TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 10 (2018) 36–54

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12304

This article is part of the topic "Coordination and Context in Cognitive Science," Christopher T. Kello (Topic Editor). For a full listing of topic papers, see http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview

Multi-Scale Contingencies During Individual and Joint Action

J. Scott Jordan,[a] Daniel S. Schloesser,[a] Jiuyang Bai,[a] Drew Abney[b]

[a]*Department of Psychology, Illinois State University*
[b]*Cognitive and Information Sciences, University of California, Merced*

Received 18 August 2016; received in revised form 26 May 2017; accepted 23 August 2017

Abstract

The present paper describes a joint action paradigm in which individuals or pairs utilized two computer keys to keep a dot stimulus moving inside a larger rectangle. Members of a pair could neither see nor hear each other. This paradigm allowed us to combine the discrete-trial type dependent variables (e.g., reaction time) commonly utilized by representational theorists, with the continuous, temporal dependence variables (e.g., RQA) utilized by dynamical theorists. Analysis revealed that individuals kept the dot in the rectangle longer than dyads and did so by moving it back and forth within the rectangle. Dyads, however, pressed their individual buttons as quickly as possible in order to keep the dot near the middle of the rectangle. These findings indicate that joint action constitutes a multi-scale phenomena that is best investigated via multiple, complementary methodologies versus single-measure, competing theories.

Keywords: Coordination; Joint action; Multi-scale; Representation; Affordances; Explanatory pluralism

1. Introduction

When an individual engages in goal-directed behavior such as playing a video game, the spatiotemporal dynamics of the actor's limbs need to be coordinated so that they do not interfere with each other. In a video game example, certain outcomes on the screen

Correspondence should be sent to J. Scott Jordan, Department of Psychology, Campus Box 4629, Illinois State University, Normal, IL 61790-4620. E-mail: jsjorda@ilstu.edu

(e.g., a character jumps and performs a particular type of kick) often require a very sophisticated, coordinated sequence of movement dynamics in which the left and right hands generate quite different key press patterns, simultaneously. These patterns of movements must be *contingent* and *complementary*; that is, the two limbs produce movements that are dependent upon and shaped by each other (i.e., contingent), and they collectively give rise to patterns that ultimately lead to the goal (i.e., complementary).

For a long time, scientific accounts of how individuals are able to generate contingent, complementary actions have ranged from strongly internalist, representational accounts that attribute the vast majority, if not all, of the causal explanation to brain dynamics (Anderson, Byrne, Douglass, Lebiere, & Qin, 2004; Meyer & Kieras, 1997) to widely distributive, dynamical accounts that spread the causal explanation across multiple scales of activity (Turvey, 2007; Van Orden & Holden, 2002). This theoretical tension between representational and dynamical accounts of coordination is also present in accounts of joint action; that is, goal achievement in which two or more actors need to generate contingent, complementary actions simultaneously in order for the group to achieve a shared goal.

2. Representational and dynamical accounts of joint action

Internalist accounts of joint action often assume that collaborators are able to generate contingent, complementary actions because each member is able to *represent* or *simulate* the other members' goals and action options. Quite often, the experimental tasks utilized by these researchers involve discrete-trial paradigms that collect a single response per trial, and the analysis involves aggregate temporal measures (e.g., average reaction time). For example, the Simon effect is an increase in reaction times (RTs) that occurs because one dimension of a stimulus (e.g., color) primes one response (e.g., press left for red) while the stimulus's spatial location primes another (e.g., the red stimulus appears on the right). Sebanz, Knoblich, and Prinz (2003) created a joint Simon task in which pairs seated next to each other saw a picture of a hand with a ring on it (the ring was red or green), and the hand was pointing to the left or the right. Each member pressed the button to their front, dependent on the color of the ring, and both members responded to a different color. Sebanz et al. reported a joint Simon effect, in that they found RTs were longer when the direction of the finger (i.e., right or left) and the direction indicated by the ring color (i.e., left or right) were different (i.e., *incompatible*) than when they were the same (i.e., *compatible*). However, this effect did not occur if participants completed the same task alone. These findings imply that the compatibility between finger-pointing direction and ring color only mattered to participants when the action option the participant did not control was actually produced by another person. Sebanz et al. accounted for this joint Simon effect by proposing that as participants completed the task with another person, they automatically represented the action option of the other participant, what has come to be known as *action co-representation*.

In contrast to internalist-driven methodologies, researchers who propose non-internalist, dynamical accounts of joint action often take continuous measurements of ongoing movement—that is, they sample movement dynamics at an extremely fast rate (e.g., 60 Hz)—and then run analyses on the degree to which the quantified movement dynamics of one individual are contingently coupled with the movement dynamics of another. In a now classic example of this methodology, Richardson, Marsh, Isenhower, Goodman, and Schmidt (2007) asked pairs of participants to sit side by side and rock in in-phase synchrony (i.e., rocking in the same direction simultaneously) or antiphase synchrony (i.e., rocking in opposite directions simultaneously) as they fixed their gaze on a red X that was situated on either the armrest of their partner's chair (focal group) or on the wall directly to their front (peripheral group). The results revealed that individuals were able to coordinate their rocking chair movements in all conditions. Experiment 2 included a *no information* condition in which participants fixed their gaze on a red X situated on a wall that was on the opposite side of the participant as the other participant (e.g., if the other participant was on their right, the red X was located on a wall to the left). Under these conditions, members of a pair did not replicate the in-phase and antiphase dynamics found in Experiment 1. Richardson et al. (2007) accounted for these findings in a non-internalist fashion by modeling the participants as coupled oscillators whose dynamics synergistically constrained each other in a continuous, contingent manner. According to this account, participants did not achieve synchrony in the *no information* condition because the movement dynamics of the two participants were effectively *de-coupled* by the removal of movement information regarding the other.

3. Competing theories or complementary methods?

Representational and dynamical accounts of joint action are often presented as competing explanatory frameworks. This might stem from their shared, perhaps implicit assumption that successful joint action requires co-actors have access to the *action* options of their partner, either in the form of internal neural representations or external informational structures (i.e., affordances). In contrast to this competitive take on the issue, some researchers model different methods as being complementary (Abney et al., 2014; Cooney & Troyer, 1994; Gilden, 1997; Jordan, 2013; Van Orden, Holden, & Turvey, 2003, 2005). From this perspective, what Abney et al. refer to as *explanatory pluralism*, the phenomena typically investigated by cognitive scientists, including joint action, tend to be multi-scale in nature, in the sense the phenomena are constituted of events taking place at multiple time scales (e.g., neural, behavioral, perceptual, and environmental), simultaneously. As a result, measures that sample regularities at different time scales should potentially compliment each other in an explanation of the multi-scale phenomenon being investigated. According to Abney et al., "the issue is *which scales* and *which theoretical frameworks* are relevant for the question at hand, and *how they relate* to each other" (p. 2).

