

SELF-OTHER RELATIONS IN INTERPERSONAL SYNCHRONY

By
Elisa Wiedemann

Submitted to Central European University - Private University
Department of Cognitive Science

*In partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Cognitive Science.*

Primary Supervisor: Natalie Sebanz

Secondary Supervisor: Günther Knoblich

Vienna, Austria
2025

COPYRIGHT NOTICE

Copyright © Elisa Wiedemann, 2025. Self-Other Relations in Interpersonal Synchrony - This work is licensed under [Creative Commons Attribution-NonCommercial-NoDerivatives \(CC BY-NC-ND\) 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



For bibliographic and reference purposes this dissertation should be referred to as:
Wiedemann, E. 2025. Self-Other Relations in Interpersonal Synchrony. Doctoral dissertation,
Department of Cognitive Science, Central European University, Vienna.

¹ Icon by [Font Awesome](https://fontawesome.com/).

AUTHOR'S DECLARATION

I, the undersigned, **Elisa Wiedemann**, candidate for the PhD degree in Cognitive Science declare herewith that the present thesis titled “Self-Other Relations in Interpersonal Synchrony” is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright.

I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Parts of this dissertation will be submitted for publication with the following co-authors:

Chapters 2-4 with Natalie Sebanz

Chapter 5 with Louise Lund Petersen, Karen Victoria Willumsen Nustad, and Dora Kampis

Vienna, 30 June 2025

Elisa Wiedemann

ABSTRACT

This thesis examines self-other relations in interpersonal synchrony. Taking an experimental approach, we investigated in a series of experiments whether performing the same movements at the same time as another person leads to an increase in self-other overlap, thereby enhancing affiliation. We found that interpersonal synchrony affects perceived, but not motor-level self-other overlap, suggesting that it is likely a social recategorization of the self in relation to others that gives rise to the effects of interpersonal synchrony. In a further qualitative study, we addressed the experience elicited by interpersonal synchrony in discrete and continuous movement contexts, finding that the experience of interpersonal coordination can be described as following a generic diachronic structure made up of three phases: an initial phase of starting, a phase of (non-)adaptation, and a phase of stable coordination. We also found evidence for some structural variations, such as the addition of a phase of experimenting, as well as inter-individual variation, particularly with respect to (non-)adaptation and experimenting. Finally, an experimental study with 18-month-olds considered the phenomenon of interpersonal synchrony in development, examining its effects of self-other alignment and its links to toddlers' development of a self concept. This study showed that interpersonal (a-)synchrony highlights the (dis-)similarity between self and other but that toddlers' responses to it differ between measures and with their self-concept development. Overall, the findings presented in this thesis suggest that interpersonal synchrony acts as a cue to group membership by prompting a social recategorization of the self in relation to one's movement partner(s) and that the context in which interpersonal synchrony occurs affects the way people experience the interaction.

Key words: interpersonal synchrony, joint action, subjective experience, development

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Natalie Sebanz, as well as my co-supervisors, advisors, and mentors for their guidance and support in making this dissertation possible: Dora Kampis, Thomas Wolf, Günther Knoblich, Laura Schmitz, and Thom Scott-Philips.

I am grateful for the support from the Department of Cognitive Science at CEU, and the Social Mind Center and Somby lab in particular, on this journey. Réka, Fanni, Candasch, and Andi – I could not have done this without you.

I would also like to express my gratitude to the Childlab at the University of Copenhagen, and all its members, for hosting me and taking care of me during my stay. I am particularly appreciative of the hands-on support I received from Louise, Karen, Nathascha, Emily, Emese, and Helle in making our project possible.

I am appreciative of all the fruitful discussions I had with colleagues, collaborators, friends, and acquaintances that kept me motivated and helped me push through.

Finally, I would like to thank my friends and family for their patience and support.

TABLE OF CONTENTS

Chapter One: Introduction.....	1
Interpersonal Synchrony	1
Mechanisms.....	2
Close Perception-Action Links	2
Shared Goals	5
Social Group Membership	5
Experience.....	6
Development	6
The Present Thesis	7
Chapter Two: Self-Other Overlap in Interpersonal Synchrony.....	8
Pilot Study.....	11
Methods.....	12
Results	27
Discussion	38
Experiment 1	45
Methods.....	46
Results	60
Discussion	68
Supplementary Information.....	74
Pilot 1: Data Treatment and Results.....	74
Pilot 2: Data Treatment and Results.....	78
Pilot 3: Data Treatment and Results.....	81
Pilots 1-3: Supplementary Tables	84

Experiment 1: Methods and Results.....	87
Chapter Three: Does interpersonal synchrony increase self-other overlap?.....	99
Experiment 1	102
Methods.....	102
Results	116
Discussion	121
Experiment 2	123
Methods.....	123
Results	127
Discussion	133
General Discussion.....	135
Supplementary Information.....	138
Experiment 1	138
Automatic Imitation Task.....	138
Experiment 2	142
Chapter Four: The Subjective Experience of Interpersonal Coordination.....	144
Methods.....	147
Participants	147
Materials and Procedure.....	148
Data Treatment.....	151
Results	154
Coordination Data	154
Interview Data.....	157
Discussion	173
Supplementary Information.....	183
Results	183

Chapter Five: Does interpersonal synchrony affect self-other alignment in 18-month-olds?	186
Methods	190
Participants	190
Materials and Procedure	192
Data Treatment	200
Results	205
Mirror Mark Test.....	205
Rhythmical Movement Task	205
Sticker Task.....	208
Imitation Task	217
Pointing Task.....	221
Questionnaire Data	224
Discussion	226
Supplementary Information.....	237
Results	237
Chapter Six: General Discussion.....	244
Bibliography.....	254

LIST OF FIGURES AND TABLES

List of Figures

Figure 1: Procedures.....	14
Figure 2: Coordination Tasks.....	17
Figure 3: Automatic Imitation Task: Trial Structure	18
Figure 4: Automatic Imitation Task: Left- vs. Right-Hand Stimuli.....	20
Figure 5: Inclusion of Other in Self (IOS) Scale.....	23
Figure 6: Compatibility Effect by Stimulus Group (RTs).....	29
Figure 7: Interference Scores	31
Figure 8: IOS Ratings – Block 1	35
Figure 9: Procedure	48
Figure 10: Coordination Task	49
Figure 11: Automatic Imitation Task: Trial Structure and Left- vs. Right-Hand Stimuli	51
Figure 12: Automatic Imitation Task: Model	53
Figure 13: Procedure – Experiment 1.....	104
Figure 14: Coordination Task – Experiment 1.....	106
Figure 15: Ratings of Connectedness.....	108
Figure 16: Automatic Imitation Task – Experiment 1: Trial Structure.....	109
Figure 17: Inclusion of Other in Self Ratings	111
Figure 18: IOS Ratings – Experiment 1	120
Figure 19: Procedure – Experiment 2.....	124
Figure 20: IOS Ratings – Experiment 2	131
Figure 21: Procedure	148
Figure 22: Tapping Task: Signed Asynchronies	155

