sdu.dk)
Centre for Human Interactivity, Department of Language & Communication, Campusvej 55
DK-5230 Odense M, Denmark

Abstract

During collaborative problem solving (CPS), coordination occurs at different spatial and temporal scales. This multiscale coordination should, at least on some scales, play a functional role in facilitating effective collaboration outcomes. To evaluate this, we conducted a study of computer-based CPS with 42 dyads. We used cross-wavelet coherence as a way to examine the degree to which movement coordination is evident at a variety of scales and tested whether the observed coordination was greater than both the amount expected due to chance and due to task demands. We found that coordination at scales less than 2s was greater than expected due to chance and at most scales (except 16s, 1m, and 2m) was greater than expected due to task demands. Lastly, we evaluated whether the degree of coherence at scales less than 2s, and the form of coordination (in terms of relative phase), were predictive of CPS performance. We found that .25s and 1s scales were predictive of performance. When including relative phase, our results suggest that higher in-phase movement coordination at the 1s scale was the strongest predictor of CPS performance. We discuss these findings and detail their relevance for expanding our knowledge on how coordination facilitates CPS.

Keywords: coordination; collaboration; problem solving; team performance; dynamical systems; synchrony.

Introduction

Collaborative problem solving (CPS) is a cognitive skill pervasive in many human interaction contexts ranging from everyday life to highly complex work environments. CPS is defined as “a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution” (OECD, 2015, p. 6). Given the increasing complexity of problems in contemporary societal practices, and the need for multiple disciplines to solve them, CPS has been recognized as an essential 21st century skill. However, while some research has examined CPS in a variety of laboratory (e.g., Berg, Johnson, Meegan, & Strough, 2003; Roschelle & Teasley, 1995) and naturalistic contexts (e.g., Fiore, Wiltshire, Oglesby, O’Keefe, & Salas, 2014; Jordan & McDaniel Jr, 2014), the state of the science is still limited. In this paper, we explore the implications of the notion that CPS is a multiscale phenomenon, and investigate the degree to which movement coordination, at various scales, plays a functional role in effective CPS.

Human interaction in any context is a dynamic, multiscale phenomenon (e.g., Dale, Fusirollo, Duran, & Richardson, 2013; Steffensen & Pedersen, 2014). For example, during a conversation, neural events transpire on the order of milliseconds, speech production and gestures over seconds, and the conversation itself on the order of minutes (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012). This point, while oversimplified, illustrates the fact that human interaction involves a variety of temporal and spatial scales (e.g., neural, physiological, bodily). Recognizing human interaction as multiscale, especially during CPS, implies that coordination, both intra- and inter-personally, must span a variety of these spatial and temporal scales in order to effectively accomplish joint goals (Eiler, Kallen, Harrison, & Richardson, 2013).

Indeed, a common question, particularly in the movement sciences, has been to understand how systems with high degrees of freedom, are able to functionally coordinate (e.g., Mitra, Amazeen, & Turvey, 1998). Coordination in this context is simply the ways in which components and processes of a system change together over time (Butner, Berg, Baucom, & Wiebe, 2014). In an interpersonal context, a wide variety of terms have been used to describe different forms of coordination (Butler, 2011) such as synchronization, co-regulation, entrainment, and coupling. Evidence for many forms of interpersonal coordination are quite pervasive amongst differing modalities (Fusaroli & Tylén, 2016; Louwerse, Dale, Bard, & Jeuniaux, 2012) and contexts (Palumbo et al., 2016). But, while many different forms of coordination have been discovered, the ways in which they facilitate effective interaction outcomes is less studied (cf., Timmons, Margolin, & Saxbe, 2015), particularly in collaborative contexts.

Prior research has suggested that coordination is required for the accomplishment of joint goals (Mills, 2014) and that stronger coordination should contribute to better collaborative results (Barron, 2000). Findings so far have been mixed, though, with regard to how coordination, albeit in different modalities and scales, relates to optimal performance on joint tasks (Gallotti, Fairhurst, & Frith, 2017). In one example, performance on a dyadic movement task was predicted by a measure of coordination that reflected interaction across multiple time scales (Davis, Brooks, & Dixon, 2016). In another example, Abney, Paxton, Dale, and Kello (2015) found that stronger coordination in bodily movements were associated with poorer performance on a movement-based dyadic problem solving task. Louwerse et al. (2012) found that as task difficulty increased, so too did coordination. Further,
coordination of bodily movements in psychotherapy, a highly collaborative endeavor (Tryon & Winograd, 2011), were shown to link to effective treatment outcomes (Ramseyer & Tschacher, 2011). While not related to performance specifically, a number of interactional benefits have been observed following periods of interpersonal coordination such as increased affiliation (Hove & Risen, 2009) and cooperative behavior (Valdesolo, Ouyang, & DeSteno, 2010; Wiltermuth & Heath, 2009).