As an example of explanatory pluralism, Abney et al. utilized data from a perceptual decision-making task (Bahrami et al., 2010) in which pairs of participants jointly determined which of two, briefly presented (i.e., 85 ms) images, each showing six Gabor patches, contained one Gabor patch whose contrast differed from the others. If participants indicated the same image (i.e., Image 1 or Image 2), they preceded immediately to the next trial. If they indicated different images, they were allowed to negotiate a joint decision via open verbal communication. One analysis of these data aggregated across all the decisions made by each individual as well as each dyad and revealed that dyads came to perform better than their individual members, but only if the members were equally good at the task. A second analysis (Fusaroli, Rączaszek-Leonardi, & Tylén, 2014) coded the verbal responses in terms of the type of confidence they revealed (certainty vs. uncertainty) and the degree of alignment pairs generated in their use of confidence types over the course of the experiment. Results revealed dyads performed better when they consistently and continually used similar types of confidence ratings, indicating that better linguistic alignment over time led to better performance. Finally, a third analysis (Fusaroli, Abney, Bahrami, Kello, & Tylén, 2013) investigated the relationship between dyadic complexity matching (CP) and performance. This was accomplished by coding the speech onsets and offsets for each member of a dyad individually and then determining the complexity (i.e., the type of contingency) entailed in each member's onset-offset pattern. Analysis revealed that dyads whose onset-offset complexity correlated most highly also performed better on the task.

According to Abney et al. (2014), the three above-mentioned analyses of the Bahrami et al. (2010) study demonstrate how analyses involving course-grained measures such as overall performance can actually complement more fine-grained analyses such as complexity matching. As a result, the different measures do not reflect opposing theoretical accounts as much as they reflect different ways of sampling the multi-scale dynamics that constitute a complex phenomenon such as joint perceptual decision making.

In light of the notion of explanatory pluralism, we present a joint action experiment in which participants used two computer keys to move a circular dot stimulus back and forth within a larger rectangle depicted on a computer screen. Participants completed the task either by themselves (i.e., the Individual condition) or with a partner (i.e., the Group condition). In the Individual case, participants controlled both computer keys, while in the Group condition each participant controlled only one of the two keys. In addition, participants in the Group condition were located in separate rooms and could neither see nor hear each other.

Because we were able to assess the duration of each key press event that occurred as participants controlled the stimulus, we were able to generate aggregated, RT-like variables, such as the average amount of time participants were able to keep the dot in the rectangle for a given button press. In addition, because these consecutive key press events followed each other in time we were able to analyze the extent to which the pattern of key press events generated by the two keys was coordinated and, potentially, contingent.

Given that members of dyads had no information regarding the actions of their partners, from the perspective of the representational-affordance debate, the dyad condition might seem trivial, as one assumes that dyads will not be able to do the task at all since they will not have access to their partner's actions via either neural representations or affordances. From the perspective of explanatory pluralism, however, control of the stimulus trajectory is a multi-scale phenomenon in which the events at different scales (e.g., the moment of each individual button press, the impact that collective button presses have on the movements of the dot, and the changes in button-press/movement patterns that emerge over time) interact continuously. As a result, contingencies may arise between more than just key presses. They might also emerge between variables that exist at different time scales (e.g., patterns of button presses, resultant patterns of stimulus movements, and overall performance). If the latter option proves true, it will not be possible to account for the results solely in terms of either internal representation construction or external affordance detection. Rather, all relevant levels of scale, from the neural to the environmental, will need to be included in the explanation.

4. Method

4.1. Participants

There were 66 participants in the present experiment. Twenty-four were in the Individual condition. Participants were recruited from the participant pool of a large Midwestern university and were given course credit for participating. Consent was obtained from each participant, all of whom ranged in age from 18 to 25.

4.2. Apparatus

Participants sat roughly 60 cm (23.62 in) in front of a computer monitor and controlled the position of a dot stimulus (i.e., 50 × 50 pixel white dot containing a smaller, 10 × 10 pixel black dot centered within the larger white dot). As seen in Fig. 1, a stationary, black rectangle was situated in the center of the screen. The

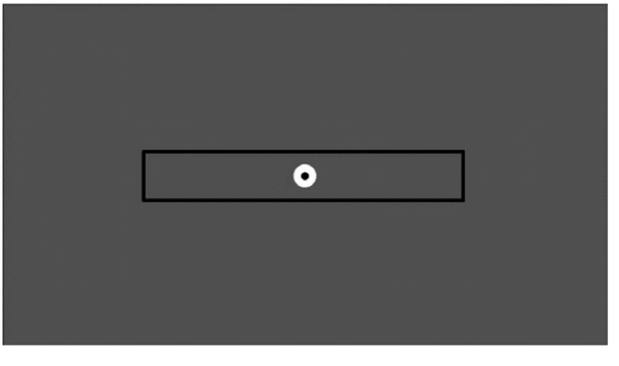

Fig. 1. A depiction of the experimental task space.

rectangle was 600 pixels wide and 100 pixels tall, with edges comprised of 10 pixel-wide black lines.

Participants controlled the dot by pressing two computer keys. Pressing the A-key incremented the dot's horizontal position to the right at a constant velocity of 375 pixels/s, as long as the key was depressed. Pressing the L-key did the same in the opposite direction. As participants controlled the dot and attempted to keep its movements within the rectangle, it became necessary to change its direction of movement (i.e., from left to right or right to left). Doing so proved challenging, for if the computer detected that both keys were being pressed simultaneously, as can be seen in Fig. 2, the dot moved both upward and in the direction it was previously moving before both keys were pressed, at a constant velocity (i.e., 375 pps) in both the *X* and *Y* dimension.

If the computer detected that neither key was being pressed, the dot moved both downward and in the direction it was already moving at a constant velocity (i.e., 375 pps) in both dimensions.