Figure 23: Sliding Task: Relative Phase	156
Figure 24: Generic Diachronic Structure of the Experience of Interpersonal Coordination .	158
Figure 25: Procedure	192
Figure 26: Imitation Task: Apparatus	196
Figure 27: Pointing Task: Toys	197
Figure 28: Pointing Task: Objects	198
Figure 29: Pointing Task: Room Layout	199
Figure 30: Synchronization Index by Condition	207
Figure 31: Sticker Choice by Condition	209
Figure 32: Sticker Choice by Condition and Mirror Mark Test Passing	210
Figure 33: Sticker Placement on Own or Companion’s Face by Condition	214
Figure 34: Sticker Placement by Condition and Mirror Mark Test Passing	215
Figure 35: Head Touch (20s) by Condition	218
Figure 36: Head Touch (20s) by Condition and Mirror Mark Test Passing	219
Figure 37: Pointing by Condition	223

List of Supplementary Figures

Supplementary Figure 1: Automatic Imitation Task: Stimuli – Practice Phase	88
Supplementary Figure 2: Automatic Imitation Task: Stimuli – Experimental Blocks	89
Supplementary Figure 3: Automatic Imitation Task: Stimuli – Experimental Blocks (Marked)	90
Supplementary Figure 4: Relation between Music and Dance Experience and Synchronous Tapping	94
Supplementary Figure 5: Relation between Music and Dance Experience and Asynchronous Tapping	96

Supplementary Figure 6: Relation between Asynchronous Tapping and Perceived Overlap with the Task Partner.....	97
Supplementary Figure 7: Automatic Imitation Task – Stimulus Images	139
Supplementary Figure 8: Automatic Imitation Task – Example Stimuli.....	141
Supplementary Figure 9: Heatmap Visualization of Correlation of Rhythmic Coordination Task Variables and Age	237
Supplementary Figure 10: Heatmap Visualization of Correlation of ECBQ Variables	243

List of Tables

Table 1: Interference Scores by Model and Condition.....	62
Table 2: Discrete Asynchronous Coordination	167
Table 3: Discrete Synchronous Coordination	168
Table 4: Continuous Asynchronous Coordination	169
Table 5: Continuous Synchronous Coordination	170
Table 6: Result Summary.....	225

List of Supplementary Tables

Supplementary Table 1: Difficulty Ratings of the Coordination Task	84
Supplementary Table 2: Enjoyment Ratings of the Coordination Task.....	85
Supplementary Table 3: Difficulty Ratings of the Automatic Imitation Task.....	86
Supplementary Table 4: Tapping Task: Mean ITI in each Trial.....	183
Supplementary Table 5: Tapping Task: Mean Asynchrony in each Trial	184
Supplementary Table 6: Sliding Task: Mean Velocity in each Trial	185
Supplementary Table 7: β Coefficients of Sticker Choice Predictors.....	238
Supplementary Table 8: β Coefficients of Sticker Placement Predictors	239
Supplementary Table 9: β Coefficients of Head Touch (20s) Predictors	241

CHAPTER ONE: INTRODUCTION

When playing team sports or making music, when rocking a child to sleep or simply walking down the street, people are constantly coordinating their movements with other people around them, either intentionally or unintentionally. This may mean accidentally falling into step or purposefully hitting a drum at the same time. The present thesis set out to examine the latter, aiming to conduct a closer investigation of self-other relations in interpersonal synchrony from different perspectives.

Interpersonal Synchrony

Interpersonal synchrony has been used as a term to refer to a variety of different behaviors, ranging from physiological to neural to behavioral alignment. In the present thesis, we take the term to refer to contexts, in which people perform the same movement at the same time as another person or people around them. Interpersonal synchrony may thus take on many forms. It can occur in a variety of contexts, such as music and dance, sports and physical activity, rituals, parent-child interactions, or marches and protests, and it encompasses a wide range of different kinds of interactions, ranging from simple motor tasks to complex multimodal interactions like music or dance. Research on interpersonal synchrony has been conducted in both naturalistic settings and in laboratory contexts; a common theme linking occurrences of interpersonal synchrony in these different environments is that they have been shown to elicit a variety of prosocial effects.

Indeed, tapping in synchrony with others has been shown to increase affiliation between co-actors (Hove & Risen, 2009). Studies have found synchronized rowing to increase subjective pain thresholds and suggest that this is indicative of social bonding (Cohen et al., 2009; Sullivan & Rickers, 2013). Dancing in synchrony has also been shown to increase pain

thresholds and social closeness (Tarr et al., 2016). Both real and imagined synchronous walking has been shown to reduce prejudice (Atherton et al., 2019). Even 14-month-old infants have been shown to help an experimenter more after bouncing in synchrony with her (Cirelli et al., 2014), and four-year-old children showed better cooperation and more intentional communication with a peer after synchronous swinging (Rabinowitch & Meltzoff, 2017).

Mechanisms

Despite widespread documentation of the existence of these effects, the literature on the underlying mechanisms remains quite scattered. A number of different processes have been proposed and a series of reviews have been conducted aiming to organize the literature (e.g., Rennung & Göritz, 2016; Mogan et al., 2017; Cirelli, 2018; Cross et al., 2019; Michael et al., 2020; Ayache et al., 2021; Basile et al., 2022; Pardo-Olmos et al., 2025), but a clear consensus has not yet been reached. In what follows, the emphasis will be on the relevance of close perception-action links in giving rise to the effects of interpersonal synchrony, as well as the role of shared goals and social group membership.

Close Perception-Action Links

One line of investigation has drawn on the idea that perception and action are closely linked. The idea of close perception-action links is not unique to the domain of interpersonal synchrony and has its roots in 19th century psychology. Building on the idea of the ideomotor principle, which states that thinking about a movement activates motor processes relevant for performing the movement (James, 1890), different models have been put forward (Shin et al., 2010). For instance, the theory of common coding (Prinz, 1990, 1997) and the theory of event coding (Hommel et al., 2001) have suggested that perceived and planned actions share common representational codes. Relatedly, the discovery of the mirror neuron system has provided

evidence for these ideas (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010; see also Brass & Muhle-Karbe, 2014).

Beyond the domain of ideomotor theory, it has been suggested that people rely on the same internal models for both the prediction of their own and others' actions (Wolpert et al., 2003). Similarly, and closely related to the discovery of the mirror neuron system, others have put forward an account that integrates this idea into a predictive coding framework (Kilner et al., 2007). Relatedly, Sebanz et al. (2005) found that when performing joint actions together, co-actors represent not only their own but also others' contributions to the joint action. The authors interpret this as evidence for close links between perception and action, and suggest that people must represent not only their own but also others' actions for successful coordination (Sebanz et al., 2006; Vesper et al., 2010).