Given the extant research, briefly reviewed here, we expect that multiscale coordination of bodily movements should have a functional relationship with performance in complex, CPS. Relatedly, a recent theoretical account of dialog proposed that high coordination in lower level behaviors (e.g., posture) may provide a necessary foundation for more variability and complementarity at higher levels of dialog (Fusaroli, Rączaszek-Leonardi, & Tylén, 2014). So, a key aspect that distinguishes the present work from prior research is that we focus on movement coordination in a challenging computer-based CPS context.

**The Current Study**

The present work is part of a larger study examining team interaction dynamics during a dyadic CPS task (Wiltshire, 2015; Wiltshire, Butner, & Fiore, 2017). Whereas in prior work, we have examined how transitions in communication structures and their complexity relate to CPS performance (Wiltshire et al., 2017), here we focus on coordination of bodily movements at various time scales and how that coordination relates to CPS performance.

We utilize cross-wavelet coherence as a way of examining coordination of human interaction that is largely unstructured, at least when compared to rhythmic movement tasks (Fujiwara & Daibo, 2016; Issartel, Bardainne, Gaillot, & Marin, 2015). This method allows for evaluation of the degree of coordination of two continuous time series and whether that coordination is in-phase or anti-phase. One of its key strengths is that it retains a high level of precision in both the time and frequency domains (Issartel et al., 2015). This method has been used previously to examine movement coordination in a variety of interactive contexts such as the exchange of jokes (Schmidt, Morr, Fitzpatrick, & Richardson, 2012), dialog (Fujiwara & Daibo, 2016), the coordination of jazz musicians (Walton, Richardson, Langland-Hassan, & Chemero, 2015), and dancers (Washburn et al., 2014).

We expect that the coordination of bodily movements will serve a functional role in facilitating effective CPS performance. However, we also expect that this functional role will vary based on time scales. In other words, movement coordination at some time scales should be more relevant to CPS than others. We thus adopt an exploratory approach to determine what scales are important in predicting effective CPS performance. Given the nature of the task, it is likely that smaller time scale movements will matter such as those that occur while controlling the computer-based task as well as during speech. When examining interpersonal coordination dynamics, it is essential to demonstrate that the observed coordination is greater than can be expected due to chance alone (Ramseyer & Tschacher, 2010), and that it is not solely due to task constraints (Strang, Funke, Russell, Dukes, & Middendorf, 2014). Thus, we advance the following research hypotheses (H) and research questions (RQ):

- **H1:** Movement coordination will be greater than chance, at least at lower scales.
- **H2:** Movement coordination will be greater than can be expected due to task constraints.
- **RQ1:** At what scales does movement coordination predict CPS performance?
- **RQ2:** Does the form of coordination (e.g., in-phase, anti-phase) at these scales relate to performance?

**Method**

**Participants**

84 undergraduate students (31 female, \(M_{age} =19.2\) years, range 18-28 years; \(\sim 67\%\) White, \(8\%\) Black, \(10\%\) Hispanic, \(10\%\) Asian, and \(5\%\) other) from a large United States university voluntarily participated in this experiment comprising 42 dyadic teams. There were five female-only teams, 17 male-only teams, and 20 mixed-gender teams. Participants must have had general video game experience using a mouse and keyboard for third-person video games, no prior history of seizures, no experience using the Moonbase Alpha simulation, and no prior acquaintance.

**Materials**

Participants sat face-to-face with each other with two desktop computers offset to one side. This setup allowed them to view the other’s face and torso. The computer screens were placed back-to-back. A Logitech HD webcam model C615 was used to record the participants from a profile view. All videos were collected in 720p resolution.