It is the overlapping key events (i.e., both keys being pressed or neither key being pressed) that made the task difficult, for if either type of overlapping event continued for too long (e.g., a few hundred milliseconds) the dot moved outside the perimeter of the rectangular box. Thus, to keep the dot moving back and forth within the rectangle, participants had to avoid producing overlapping key events, particularly as they attempted to change the dot's direction. However, because the status of the keyboard was sampled every 16 ms, it was practically impossible for participants to completely avoid overlapping key events.

4.3. Procedure

Participants were read their rights as participants and were asked if they had any questions. If not, they were asked to sign an informed consent form.

Participants were pseudo-randomly assigned to either the Individual or Group condition. Assignment was pseudo random because of the need to ensure similar group sizes.

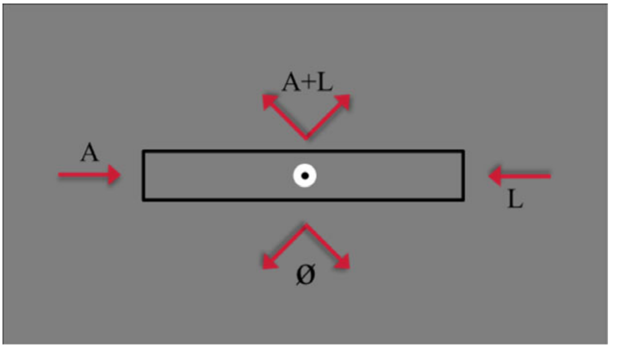

Fig. 2. The individual-press (i.e., A-key and L-key) and overlapping-press (A + L keys and no-key) functions.

Participants in the Individual condition completed the dot control task alone. Group participants completed the task in pairs in which each member controlled only one of the two keys (i.e., either the L or the A-key). Additionally, Group participants completed the task in separate lab rooms, which prevented them from seeing or hearing each other.

After being assigned to a condition and escorted to the appropriate lab space, participants were read instructions about the task (i.e., keep the dot in the rectangle) and the functions of the key they controlled. Participants in the Group condition were told only about the specific key they would control throughout the dot control task. Group participants were further told (a) their partner in the other room would be controlling a key and (b) that key would have a certain type of effect on the dot. However, they were not told the specific key their partner would be using. This manipulation was created so that participants would not attempt to use the other control-key during the course of the task.

Participants then pressed the Enter key. Following a visible 3-s countdown, they completed a single, 8-min trial. Subsequent key presses altered the dot's movements in the above-mentioned fashion, with the dot's velocity being 375 pps. Upon completion of the trial, participants were then thanked for their time, de-briefed about the experiment, and awarded extra-credit for their participation.

5. Results

5.1. Aggregated, RT-like, performance variables

Every time the state of the computer keyboard was changed (i.e., a key was pressed or released), the program recorded the *X–Y* coordinates of the dot as well as the time of the event. Given these moments involve both key *presses* and key *releases,* we refer to them as key *events*.

Key events were coded as occurring either in the rectangle (i.e., an in-box event) or out of the rectangle (i.e., an out-box event). Fig. 3 illustrates the spatial distribution of events on the computer screen, in relation to the location of the rectangle, for the best and worst individuals and dyads. In-box events were coded in terms of the quadrant of the target area in which they occurred, with the left-most 150 pixel quadrant being labeled Quadrant 1, and the remaining being labeled consecutively, from left to right. We also coded each in-box key event in terms of whether or not the dot changed its direction on that event. A key event was coded as a *turn press* if the *x*-coordinate of the event was larger, or smaller, than the *x*-coordinate of both the preceding and following key events. Finally, A-key events and L-key events were coded as *alone* presses, while AL and no-key events were coded as *together* press.

In-box time is the percentage of *total time* the dot was inside of the rectangle. We then coded each key event in terms of whether it occurred in the first 2 min of the 8-min trial, or in one of the subsequent 2-min portions, resulting in four, 2-min intervals we refer to as a block. *Coast time* is the amount of time, in milliseconds, that *followed* a particular

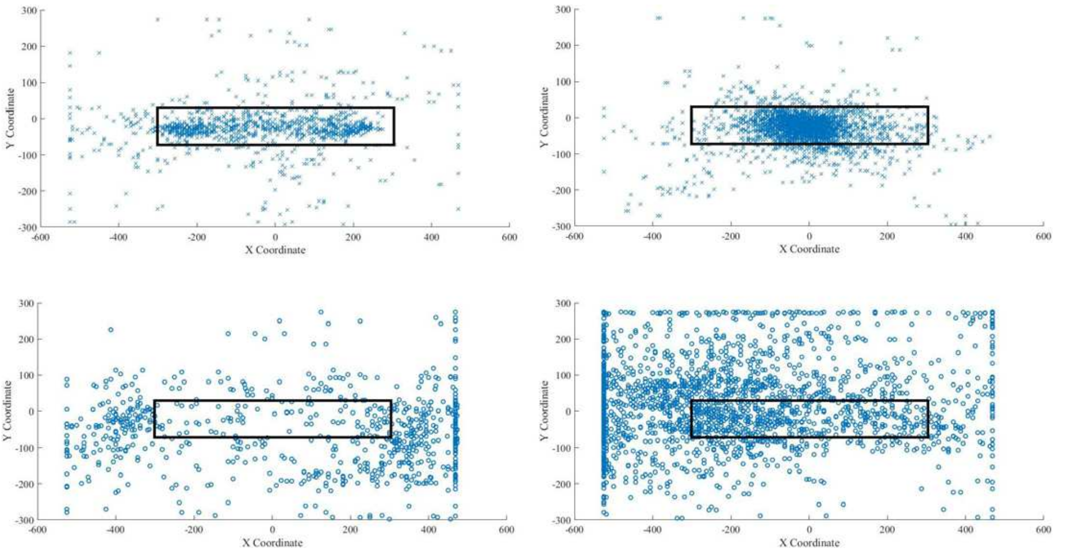

Fig. 3. X–Y plots of the spatial distribution of event locations on the computer monitor. Each dot represents the location of a change in key press event. Top left: Individual having the highest in-box time. Bottom left: Individual having the lowest in-box time. Top right: Dyad having the highest in-box time. Bottom right: Dyad having the lowest in-box time.

in-box, key event. Coast time allowed us to assess the extent to which participants developed the ability to let the dot coast back and forth across the inside of the rectangle.