Applied to the domain of interpersonal synchrony, Hove (2008) proposed a model in line with ideomotor conceptions of direct perception-action links and drawing on evidence of the existence of the mirror neuron system. This model suggests that increased affiliation between co-actors following interpersonal synchrony is the result of an increase in motor-level self-other overlap that occurs during interpersonal synchrony. In particular, this proposal suggests that two mechanisms give rise to this increase in self-other overlap, and thereby enhance affiliation. First, the matching of the sensory consequences of another person's action with the sensory consequences of one's own action may reduce a person's ability to distinguish between their own and the other's actions. Second, the priming of one's own motor system when observing another person perform an action, paired with a lack of inhibition of this motor response, may further reduce a person's ability to distinguish between their own and the other's actions.

This model is presented as an extension of the shared circuits model (Hurley, 2008), which proposes a layered architecture of cognition that operates on unsegregated information about perception and action at layer 1, then includes mechanisms for distinguishing between perception and action in layer 2 and between self and other in layer 4. Hove (2008) put forward the argument that by incorporating a temporal dimension, this model could be applied to the domain of interpersonal synchrony. Moving in synchrony with others would thus reduce one's ability to distinguish between perception and action in layer 2 and between self and other in layer 4, thereby leading to an increase in empathy and affiliation between synchronous movement partners. Indeed, Hove & Risen (2009) found evidence for this proposal, with synchronous tapping leading to greater affiliation with the co-actor than asynchronous tapping. Other studies have also found interpersonal synchrony to increase affiliation (e.g., Crossey et al., 2021; Basile et al., 2022). However, others did not replicate this effect, while finding evidence for an increase in other measures such as perceived self-other overlap or self-esteem (Lumsden et al., 2014). This suggests that while there is some support for the model put forward by Hove (2008), more work is needed to establish its robustness.

Another quite different approach to the topic builds on the idea that cognition is not primarily based on mental representations but that it is instead enacted via embodied interactions in the world (Varela et al., 1991). Rather than relying on internal models, Gallagher (2006) has proposed that people rely on direct social perception to understand others' actions. Moreover, De Jaegher & Di Paolo (2007) have proposed that in the context of interpersonal coordination, the social interaction itself takes on a form of autonomy through dynamic coupling between autonomous individuals, influencing the agents involved for as long as the interaction persists. Similarly, Fuchs & De Jaegher (2009) have argued that participants in an interaction physically and affectively overlap, co-regulating each other via their movements. Though different in emphasis, what these proposals have in common is the idea that (social)

cognition may not fundamentally rely on the centralized processing of mental representations to operate but may instead use features of its environment as a scaffolding for – or even as an integral part of – its architecture.

Shared Goals

Others have proposed that the presence of shared goals play a key role in producing the effects of interpersonal synchrony (Reddish et al., 2013; Cross et al., 2019). Research in the context of joint action has found that the presence of a shared goal results in more fluent and efficient coordination (Vesper et al., 2010), that it prompts task co-representation (Sebanz et al., 2005), and that it improves error monitoring and adaptive timing in the context of musical joint action (Keller et al., 2014). However, phenomena such as emergent, or incidental coordination highlight the fact that shared goals are not a necessary condition for successful coordination to occur (Renfrew et al., 2008). Yet studies have found the presence of shared goals to boost effects of interpersonal synchrony. For instance, Reddish et al. (2013) found interpersonal synchrony with a shared goal to lead to increased cooperation compared to incidental synchronization. Similarly, Cross et al. (2021) found interpersonal synchrony with a shared goal to elicit greater motor-level self-other overlap than incidental synchronization.

Social Group Membership

Another approach to explaining the effects of interpersonal synchrony has made reference to a process of social recategorization in terms of group membership (Cross et al., 2019). Social group membership has been shown to lead to ingroup favoritism, with people preferring to cooperate and coordinate with people who are part of their own social group (Tajfel & Turner, 1979; Hogg, 2016). In the context of interpersonal synchrony, it has also been argued that interpersonal synchrony may act as a cue to group membership, similar to the way music and dance function as coalition signals (Hagen & Bryant, 2003). This idea has found

some empirical support (Cross et al., 2020; Tarr et al., 2015; Cirelli et al., 2016), but other studies failed to find evidence for these claims (Reddish et al., 2014, 2016), suggesting that more research is needed.

Experience

Literature on interpersonal synchrony has often thematized ritualistic synchrony (Gelfand et al., 2020), emphasizing its affective and physiological effects (e.g., Fischer et al., 2014), and the types of experiences it elicits (e.g., Ehrenreich, 2006). While experimental and lab studies may reference this as the starting point of their investigation (e.g., Reddish et al., 2013; Biswas & Brass, 2024), they typically focus on other aspects of interpersonal synchrony in their studies or fail to find evidence for its effect on affective responses (see Mogan et al., 2017). As interpersonal synchrony has been posited to elicit a variety of different experiences, including not only affective responses, but also social closeness and connectedness (see Rennung & Göritz, 2016), investigating the experiences interpersonal synchrony elicits in a lab context is highly relevant. It allows researchers to understand whether the phenomenon lab studies investigate is comparable to the phenomenon documented in ethnographic reports.

Development

While interpersonal synchrony in neurotypical adults has received much attention, much less is known about its developmental origins. Evidence for prosocial effects of interpersonal synchrony has been documented from around 14 months of age (Cirelli et al., 2014), though it already begins to influence infants' social behavior before this (Tunçgenç et al., 2015). These effects are thought to be the result of interpersonal synchrony acting as a cue to affiliation (Fawcett & Tunçgenç, 2017) or social group membership (Cirelli, 2018). However, there is a gap in the literature on interpersonal synchrony between the ages of 15 months to 2.5 or 3 years, with only a handful of studies conducted during this period (e.g.,

Kirschner & Tomasello, 2009; Bazhydai et al., 2022). This is worth noting, as several important developments occur during this period, including that of a self-concept (Rochat, 2018). As one's categorization of oneself in relation to others is closely linked to one's own subjective self-awareness, an investigation of the relationship between interpersonal synchrony and the development of a self-concept appears crucial to gain a better understanding of the processes underlying the prosocial effects of interpersonal synchrony in development.