**Task**

NASA’s Moonbase Alpha is a complex, CPS task (NASA, 2011) that places team members in a simulated scenario where a meteor strike damages critical life support systems of a moonbase. The goal of the Moonbase Alpha task is for participants to fully restore oxygen to the settlement in 25 minutes or less. Both team members must work together to solve the problem by figuring out how to fix and/or replace damaged components of the life support system such as solar panels, power cables, couplers, and a power distributor. A variety of tools and coordination strategies must be employed to complete the task; however, there are no predefined guidelines for how to completely repair the settlement in the given timeframe.

**Procedure**

Participants were briefed about the nature of the experiment and asked to introduce themselves to each other by
providing a greeting and sharing their name with the other participant. Participants were then given an informed consent document to review and asked to complete a biographical questionnaire.

Participants were then provided a PowerPoint tutorial that covered the basics of the Moonbase Alpha simulation, which were derived from the simulation’s instruction manual (NASA, 2011). Further, participants were told that they would be tested on the content. After completing the PowerPoint, they received a 10-item multiple-choice knowledge assessment (see Wiltshire, 2015).

After completion of the knowledge assessment, the necessity for communication to complete the task was reiterated. Participants were then instructed to begin the simulation. A short video introduced the problem (i.e., the moonbase was damaged by a meteorite and life support functions need to be restored) before participants began the 25-minute task. The task was considered complete either when time ran out or once participants fully restored oxygen, whichever came first.

CPS Performance
Problem solving performance was determined by a rescaled combination of three variables: (a) the total time taken to restore life support (0-25 minutes), (b) the total percentage of oxygen restored (0-100%), and (c) a ratio of completed object repairs to the total possible repairs (0-25; only for teams that restored zero oxygen). The rescale function in R (R Core R Core Team, 2016) was used to place teams whose performance restored no oxygen at all into a range of 0-33 as a function of their ratio of object repairs/total possible object repairs. Those teams that restored some, but not all, oxygen were rescaled to fit a range of 34-66. Lastly, for those teams that restored all oxygen, the time to complete the task was inversely rescaled to fit the range of 67-100 (with lower times leading to higher scores).

Analytic Strategy
Frame-Differencing We used Paxton and Dale’s (2013) video frame-differencing technique to extract a time series representing the level of bodily movement for each participant from all videos at an 8 Hz sampling rate. For the current task, this measure of bodily movement captures behaviors such as speech, postural sway, gestures, adjustment of position, hand movements controlling the mouse and keyboard, and shifting of the legs and/or feet. In general, this technique provides an objective measure of the amount of movement a given participant is exhibiting moment-by-moment over the duration of the task with higher values corresponding to more movement. The trade-off when using this type of method is that there is a loss of specificity with regard to the types of movements that are coordinated, but movement can be extracted with relatively little effort and time compared to more specific movement coding systems (cf. Louwerse et al., 2012; Paxton & Dale, 2013).

Cross-Wavelet Coherence We examined dyadic movement coordination with the cross-wavelet transformation method by using the wtc function from the biwavelet package (Gouhier, Grinsted, & Simko, 2016) in R. This is a spectral decomposition method that allows for examination of time localized oscillations in a variety of frequencies and how the spectrum changes in those frequencies over time (Issartel, Marin, Gaillot, Bardainne, & Cadopi, 2006). This method is known to be robust to nonstationary time series (Issartel et al., 2015). We extracted the average coherence and average relative phase values from the following frequency ranges: .25s, .5s, 1s, 2s, 4s, 8s, 16s, 32s, ~1m, ~2m, and ~4.5m within +/- .5 scales (frequency scales are converted to time domain by multiplying them by the 8 Hz sampling rate and dividing by 60 for minutes). Coherence is the spectral equivalent to a cross-correlation. Values of 0 convey no coordination and a value of 1 conveys absolute coordination (Schmidt, Nie, Franco, & Richardson, 2014). Relative phase indicates whether the oscillations are in-phase (0°), anti-phase (180°), or exhibiting a lag (between 0° and 180°).

Surrogate and Virtual Pairs Analyses Surrogate analysis was conducted by computing a shuffled transformation of each observed movement time series and repeated the cross-wavelet analyses for each dyad. This effectively destroys the temporal pattern in the data while preserving the distributional properties (Louwerse et al., 2012). Any measures of coordination applied to these are widely interpreted as the degree of coordination expected due to chance (Ramseyer & Tschacher, 2010).