5.2. In-box time

A mode (Individual vs. Dyad) by block (1–4) mixed ANOVA in which block was a within-subject variable, yielded a main effect of mode, $F(1, 43) = 10.21$, $p = .003$, partial $\eta^2 = 0.19$. The average in-box time of individuals ($M = 53.15$, $SD = 21.06$, $n = 24$) was significantly longer than that of dyads ($M = 34.35$, $SD = 17.98$, $n = 21$). There was also a main effect of block, $F(3, 129) = 10.38$, $p < .001$, partial $\eta^2 = 0.19$. As seen in Fig. 4, the average in-box time increased from block 1 to block 4. There were no other statistically significant effects.

5.3. Coast time

A mode (Individual vs. Dyad) by block (1 through 4) by quadrant (1 through 4) by type of key event (alone vs. together) mixed ANOVA (block, quadrant, and type of key event were within-subject variables) on coast time revealed a main effect of mode, $F(1, 32) = 125.64$, $p < .001$, partial $\eta^2 = 0.80$, and a main effect of block, $F(3, 288) = 5.94$, $p = .001$, partial $\eta^2 = 0.16$. As can been seen in Fig. 5, individuals ($M = 443.89$, $SD = 81.12$) had significantly longer coast ties than dyads ($M = 161.96$, $SD = 67.92$),

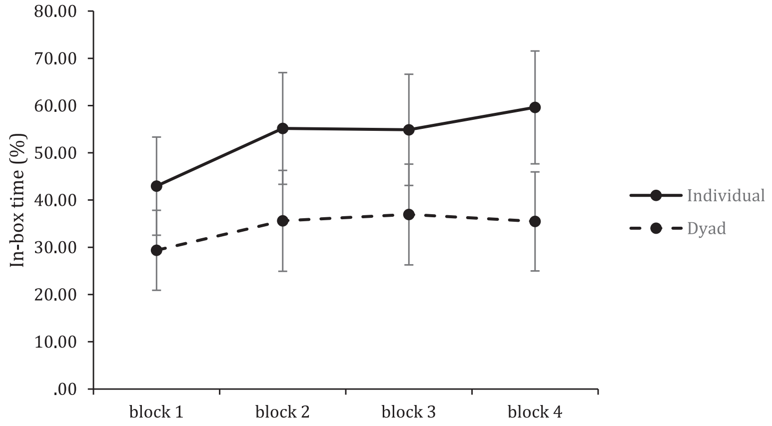

Fig. 4. In-box time of individuals and dyads as a function of block. The length of the error bars equals one standard deviation.

and overall average coast time increased significantly across blocks. There was also a main effect of quadrant, $F(3, 288) = 95.23$, $p < .001$, partial $\eta^2 = 0.75$. As seen in Fig. 6, coast times were significantly larger when they followed a key event produced in the outer two quadrants (i.e., quadrants one and four). Finally, there was a main effect of type of key press, $F(1, 288) = 98.31$, $p < .001$, partial $\eta^2 = 0.75$. The average coast time of alone presses ($M = 472.34$, $SD = 314.89$) was significantly longer than that of together press ($M = 151.51$, $SD = 82.33$).

As regards interactions, the ANOVA also revealed a host of two-way interactions, including an interaction between mode and quadrant, $F(3, 288) = 49.93$, $p < .001$, partial $\eta^2 = 0.61$, mode and type of key press, $F(1, 288) = 67.03$, $p < .001$, partial $\eta^2 = 0.68$,

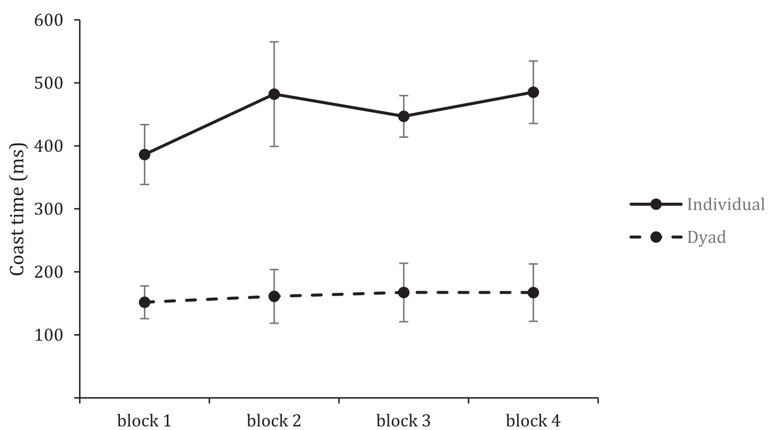

Fig. 5. Coast time of individuals and dyads as a function of block. The length of error bar is one standard deviation. The length of error bar is one standard deviation.

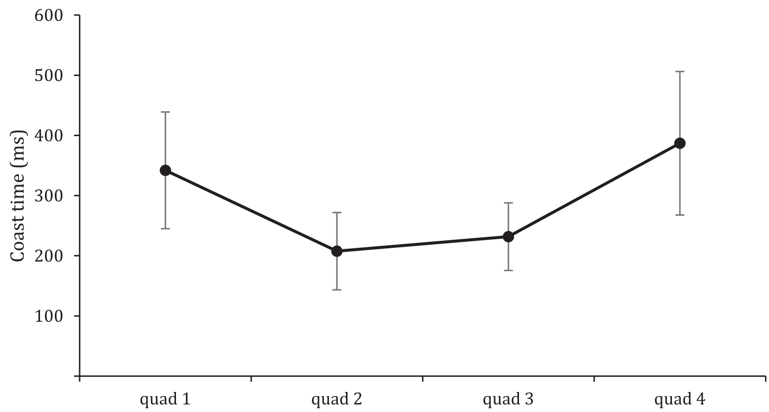

Fig. 6. Coast time as a function of quadrant. The length of error bar is one standard deviation.

block and type of key press, $F(3, 288) = 13.97$, $p < .001$, partial $\eta^2 = 0.30$, and quadrant by type of key press, $F(3, 288) = 82.93$, $p < .001$, partial $\eta^2 = 0.72$.

All of the above-mentioned two-way interactions were involved in at least one of two, three-way interactions. We therefore illustrated the two three-way interactions only. These include an interaction between mode, block, and type of key press, $F(3, 288) = 6.26$, $p = .001$, partial $\eta^2 = 0.16$. As seen in Fig. 7, for individuals, coast times for *alone* presses increased over blocks, and those for *together* presses decreased, while for dyads, coast times did not vary with either block or type of key press.