The Present Thesis

The present thesis aimed to investigate interpersonal synchrony from different perspectives. First, the effects of interpersonal synchrony on motor-level self-other overlap were investigated (Chapters 2 and 3). In doing so, we draw on the outlined approaches to interpersonal synchrony, testing the proposal put forward by Hove (2008), while also including a shared goal and acknowledging the potential alternative mechanism of social recategorization in terms of group membership. Next, we aimed to tackle the role of experience in interpersonal synchrony by investigating the subjective experiences our tasks elicit (Chapter 4). Finally, we turned to development and tested for an effect of interpersonal synchrony on self-other alignment, while examining its relationship with self-concept development (Chapter 5). Overall, this thesis aimed to contribute to a better understanding of self-other relations in interpersonal synchrony by investigating the interpersonal synchrony in a lab setting, using qualitative and quantitative methods, and targeting different populations.

CHAPTER TWO: SELF-OTHER OVERLAP IN INTERPERSONAL SYNCHRONY

Moving in time with others is a widespread phenomenon. It often happens unintentionally – when walking together, we fall into step; when applauding at a concert, our clapping becomes indistinguishable from the clapping of other audience members. But it can also be brought about intentionally – for instance, when learning a new dance routine or taking part in an exercise class, we do our best to copy the instructor’s movements while also trying to move in time with the rest of the group. These different forms of interpersonal synchrony have been shown to lead to a variety of prosocial effects, including increased affiliation (Hove & Risen, 2009), more cooperation (Wiltermuth & Heath, 2009), and better coordination (Valdesolo & DeSteno, 2011) between moving partners.

There are a number of mechanisms that may underlie these effects (see Mogan et al., 2017 for a review). Some have proposed that interpersonal synchrony acts as a cue to group membership and leads us to reclassify ourselves as members of the same social group as our synchronous counterparts (e.g., Cross et al., 2019; Good et al., 2017). This in turn is thought to lead us to exhibit in-group favoritism and act more pro-socially towards others, as we identify them as members of our own social group. Others have argued that performing the same movements at the same time as another person leads to an increase in self-other overlap between synchronous movers, or a reduction in their ability to distinguish between themselves and their synchronous counterpart (Hove, 2008). This increase in self-other overlap is then thought to lead to stronger affiliation between co-actors and more prosocial behavior (Hove & Risen, 2009). While these mechanisms need not be mutually exclusive, only few studies (e.g., Cross et al., 2017, 2021; Launay et al., 2013) have directly put them to the test. The present research aims to test whether moving in synchrony with another person leads to an increase in

self-other overlap with them and to more prosocial behavior towards them, compared to moving asynchronously, and to thereby contribute to a better understanding of how interpersonal synchrony affects self-other relations.

Hove (2008) proposed that there are two routes by which moving in synchrony with another person increases self-other overlap, and thereby enhances affiliation. First, observing another person perform an action whose sensory consequences match the predicted sensory consequences of one's own actions may reduce one's ability to distinguish between one's own and the other's actions. Second, observing another person perform an action primes one's own motor system to mirror this action which leads one to perform the same action at the same time as the other person – if one's own motor output is not actively inhibited. This mirroring tendency may also lead to a decrease in one's ability to distinguish between one's own and the other's actions. The resulting increase in motor-level self-other overlap is then thought to lead to increased affiliation between co-actors due to shared representational codes.

The idea that observing another person's actions primes our own motor system in response is supported by research on the mirror neuron system (Rizzolatti & Sinigaglia, 2010) and is consistent with the framework of common coding for perception and action (Prinz, 1990). Studies have shown that observing simple finger movements automatically primes an imitative response (see Cracco et al., 2018 for a review), and these automatic imitative tendencies have been proposed to serve as an index of mirror neuron system activation (Heyes, 2011). Moreover, this imitative response may be reduced if the focus is placed on oneself (Spengler et al., 2010; but see Khemka et al., 2021), and it has also been linked to activation in the right temporo-parietal junction (rTPJ), a brain area associated with switching between representations of self and other (Sowden & Catmur, 2015).

Evidence that people are affected by (observing) the actions of others exists not only in the domain of action perception, but is visible in joint action contexts as well. When people act together in order to reach a shared goal, they represent not only their own contributions towards reaching the shared goal but also their co-actors' (Sebanz et al., 2005). This co-representation of others' contributions may interfere with people's individual performance but might serve to facilitate reaching the shared goal. In the context of interpersonal synchrony, the presence of a shared goal has been shown to lead to increased cooperation compared to incidental synchronization without a shared goal (Reddish et al., 2013), and has also been linked to stronger automatic imitative tendencies (Cross et al., 2021).

Most studies aiming to test whether interpersonal synchrony leads to an increase in self-other overlap have relied on questionnaire items, such as the Inclusion of Other in Self (IOS) scale (Aron et al., 1992), a 7-point rating scale in which participants are asked to select which of seven images depicting two circles with different degrees of overlap best describes their relationship with the other person, or ratings of participants' liking of their task partner as their primary measure of self-other overlap (e.g., Hove & Risen, 2009; Reddish et al., 2016; Wiltermuth, 2012a; 2012b). However, Cross et al. (2021) measured self-other overlap in a more direct way. In their study, participants first performed a coordination task, in which they moved either synchronously or asynchronously with a task partner. Then, participants performed an automatic imitation task, in which they responded to videos of their task partner's and the experimenter's hands. This allowed the authors to quantify participants' imitative tendencies for their task partner's and the experimenter's finger movements, respectively, and to directly compare motor system activation in both cases. Their results showed stronger imitative tendencies for the task partner's hand than the experimenter's hand following intentional synchronization compared to asynchronous movement. This suggests that interpersonal

synchrony, when achieved intentionally, may indeed induce stronger mirroring tendencies and a decrease in self-other distinction.

Previous research thus provides initial support for the proposal that interpersonal synchrony may lead to an increase in self-other overlap (Hove, 2008), yet robust empirical evidence is lacking; the findings by Cross et al. (2021) so far have not been replicated. Therefore, the present study aimed to further investigate how intentional synchronization influences self-other overlap, both on a motor-level and on the level of subjective self-perception. We first report three pilot studies, using a within-subjects design, and then the resulting experiment, using a between-subjects design. Each study includes a coordination task, in which people either intentionally synchronized their movements with a partner, or in which they intentionally moved at a different pace than their partner (similar to the protocol of Cross et al., 2021), and two measures of self-other overlap. The final experiment also included a prosocial helping task to allow us to compare our results to previous findings. We predicted that intentional synchronization would lead people to show greater self-other overlap with their partner, compared to simultaneous movement at different paces, and to display more prosocial helping behavior.