For the virtual pairs analysis, 42 randomized combinations of individuals who did not interact together were created. Because these individuals did not interact with each other, but were performing the same task, coordination measures calculated from virtual pairs have been interpreted as the coordination that can be expected due to the task demands (Strang et al., 2014). Thus, cross-wavelet analyses were conducted on these virtual pairs. Where time series were of unequal length, the longer time series was truncated to the length of the shorter series. Separate paired-sample t-tests were used to compare between observed coherence and surrogate coherence as well as between observed coherence and virtual pairs coherence for each time scale.

Examining Relationship Between Coordination and Performance Our approach to answering RQs 1 and 2 was exploratory based on the results from H1. Specifically, we first conducted a linear multiple regression model with the observed coherence values at scales that were significantly greater than chance as predictors of performance. Then, we took those values that were significant predictors of performance and included them in a second multiple regression model with the relative phase values for those respective scales.

Results
In order to examine H1, that the observed coordination would be greater than chance at some scales, the coherence
values for the surrogate data were compared to the coherence of the observed data. Results (see Table 1) suggested that coherence was significantly greater than chance at the .25s, .5s, 1s, and 2s frequency scales.

Table 1: Paired sample t-tests comparing observed to surrogate coherence and to virtual pairs coherence.

<table>
<thead>
<tr>
<th>Freq. Scale</th>
<th>Observed Coherence</th>
<th>Surrogate Coherence</th>
<th>Virtual Pairs Coherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>.25s</td>
<td>0.84 (.04)</td>
<td>0.30 (.02)***</td>
<td>0.72 (.05)***</td>
</tr>
<tr>
<td>.5s</td>
<td>0.93 (.03)</td>
<td>0.54 (.09)***</td>
<td>0.79 (.17)***</td>
</tr>
<tr>
<td>1s</td>
<td>0.42 (.06)</td>
<td>0.29 (.02)***</td>
<td>0.30 (.01)***</td>
</tr>
<tr>
<td>2s</td>
<td>0.37 (.05)</td>
<td>0.33 (.03)***</td>
<td>0.30 (.03)***</td>
</tr>
<tr>
<td>4s</td>
<td>0.28 (.02)</td>
<td>0.29 (.01)</td>
<td>0.26 (.02)***</td>
</tr>
<tr>
<td>8s</td>
<td>0.28 (.03)</td>
<td>0.29 (.04)</td>
<td>0.26 (.02)**</td>
</tr>
<tr>
<td>16s</td>
<td>0.30 (.04)</td>
<td>0.31 (.04)</td>
<td>0.29 (.04)</td>
</tr>
<tr>
<td>32s</td>
<td>0.32 (.07)</td>
<td>0.31 (.07)</td>
<td>0.29 (.06)*</td>
</tr>
<tr>
<td>1m</td>
<td>0.33 (.08)</td>
<td>0.34 (.08)</td>
<td>0.31 (.07)</td>
</tr>
<tr>
<td>2m</td>
<td>0.38 (.13)</td>
<td>0.36 (.11)</td>
<td>0.33 (.11)</td>
</tr>
<tr>
<td>4.5m</td>
<td>0.48 (.18)</td>
<td>0.43 (.21)</td>
<td>0.36 (.17)*</td>
</tr>
</tbody>
</table>

Note. Values are mean and standard deviation. * p < .05; ** p < .01; *** p < .001

To better understand the relationship between coherence and performance, we present two examples of cross-wavelet coherence plots in Figure 1. The top example is derived from the top performing team and the bottom example is derived from the lowest performing team. The y-axis corresponds to the frequency scale (which when divided by 8 can be related to time in seconds). The x-axis corresponds to the time on task with each point corresponding to 1/8 of a second (or one video frame). The colors correspond to the amount of coherence with warmer colors indicating high coherence. Arrows indicate phase relationships with right arrows conveying in-phase and left arrows conveying anti-phase. Arrows shifted up or down convey a lag in the oscillations between participants.

Next, we turn to RQs 1 and 2. In our first model, we included the four scales that were significantly more coordinated than expected due to chance alone as predictors of CPS performance (.25s, .5s, 1s, and 2s). Overall, this model accounted for a significant 30.2% ($R^2_{adj} = .226; F(4,37) = 4.00, p = .009$) of the variability in CPS performance with coherence at the .25s ($\beta = -.584, p = .017$) and 1s scales ($\beta = .789, p = .003$) as significant predictors of performance. The .5s and 2s scales were not significant ($p > .05$).