The second three-way interaction was between mode, quadrant, and type of key press, $F(3, 288) = 60.25$, $p < .001$, partial $\eta^2 = 0.65$. As seen in Fig. 8, participants in the individual condition generated their largest coast times for alone presses that occurred in the outer two quadrants of the rectangle (i.e., quadrants one and four). Coast times for together presses were significantly smaller than coast times for alone presses, and they did not vary significantly across quadrants. For dyads, coast times did not vary with block or quadrant.

These differences between individual and dyad coast-time patterns are further clarified by the frequency of alone and together presses. First, as can be seen in Table 1, Dyads produced significantly more in-box key events (i.e., alone and together presses) than Individuals, which is consistent with the finding their coast times are significantly shorter than those of individuals. Second, for both groups, roughly 50% of the in-box key events were alone presses (i.e., either the A- or L-key alone), with the remainder being together presses (i.e., either neither key was pressed, or both keys were). Despite this even split between together and alone presses, however, 94.1% of the coast time produced by Individuals resulted from alone presses, while the percentage of coast time was pretty much evenly distributed between alone and together presses for dyads. These data imply that individuals used A and L events to move the dot back and forth within the rectangle, while for dyads, alone events were not functionally different from together events.

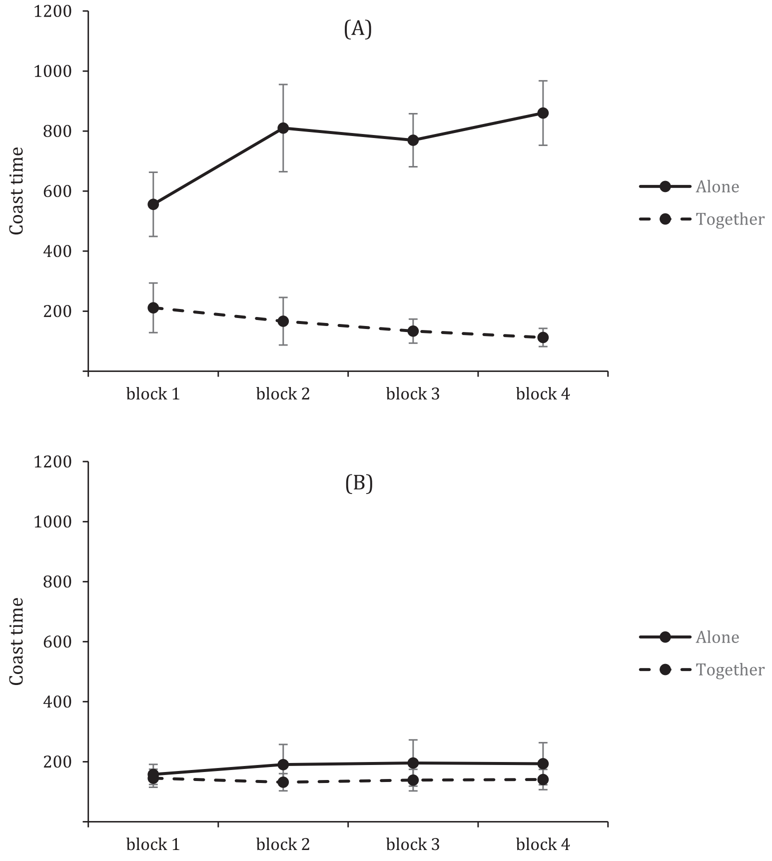

Fig. 7. A (top) illustrates coast time as a function of type of key press (i.e., alone or together) and block, for individuals, while B (bottom) does so for dyads. The length of the error bar is one standard deviation.

This interpretation is consistent with correlational analyses, which revealed that coast time and in-box time were positively correlated for individuals overall, $r(24) = .51$, $p < .05$, and negatively for dyads in Block 4, $r(22) = .467$, $p < .05$. Since individuals used alone presses to make the dot move back and forth, in-box time was larger for individuals who generated *larger* coast times. Since dyads seem not to have discriminated alone from together presses, by Block 4, in-box time was larger for dyads whose coast times were shorter; that is, they pressed their individual buttons faster.

5.4. Turning point

A mode (Individual vs. Dyad) by block (1–4) by quadrant (1–4) mixed ANOVA on turning point yield a main effect of quadrant. However, because quadrant was defined by horizontal location, the main effect of quadrant on turning point is actually an artifact of creating the quadrant variable. The ANOVA also revealed an interaction between mode and

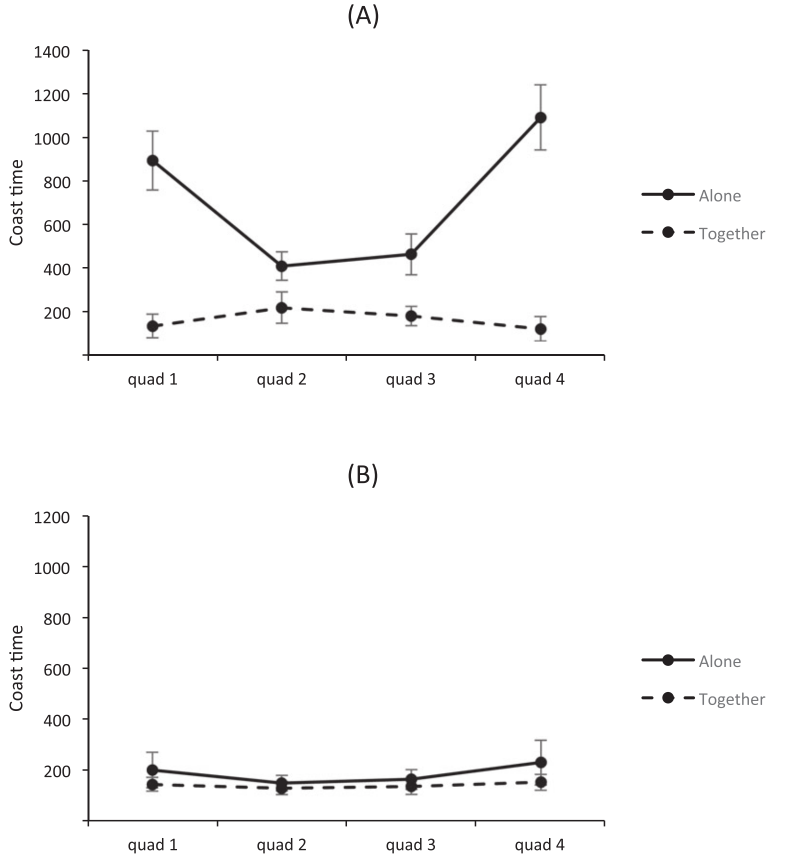

Fig. 8. A (top) illustrates coast time as a function of type of key press (i.e., alone or together) and quadrant, for individuals, while B (bottom) does so for dyads. The length of the error bar is one standard deviation.

quadrant, $F(3, 252) = 22.97$, $p < .001$, partial $\eta^2 = 0.45$. As seen in Fig. 9, participants in the Individual condition produced turn points that were nearer to the outer edge of the quadrant, regardless of quadrant.