Pilot Study

Across three pilot experiments, we examined the effects of intentional synchronization on self-other overlap. In an initial attempt at a conceptual replication of Cross et al. (2021), participants performed sliding movements on a large touchscreen either synchronously or asynchronously with their task partner before completing our measures of self-other overlap. As in Cross et al. (2021), participants wore colored gloves to make hand identity more salient in this pilot study. Unlike Cross et al. (2021), however, in pilot experiment 1, participants stood next to each other rather than facing one another during the coordination task. As we were

unable to replicate their findings in pilot experiment 1, we modified the spatial arrangement of our coordination task in pilot experiment 2, making it more similar to the one used in Cross et al. (2021) by having participants stand facing one another. We did not replicate their findings in this pilot experiment either. In pilot experiment 3, we made more substantial changes to our protocol, moving from a continuous to a discrete coordination task, in which participants tapped synchronously or asynchronously with a confederate. We also removed the gloves and instead used images of the confederate's and the experimenter's bare hands in the automatic imitation task. All three pilot experiments followed a within-subjects design. In what follows, we report methods and results collectively for the three pilot experiments to serve as a basis for the logic we followed in our implementation of experiment 1.

Methods

Participants

40 participants (26F, 12M, 20-55 years old) were recruited from the Sona System database (<https://www.sona-systems.com/>, last accessed 2023-01-25) of Central European University (CEU) PU to take part in the pilot studies reported here. An additional 4 participants were recruited but had to be excluded due to technical error. In pilot studies 1 ($n = 16$) and 2 ($n = 16$), participants were recruited to take part in pairs of two. In pilot study 3 ($n = 8$), participants completed the experiment with a confederate who acted as their task partner. All participants were right-handed, had normal or corrected-to-normal vision, no motor impairments, no history of neurological impairments or diagnoses, and spoke (at least) basic English. The pilot studies took place in the laboratories of the Social Mind Center of CEU PU in Vienna. All procedures used in these studies were approved by the Psychological Research Ethics Board (PREBO) at CEU PU Vienna.

Materials and Procedure

Participants performed a coordination task, an automatic imitation task (based on the paradigm in Brass et al., 2000), and Inclusion of Other in Self (IOS) ratings (Aron et al., 1992). The overall procedure differed somewhat between pilots (see Figure 1). In pilot 1, participants performed one block of the coordination task, one block of the automatic imitation task, IOS ratings, before repeating this same procedure again and completing a final questionnaire at the end of the experiment. Pilot 2 was identical to pilot 1 in structure but included a baseline IOS rating before the first block of the coordination task to allow us to quantify how participants' perceived self-other overlap with their task partner and with the experimenter changed due to the experimental manipulation. Pilot 3 included multiple shorter blocks of the coordination task, each followed by either an automatic imitation task (blocks 1, 2, 4, and 5) or IOS ratings (blocks 3 and 6). This way, each dependent measure was directly preceded by a block of the coordination task. The final questionnaire was administered at the end of the experiment.

Make-Believe Generation of Stimuli

At the start of the experimental session, participants provided informed consent to the procedures. They were then told that experimenter 2 would take pictures of their hands and that these pictures would later be used as stimuli for one of the tasks. In pilot studies 1 and 2, both participants and experimenter 1 were assigned different colored gloves (white, green, or black) to wear for the duration of the experiment so that their hands could easily be distinguished. Experimenter 2 then took three pictures of each participant's as well as experimenter 1's right hand: one in resting position and two performing simple finger movements (raising index finger and raising middle finger). In pilot study 3, a confederate acted as one of the participants, and we did not use colored gloves because we thought that seeing natural hands might increase imitative tendencies. Experimenter 2 took pictures of the participant's, the confederate's and experimenter 1's right hand.

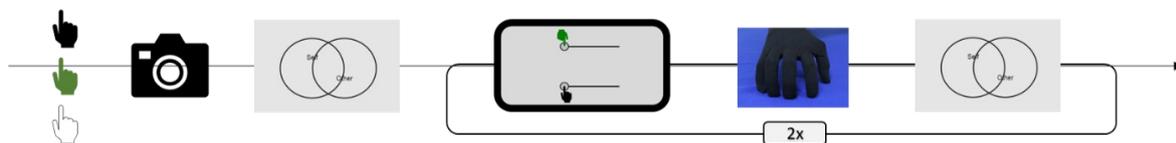
Figure 1

Procedures

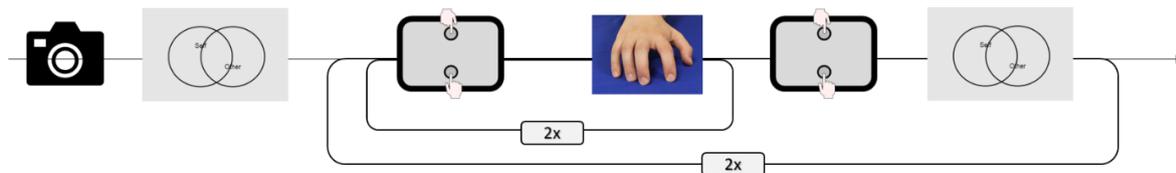
A



B



C



Note. Procedure in pilot study 1 (A), pilot study 2 (B), and pilot study 3 (C). In pilot studies 1 and 2, each participant was assigned a colored glove (black, green, or white) to wear for the duration of the experiment (A and B); experimenter 1 wore a glove of the remaining color. Then experimenter 2 took pictures of both participants' and of experimenter 1's right hand. After this, in pilot study 2, participants performed a baseline Inclusion of Other in Self (IOS) rating (B). Then, in pilot studies 1 and 2, participants performed the coordination task together on a large touchscreen, followed by an automatic imitation task and IOS ratings (A and B). This procedure was repeated twice, once in each condition (synchronous vs. asynchronous coordination). In pilot study 3, experimenter 1 immediately took pictures of the participant's, the confederate's, and experimenter 1's hand (C). Participants then completed baseline IOS

ratings and performed the coordination task together with the confederate, followed by the automatic imitation task, twice (once with images of the experimenter's hand as stimuli in the automatic imitation task, once with images of the confederate's hand as stimuli in the automatic imitation task). They then performed another block of the coordination task, followed by IOS ratings, before the same procedure was repeated in the other condition (synchronous or asynchronous coordination).

Coordination Task

The coordination task followed the same general structure in all three pilots. Participants heard a metronome through a set of headphones (either 59bpm or 83bpm) and were instructed to move in time with it. In pilot studies 1 and 2, participants were instructed to perform continuous movements along a 30cm track presented on a large touchscreen, reaching one end of the track each time they heard a tone. In pilot study 3, the movements were discrete, as participants were instructed to tap on a virtual button each time they heard a tone. We switched from continuous to discrete movements in the coordination task, as we reasoned that tapping movements would be more similar to the movements participants were asked to perform during the automatic imitation task and should thus lead to a stronger modulation of participants' imitative tendencies.