Regarding RQ1, these results suggest that whereas stronger movement coordination at the 1s scale is a strong predictor of better CPS performance, stronger coordination at the .25s scale is associated with poorer performance.

Next, we sought to better understand the form of coordination at these scales. Thus, we conducted a second model that included coherence as well as relative phase at .25s and 1s scales. This model accounted for a significant 34.6% ($R^2_{adj} = .276; F(4,37) = 4.90, p = .003$) of the variability in CPS performance with coherence ($\beta = -.614, p = .007$) and relative phase ($M = 3^\circ, SD = 2^\circ; \beta = -.294, p = .038$) at the 1s scale as significant predictors of performance. Now, however, coherence at .25s was not significant ($\beta = -.412, p = .06$) nor was relative phase at .25s ($p > .05$). Thus, these results suggest that movement coordination at the 1s scale is a primary predictor of performance and further, that relative phase values at the 1s scale closer to 0° (more in-phase) are associated with better CPS performance.

**Discussion**

In this work, we investigated the multiscale, movement coordination dynamics that emerge in computer-based CPS. We found that movements in .25s-2s scales were significantly more coordinated than chance and that all but the 16s, 1m, and 2m scales were more coordinated than expected due to task demands. We also observed that where coordination was greater than chance, both .25s and 1s were associated with CPS performance. However, when also accounting for relative phase, it appeared that higher in-
phase coordination at the 1s scale was the best predictor of CPS performance. Thus, some significant variability in CPS performance, in this context, appears to be explained by specific, low-scale patterns of coherence.

Given the low specificity of the movement data extracted from video, the question remains as to what is coordinated at these low scales and why they matter. In general, interactional phenomena that play out on (and below) a 0.25s timescale differ qualitatively from phenomena at a 0.5s timescale and beyond. For example, how interlocutors orient to each other’s behavior as meaningful for the interaction depends on timing. Short pauses in interaction (typically < 0.25s) are treated as idiosyncratic variation in speech; pauses around 0.5s mark a transition space where the next speaker can take the word; and longer pauses (> 1s) “are often treated as flagging something unusual or troublesome about the interaction” (Mushin & Gardner, 2009, p. 2035). Although, in addition to capturing these aspects of dialog, the observed coordination also captures mouse and keyboard movements, which likely unfold at these low scales as well. In general, many of the modalities captured by our movement measure can be argued to be task relevant as they capture dialogical events and computer input required for collaboration, but future work could consider more specific modalities such as how mouse movements are coordinated.

As far as future work is concerned, it is important to note some observable differences in coherence between the high and low performing teams in Fig. 1. There appear to be differences at higher scales, although average coherence was not generally above chance at these scales. However, we can speculate that participants’ performance may reflect their ability to create functional coherence across scales (e.g., between bodily ability and task demands), which could be assessed with fractal analyses (Davis et al., 2016). Further, it may be that successful CPS performance relies on higher-order transitions (Wiltshire et al., 2017) in coordination at one or more slow scales, as could be tenuously suggested by the pattern of high-low-high coherence near the 2m scale across the duration of the task. Thus, future work should also consider extracting not only specific scales, but also time ranges that could be theoretically important to CPS.

More generally, research of this nature is important because it advances an efficient means of unobtrusively examining coordination processes during collaboration with a goal of working toward systems that can elicit forms of coordination that enable effective collaboration (Fiore & Wiltshire, 2016; Kim, Chang, Holland, & Pentland, 2008; Wiltshire & Fiore, 2014). However, more work is necessary to understand if movement coordination is related to CPS performance in larger teams (de Montjoye, Stopczynski, Shmueli, Pentland, & Lehmann, 2014), with different roles and disciplinary expertise (Bergmann, Dale, Sattari, Heit, & Bhat, 2016), and when the teams are not co-located. Of course, such pursuits may require considering alternative modalities in which multiscale coordination might also occur. We expect that such endeavors are essential to advancing our knowledge of the way that coordination during human interaction relates to collaborative cognition.

Acknowledgments

This research was supported by Grant CF16-0455 from the Carlsberg Foundation awarded to Travis J. Wiltshire. We greatly appreciate the efforts of Stephen M. Fiore who was instrumental in the design and collection of data used for the present paper.

References