5.5. *Non-linear analyses (CRQA for events series of button presses)*

We conducted categorical cross-recurrence quantification analysis (C-CRQA) on the time series of consecutive A and L presses to examine the dynamic means by which aggregate outcome variables such as coast time and in-box time were produced. C-CRQA examines the degree to which two time series of categorical or continues events (e.g., A and L button presses in the present experiment, or eye movements made by two interacting persons) are organized in relation to each other over time. Richardson and Dale (2005), for example, found that during conversation, the eye-movement patterns produced by the listener follow the patterns produced by the speaker, with a time lag of roughly

Table 1
Percentage coast time and event totals and percentages for alone and together presses by individuals and dyads

	Individuals		Dyads	
	Key Presses			
Variables	Alone	Together	Alone	Together
% Coast time	94.1%	5.7%	65.9%	34.0.3%
Event count	5,124	4,183	18,192	14,324
Event percentage	56%	44%	56%	44%

Note. % Coast time and Event percentages do not equal 100% due to key press mistakes (e.g., participant presses the m key accidently, instead of the L-key).

2 s. In short, the eye-movement dynamics of the participants were *recurrent* with a time lag of roughly 2 s.

In the present experiment, we wanted to determine the extent to which A presses and L presses were recurrent and at what, if any, time lag. If participants learn to make the dot move back and forth within the rectangle, A and L presses should not co-occur during coast times, as participants are holding down only one of the two buttons. In addition, A and L presses should predominantly co-occur as participants change the dot's direction (i.e., when switching from an A press to an L press, or vice versa). As a result, A and L presses should most often co-occur with a time lag roughly equal to the average coast time.

To measure the recurrence between the A- and L-keys, we created separate event series of button presses ("A" button and "L" button) for each individual and each dyad. We then submitted the two time series from each trial to a C-CRQA and obtained a diagonal cross-recurrence profile (DCRP; Abney, Warlaumont, Oller, Wallot, & Kello, 2016; Abney, Paxton, Dale, & Kello, 2015; Coco & Dale, 2014; Dale, Warlaumont, &

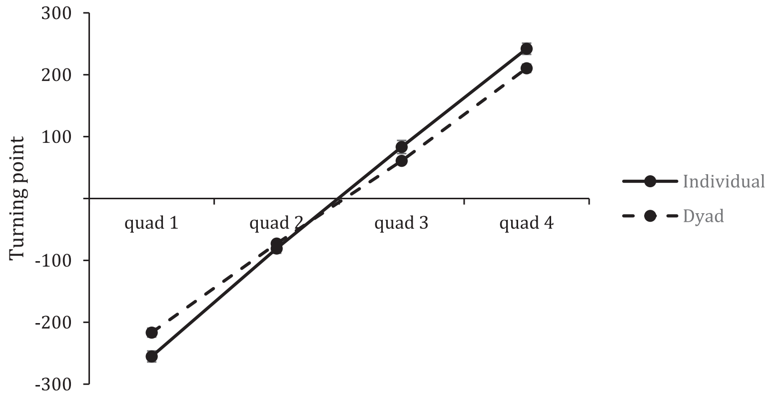

Fig. 9. Turning point as a function of mode and quadrant. The length of the error bar is one standard deviation.

Richardson, 2011; Richardson & Dale, 2005; Warlaumont, Richards, Gilkerson, & Oller, 2014). C-CRQA is distinct from CRQA correlation in a variety of ways. Most important to the current study, C-CRQA does not consider the absence of a behavior across two time series to count towards a "shared absence" and therefore contribute to the anti-correlation as is done in cross-correlation. As a result, all samples in a time series without an A or an L were coded as different values for the two time series.

Based on the observed average coast time and in-box time, we measured the temporal coordination between the time series within a 3 s window. As a result, the DCRP provided a profile of the recurrence rate at each lag from −3 s to +3 s in 100 ms increments. Recurrence rate is thus the proportion of time the two time series co-occur at a certain time lag in the profile. From the DCRP, we derived two measures: MAXREC and MAXLAG. MAXREC is the maximum recurrence rate found in the DCRP. MAXLAG is the specific lag on the DCRP where MAXREC occurred. See Fig. 10 for the averaged DCRPs for the Individuals and Dyads.

A one-way ANOVA with MAXREC as the dependent variable and Mode (Individuals vs. Dyads) as the predictor variable suggested that there was a significant difference across Individuals ($M = 0.12$, $SD = 0.03$) and Dyads ($M = 0.16$, $SD = 0.05$), $F(1, 43) = 13.74$, $p < .001$. A one-way ANOVA with MAXLAG as the dependent variable and Mode as the predictor variable suggested that there was a reliable difference for Individuals ($M = +928$ ms, $SD = 8.16$) and Dyads ($M = -430$ ms, $SD = 15.49$), $F(1, 43) = 10.76$, $p = .002$.