The metronome was only present for the first 20 seconds of each trial, then it stopped. Participants were instructed to continue moving at the pace of the metronome once it stopped until they received a signal indicating that the trial was over. In synchronous coordination blocks, both participants heard identical metronomes (both heard 59bpm or both heard 83bpm). They were additionally instructed to try to move in time with each other once the metronome stopped. In asynchronous coordination blocks, participants heard different metronomes (one heard 59bpm, the other heard 83bpm). In these blocks, they were additionally instructed to try

to stick to their own pace. In both conditions, the emphasis was placed on the second part of the instructions to ensure that participants understood the difference between conditions.

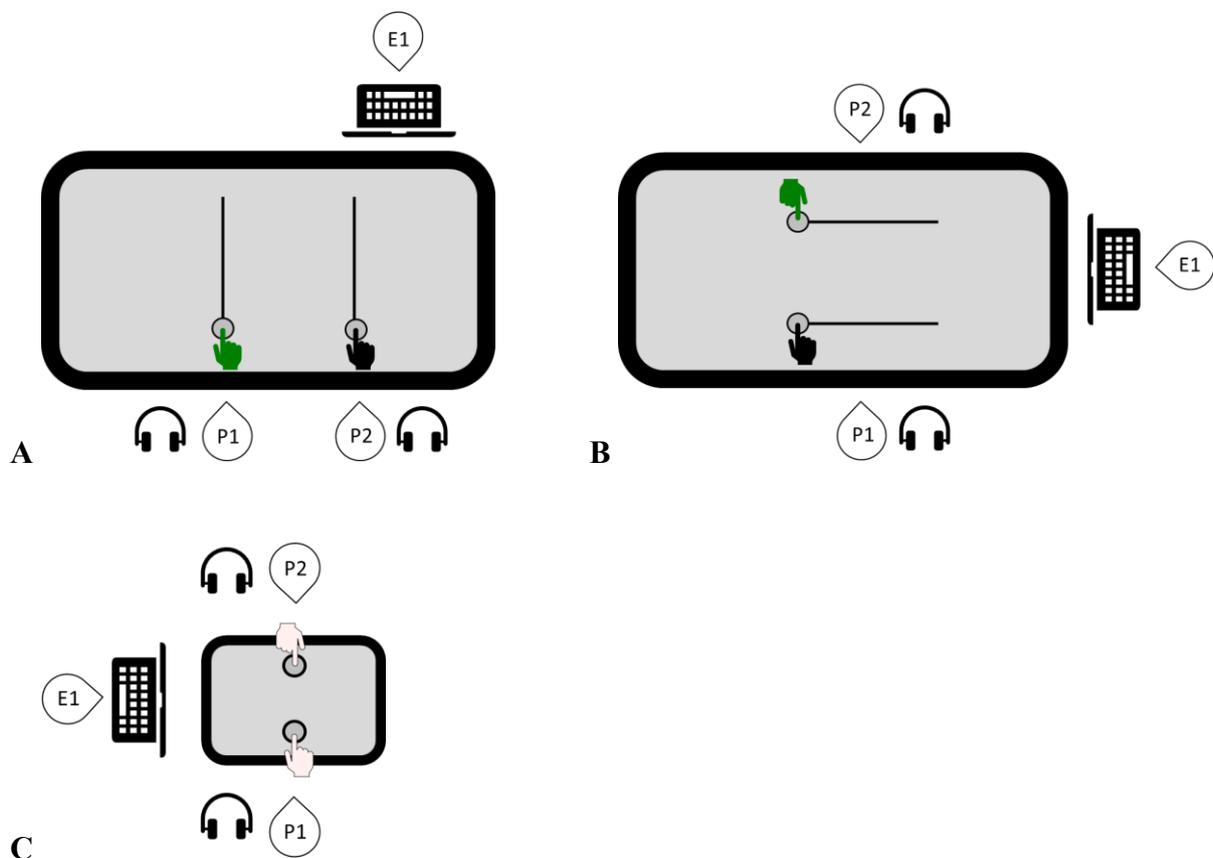
Participants completed four practice trials (first two synchronous trials, then two asynchronous trials) lasting 100 seconds each in pilot studies 1 and 2, and 40 seconds each in pilot study 3, before starting the experimental blocks. In the first two pilot studies, each block consisted of eight trials à 100 seconds, and participants completed one synchronous and one asynchronous coordination block. The order of these blocks was counterbalanced. In the third pilot study, each block contained only two trials à 60 seconds, and participants completed six blocks of the coordination task (i.e., three synchronous coordination blocks, and three asynchronous coordination blocks). The order of synchronous vs. asynchronous coordination blocks was also counterbalanced in this pilot study, but participants always completed three blocks of one type of coordination before completing three blocks of the other type of coordination.

In pilot studies 1 and 2, the coordination task was performed on a 46-inch iiyama ProLite touchscreen (Figure 2, top row). Participants stood either next to each other (pilot 1) or opposite one another (pilot 2) facing the touchscreen. They moved a virtual knob back and forth along a 30cm long track. We changed the layout of the task in pilot study 2 (see Figure 2) to introduce continuity in participants' perspective on their task partner's hand between the coordination task and the automatic imitation task. In pilot study 3, participants performed the coordination task on a 12.9-inch iPad Pro together with a confederate. This allowed us to use real images of the task partner's and the experimenter's hand and to keep these constant across participants, as all participants completed the task with the same person. The confederate was instructed to act as though she were a participant in the study. The participant and the confederate sat facing each other on either side of a large table. As in pilot study 2, this layout

was chosen for the sake of maintaining continuity of perspective on the task partner's hand between the coordination task and the automatic imitation task. The coordination task in this pilot study consisted in tapping on a virtual button on a tablet that was placed on the table between the participant and the confederate (see Figure 2C).

Figure 2

Coordination Tasks



Note. Coordination task in the first (A), second (B), and third pilot study (C). In the first two pilot studies (A and B), two participants (P1 and P2) performed the coordination task together on a large touchscreen wearing an assigned colored glove (green, white, or black). Participants in the first pilot study stood side by side (A); participants in the second pilot study stood facing one another on opposite sides of the touchscreen (B). They moved a virtual knob along a 30cm track vertically in pilot 1 (A), or horizontally in pilot 2 (B). In the third pilot study (C), one