Regarding the significant difference in MAXLAG between the Individual and Group conditions, Fig. 10 reveals that for the Dyad aggregate DRCP there is a slight, bimodal distribution. It seems the average MAXLAG of −428 ms essentially reflects a "washing

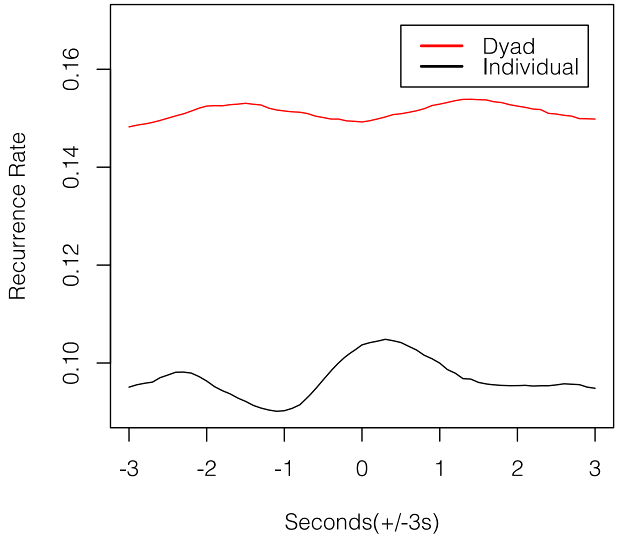

Fig. 10. Diagonal cross-recurrence profile averaged across Individuals and Dyads.

out" because of this bimodality. For the Individual aggregate DRCP, the average MAXLAG of +928 ms, while not exactly reflecting the largest DRCP in the figure, is, nonetheless, much closer to the peak value than was the case in the Group data. This implies there may be only one local maxima, which indicates that for participants in the Individual condition, A events preceded L events more frequently with a lag of 928 ms. This is the type of MAXLAG that would emerge if one key event (e.g., an A event) reliably preceded another event (e.g., an L event), and this is the type of pattern that has to be in place if one is to move the dot back and forth across the screen consistently. For Dyads, the bimodal relationship seems to imply that both the A and L buttons led each other. In other words, A events and L events tended to occur at a somewhat reliable temporal distance from each other, but there is no pattern as to which one tended to precede the other. This is the type of pattern one would expect if participants were basically pressing their keys as quickly as possible (i.e., at a rate of roughly five presses per second).

To determine if MAXREC or MAXLAG were functional for the task across group type, we performed separate one-way ANOVAS with task performance (i.e., in-box time) as the dependent variable and Mode and MAXREC or Distance from Synchronization as predictor variables. Distance from Synchronization (DFS) was derived from MAXLAG by computing the absolute difference between MAXLAG and the index of synchronization in the DCRP. Smaller values of DFS suggest that MAXLAG was temporally near to the index of synchronization, whereas larger DFS values suggest that MAXLAG was temporally further away (in either direction) from the index of synchronization. For MAXREC, we observed no reliable effects for the main effects of MAXREC ($F[1, 41] = 1.65$, $p = .21$), Mode ($F[1, 41] = 0.16$, $p = .72$), nor the interaction, $F(1, 41) = .01$, $p = .94$. For DFS, we observed no reliable effects for the main effects of DFS ($F[1, 41] = 1.19$, $p = .28$), Mode ($F[1, 41] = 1.37$, $p = .25$), nor the interaction, $F(1,41) = 0.04$, $p = .84$.

6. Discussion

Overall, the analysis of in-box time, coast time, and turning points indicates that individuals maximized their ability to keep the dot in the rectangle by learning how to make it coast back and forth the inside of the rectangle. This pattern gave rise to significantly larger in-box times, as well as significantly longer coast times that increased across the four blocks, and were larger for *alone* presses occurring in the outer quadrants of the rectangle. It also gave rise to the positive correlation between coast time and in-box time for individuals.

The idea that individuals learned to move the dot back and forth within the rectangle is further supported by the C-CRQA. As seen in Fig. 10, MAXLAG was +928 ms. This means that A and L events were most likely to recurse on each other with a time lag of just under 1 s. This interval is roughly equal to the average coast time of *alone* presses produced in the outer two quadrants of the rectangle (see Fig. 8A), and it indicates that

A and L events were most likely to co-occur as individuals were attempting to switch the direction of the dot by switching from an A press to an L press, and vice versa.

Collectively, these data imply that individuals were able to develop contingent coordination between A and L presses, for the only way to produce a long coast time (i.e., the time it takes the dot to move across the length of the rectangle) was to produce an alone press. Generating an alone press was manageable for individuals because the action options (i.e., the A- and L-keys) were controlled by a single person who had access to *when* each key would, or would not, be pressed. In short, the right hand knew what the left hand was doing.

As regards dyads, they produced coast times that equaled roughly five presses per second (i.e., 200 ms) when the dot was in the rectangle, and these coast times did not vary with type of key presses (i.e., alone vs. together), block, or quadrant. In addition, dyads produced over four times as many key events within the rectangle as individuals. Finally, dyads that pressed their buttons most quickly (i.e., with the shortest coast times) produced the largest in-box times. Collectively, these data indicate that members of pairs pressed their keys pretty much as quickly as possible when the dot was in the rectangle. As a result, their "coast times" are probably more a measure of how quickly they were able to generate consecutive key presses than a measure of how long they let the dot coast within the rectangle. This idea finds further support in the fact that MAXREC was significantly higher for dyads than for individuals. Given that dyad members were pressing their buttons so quickly, A and L presses were significantly more likely to co-occur, at all time lags, than was the case for individuals, who were actually attempting to avoid A and L co-occurrences.

While the dyad data do indicate that the key events produced by pair members were not contingent on each other (i.e., they pressed their individual keys as quickly as possible, with little, if any, regard for the presses of their partner), the negative correlation between coast time and in-box time indicates that dyad presses may have been complementary. That is, dyads that produced faster button presses (i.e., shorter coast times) kept the dot in the rectangle longer (i.e., larger in-box times). This negative correlation implies that producing button presses as quickly as possible was a goal-directed behavior. Specifically, rapid button presses by both partners would have resulted in rapidly changing key events (e.g., a 200 ms A event, followed by a 100 ms AL event, followed by a 100 ms "no key" event), which may have served to keep the dot oscillating rapidly back and forth, and up and down, near the middle of the rectangle. Such a pattern of dot movements would have emerged spontaneously as dyads pressed their individual buttons in the attempt to get a grip on the movements of the dot (see Fig. 3, top right, for an example). While this is not the same strategy that was used by participants in the Individual condition, it still constitutes a strategy. Thus, while dyad key presses were not contingent on each other, they seem to have been contingent on the larger time-scale pattern of dot movements produced by the dyad's combined presses. As a result, they seem to have been complementary, in that, participants pressed their key as quickly as possible in order to keep the dot oscillating near the middle of the rectangle.

For those who are familiar with video games, this pattern of behavior is often referred to as *key mashing*, in which one simply presses various control keys as quickly as possible in the hopes of generating useful effects on the screen. In the present experiment, the "hoped for" useful effect would have been to keep the dot near the middle of the rectangle. Thus, while the right hand did not know what the left hand was doing, each hand did, in fact, know what it was doing *in relation to the larger scale pattern of movement effects the collective button presses produced*.

Cleary, the task constraints gave rise to different patterns of behavior in the Individual and Group conditions. One could attempt to account for these differences in terms of an inability to generate internal, co-representations, or a lack of external information about partner dynamics. From either of these perspectives, the present results might seem trivial, the idea being that groups did worse because the left hand did not know what the right hand was doing.

While such an account is partially correct, it misses the fact that despite the inability of dyads to generate contingent button presses, they were still able to couple the timing of their button presses with the pattern of dot movements that was produced by their collective button presses, and persisted at a larger time-scale that of any single button press. Such a finding is consistent with the notion of *explanatory pluralism*. For since the task was multi-scale in nature (i.e., individual button presses, dot movement patterns, and changes in button press—dot movement patters across the experiment), different forms of analysis were appropriate for phenomena at different time scales (e.g., C-CRQA for button press recurrence, and coast time and in-box time for aggregate, larger time-scale outcome measurement).

To be sure, we did not find correlations between outcome measures (i.e., in-box time or coast time) and MAXREC. This lack of correlations, however, might provide an example of how different measures collected from the same phenomenon can be independent of each other. For example, the goal-directed behavior of dyads manifested as very rapid button presses that were independent of each other in the sense group members were simply pressing as quickly as possible, with no regard for their partner's presses. Given A and L presses were independent, there was no *pattern* of recursion between them. The two independent events were equally likely to co-occur at most time lags. Despite the independence of A and L presses, however, the faster they were made, the more rapidly a series of somewhat random key press events occurred (i.e., A, L, AL, or no-key), resulting in an oscillation of the dot near the middle of the rectangle.

The lack of correlation between MAXREC and performance measures is more difficult to explain for individuals. One would simply assume that those with larger coast times and larger in-box times would also have larger MAXREC. However, MAXREC can be positive or negative, depending on which event, A or L, tends to be leading the other. If some participants had negative MAXREC, while others has positive MAXREC, correlations with coast time and in-box time would be reduced. Despite this lack of correlations with MAXREC, in the end, all of these measures clearly complemented each other to explain how (a) individuals were able to achieve the goal of making the dot move back and forth within the rectangle, and (b) dyads pressed their buttons as quickly as possible in order to keep the dot near the center of the rectangle.

The notion of accounting for psychological phenomena in terms of multi-scale, contingent interactions dates as far back as Dewey's (1929) interactionist perspective, and it is also present in contemporary cognitive science (Gilden, 1997; Jordan, 2013; Jordan & Day, 2015; Oyama, 1985; Van Orden & Holden, 2002; Van Orden et al., 2003, 2005). Given the availability of such multi-scale theorizing, as well as techniques capable of addressing multi-scale dynamics, it seems increasingly difficult to provide explanations of psychological phenomena, let alone joint action, that rely on one level of explanation (i.e., neural representations) or assume that one level should be effectively removed from a multi-scale explanation (i.e., neural dynamics) because it has been inappropriately conceptualized in the past (i.e., neural representations).

To be sure, it may be the case that internalist and dynamicist theorists have been led to their theoretical positions by the assumptions that surround the dependent variables they measure. But given the present paradigm's ability to extract RT-like variables (i.e., coast time) from ostensibly continuous data, while simultaneously conducting multi-scale RQA analyses on those very same data, it seems unnecessary to assume RTs and RQA measure completely different phenomena or should lead to qualitatively different theoretical conclusions. On the contrary, one could argue, as do explanatory pluralists (Abney et al., 2014), that RTs and RQA simply constitute different means of sampling the same, multi-scale phenomenon.

References

Abney, D. H., Dale, R., Yoshimi, J., Kello, C. T., Tylén, K., & Fusaroli, R. (2014). Joint perceptual decision-making: A case study in explanatory pluralism. *Frontiers in Psychology*, *5*, 330. https://doi.org/10.3389/fpsyg.2014.00330.

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2015). Movement dynamics reflect a functional role for weak coupling and role structure in dyadic problem solving. *Cognitive Processing*, *16*(4), 325–332.

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2016). Multiple coordination patterns in infant and adult vocalizations. *Infancy*, *22*(4), 514–539.

Anderson, J. R., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060. https://doi.org/10.1037/0033-295X.111.4.1036.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science*, *329*, 1081–1085. https://doi.org/10.1126/science.1185718.

Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An R package. *Frontiers in Psychology*, *5*, 510.

Cooney, J., & Troyer, R. (1994). A dynamic model of reaction time in a short-term memory task. *Journal of Experimental Child Psychology*, *58*, 200–226.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. *International Journal of Bifurcation and Chaos*, *21*(04), 1153–1161.

Dewey, J. (1929). *The quest for certainty: The study of the relation of knowledge and action*. New York: Minton, Balch & Company.

Fusaroli, R., Abney, D. H., Bahrami, B., Kello, C. T., & Tylén, K. (2013). Conversation, coupling, and complexity: Matching scaling laws predicts performance in a joint decision task. Poster presented at the 35th Annual Conference of the Cognitive Science Society. Berlin, Germany.

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas Psychology*, *32*, 147–157. https://doi.org/10.1016/j.newideapsych.2013.03.005.

Gilden, D. L. (1997). Fluctuations in the time required for elementary decisions. *Psychological Science*, *8*(4), 296–301.

Jordan, J. S. (2013). The wild ways of conscious will: What we do, how we do it, and why it has meaning. *Frontiers in Psychology*, *4*, 574.

Jordan, J. S., & Day, B. (2015). Wild systems theory as a 21st century coherence framework for cognitive science. In T. Metzinger & J. M. Windt (Eds). *Open MIND*: 21(T). Frankfurt am Main: MIND Group. https://doi.org/10.15502/9783958570191

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I Basic mechanisms. *Psychological Review*, *104*(1), 3.

Oyama, S. (1985). *The ontogeny of information: Developmental systems and evolution*. Cambridge, UK: Cambridge University Press.

Richardson, D. C., & 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.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*(6), 867–891.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, *88*, 11–21.

Turvey, M. T. (2007). Action and perception at the level of synergies. *Human Movement Science*, *26*(4), 657–697.

Van Orden, G. C., & Holden, J. G. (2002). Intentional contents and self-control. *Ecological Psychology*, *14* (1–2), 87–109.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, *132*, 331–350.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2005). Human cognition and 1/f scaling. *Journal of Experimental Psychology: General*, *134*, 117–123.

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324, 0956797614531023.