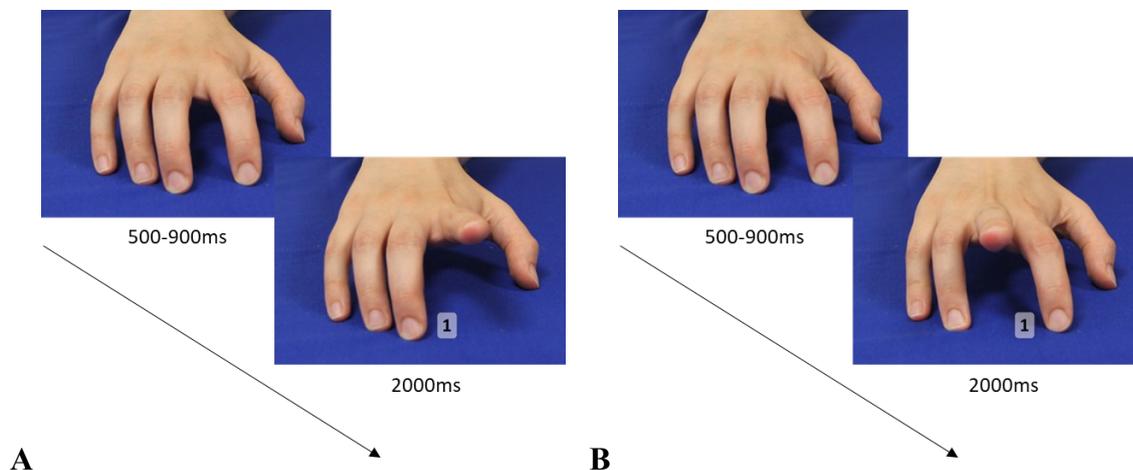
participant (P2) performed the coordination task on a tablet together with a confederate (P1). They sat on opposite sides of the tablet, and their task was to tap on a virtual knob. Experimenter 1 (E1) sat behind a laptop (in all pilot studies) and controlled the start of each trial. Participants were instructed to move either synchronously or asynchronously with their partner. For the first 20 seconds of each trial, they were paced by a metronome played through a set of headphones, then they continued without the metronome while trying to keep the instructed pace.

Automatic Imitation Task

We used a two-alternative forced-choice automatic imitation task, with key release serving as our dependent measure. Participants were instructed to press two keys marked with yellow stickers on their computer keyboard (V and B) with the index and middle finger of their right hand while viewing the images, and to release either their index or their middle finger in response to numbers displayed on the screen (i.e., to lift their index finger in response to a 1 and their middle finger in response to a 2, as in Brass et al., 2000). The numbers on the screen were always presented together with pictures of a hand (see Figure 3).

Figure 3

Automatic Imitation Task: Trial Structure



Note. Example of an effector-compatible trial (A) and an effector-incompatible trial (B). In the automatic imitation task, participants are shown an image of a hand in resting position (top left image in both A and B) for 500 to 900ms, then they are shown an image of the same hand with either a raised index finger (bottom right image in A) or a raised middle finger (bottom right image in B). The second image is displayed together with a number cue (either a 1 or a 2) for a maximum of 2 seconds, or until a response is made. Participants are instructed to lift their index finger in response to a 1 and to lift their middle finger in response to a 2. The depicted finger movement is either compatible (A) or incompatible (B) with the cued response. The images shown in this figure were used as stimuli in pilot study 3 (experimenter 1's hand).

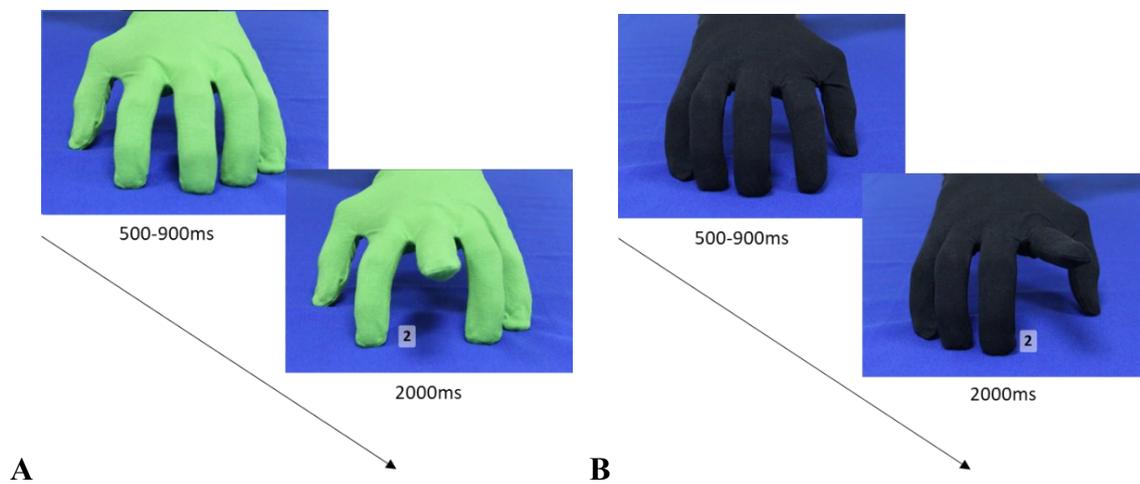
In this task, there was one within-subject manipulation (Effector Compatibility) and there were two between-subject factors we manipulated (Stimulus and Model). We manipulated Effector Compatibility by varying the finger movement participants were shown together with the number cue. In half of the trials, participants saw a movement that was compatible with the prompted response (e.g., a raised index finger presented together with a 1, as in Figure 3A). In the other half, they were shown a movement that was incompatible with the one they were prompted to perform (e.g., a raised middle finger together with a 1, as in Figure 3B). These compatible vs. incompatible trial types were presented in random order in each block.

The between-subjects manipulation of Stimulus consisted in showing half of the participants in our first two pilot studies left hand images in the automatic imitation task, while the other half saw right hand images. Participants always responded with their right hand. In left hand images, effector-compatible movements were also spatially compatible (e.g., a depicted index finger movement was on participants' left, as was their own index finger), so the two types of compatibility were confounded (see Figure 4A). In right hand images, spatial

compatibility and effector compatibility were opposed (e.g., a depicted index finger movement was further to the right, while their own index finger was further left; see Figure 4B). We included this manipulation in pilot studies 1 and 2 to control for effects of spatial compatibility, following the recommendations in Cracco et al. (2018). Pilot study 3 contained only right-hand images.

Figure 4

Automatic Imitation Task: Left- vs. Right-Hand Stimuli



Note. Examples of left-hand stimuli (A) and right-hand stimuli (B) used in pilot studies 1 and 2. Participants were assigned to one of two Stimulus groups and saw either left-hand stimuli (A) or right-hand stimuli (B) in both blocks of the automatic imitation task. Left-hand images were mirrored versions of the right-hand stimuli.

In addition to the factors Effector Compatibility and Stimulus, we also manipulated the factor Model: in half of the experimental trials, participants were told they would be responding to images of their task partner's hand; in the other half they were told they would be shown experimenter 1's hand. In pilot studies 1 and 2, participants were not shown pictures of their task partner's or experimenter 1's hand but instead saw images of a hand wearing the same

colored glove that their task partner had been assigned (e.g., black) and that experimenter 1 had been assigned (e.g., green; see Figure 4). In pilot study 3, participants were shown images of the confederate's hand and of experimenter 1's hand (see Figure 3). All stimuli were prepared ahead of time, and the same set of stimuli was used for all experimental sessions.

Participants first completed two rounds of practice trials, with each round consisting of eight trials. In the first round, they were shown an image of a hand in resting position for 500 to 900ms, then an image of the same hand with the number cue for a maximum of 2 seconds. The second image disappeared once one of the two response keys was released. This round of practice trials was repeated as often as necessary until participants had learned the correct cue-response mapping and responded correctly on at least 75% of trials. They then moved on to the second round of practice trials, in which they were shown the same image of a hand in resting position for 500 to 900ms, then an image of the hand with the number cue, this time with a raised index or middle finger, for a maximum of 2 seconds. Once they succeeded in giving a correct response on at least 75% of the practice trials in this round, they proceeded to complete the first experimental block of the automatic imitation task in pilot studies 1 and 2, or completed the first block of the coordination task in pilot study 3. The stimuli used in these trials were adapted from Brass et al. (2000).

In experimental blocks, participants were shown 24 trials depicting their task partner's hand and/or 24 trials depicting the experimenter's hand. These trial types were randomly intermixed in each block in pilot studies 1 and 2 – participants saw images of hands wearing two different colored gloves in random order – resulting in 48 trials per block. The trial types were blocked in pilot study 3 – each block consisted of 24 trials depicting either the confederate's or the experimenter's hand. This way participants did not need to switch back and forth between representations of the different hand identities within each block, allowing

us to measure how our manipulation affected their response to each model separately. Each trial followed the same structure as the second round of practice trials: participants were shown an image of a hand in resting position for 500 to 900ms, then an image of the same hand with a raised index finger or middle finger, together with the number cue, for a maximum of 2 seconds, or until participants responded (see Figure 3).

At the end of the final block, participants were shown two images of hands and were asked to identify which one was their task partner's or the experimenter's hand. The images they were shown corresponded to the images they had responded to in the task. This question served as a manipulation check to make sure that participants were able to detect the hand's identity. Participants were also asked to give a rating of how confident they were that their response was correct. The responses to this question allowed us to gauge how difficult it was for them to differentiate the confederate's and the experimenter's hand, particularly in pilot study 3. In pilot studies 1 and 2, participants were additionally asked to rate their confidence that the images they had responded to were actually the images that had been taken at the start of the session. This gave us an indication of whether participants believed the information they had been given concerning the hand stimuli in the automatic imitation task.

Inclusion of Other in Self (IOS) Ratings

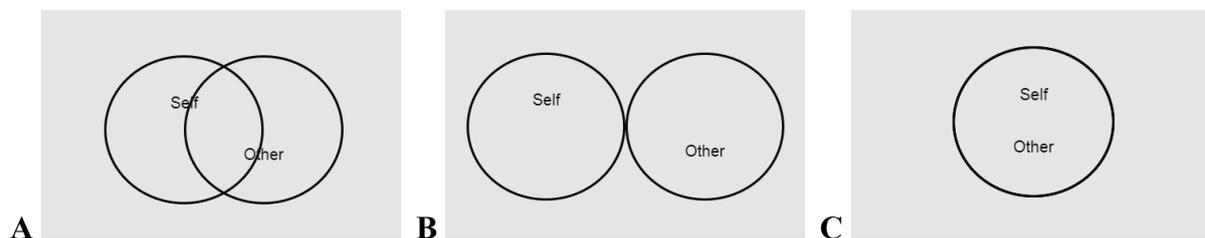
We used a continuous version of the IOS scale designed by Aron et al. (1992). Participants were instructed to arrange two circles, labeled 'Self' and 'Other', in the way that best described their relationship with their task partner, and with experimenter 1 (see Figure 5). The circles were arranged in such a way that they had an overlap of 50% at the start of each rating. We chose to start at the midway point of the scale each time so as not to bias participants in either direction. Participants performed one rating of overlap with their task partner and one

rating of overlap with experimenter 1 in each block. The ratings were performed in random order.

Participants performed the IOS ratings by moving the circle labeled ‘Self’ left or right until they reached their desired arrangement. When participants moved the circle labeled ‘Self’ closer to the other circle (creating more overlap), the circle labeled ‘Other’ also moved closer to the center by the same number of pixels. When participants moved the circle labeled ‘Self’ further away from the other circle, the circle labeled ‘Other’ automatically moved away from the center, too. This way the circles stayed centered on the screen and only the percentage of overlap changed. The final percentage of overlap was recorded and entered into the analyses.

Figure 5

Inclusion of Other in Self (IOS) Scale



Note. Participants performed the IOS ratings by arranging two circles (one labeled ‘Self’, one labeled ‘Other’) in the way that best described their relationship to the task partner, and experimenter 1, respectively. At the start of each trial, the circles were presented with 50% overlap (A). Participants could then move the circles apart to a minimum of 0% overlap (B), and together to a maximum of 100% overlap (C). Participants moved the circle labeled ‘Self’ by dragging it using a mouse in pilots 1 and 2 or using the arrow keys on their computer keyboard in pilot 3.

Final Questionnaire

At the end of the experiment, participants filled out a final questionnaire. They were asked to rate their liking of their task partner and experimenter, their enjoyment of the coordination task in general as well as in the synchronous and asynchronous condition separately. They were also asked about the difficulty of the coordination task – in general, in synchronous blocks, and in asynchronous coordination blocks – and about the difficulty of the automatic imitation task – in general, with regard to ignoring the depicted finger movement, and with regard to ignoring the depicted finger movement when they responded to the experimenter’s and their task partner’s hand, respectively. All ratings were performed on a scale from 1 to 7. The questionnaire additionally included questions about participants’ experience with music and dance, probed their hypotheses as to what the experiment was about, and gave them the opportunity to provide comments and feedback. Participants’ responses to these items are not reported in the present chapter.

Data Treatment

Coordination Data

The coordination data were analyzed separately for each pilot study and compared between synchronous and asynchronous coordination blocks. The analysis procedures and results for each pilot study can be found in the supplementary information.

Automatic Imitation Data

Participants’ responses in the automatic imitation task were analyzed in terms of their reaction times (RTs) and error rates (ERs). RTs are thought to measure of covert imitative tendencies, with ERs providing an index of overt imitation (Cracco & Brass, 2019). As our hypotheses mainly target covert imitative tendencies, predicting a modulation thereof, we will mainly focus on RTs in our analysis.

Participants who failed to correctly identify their task partner's or the experimenter's hand after the final block of the automatic imitation task were excluded from the analyses. Trials in which participants responded before the number cue appeared were coded as invalid and excluded from the analysis. Trials that timed out because participants did not make a response were also excluded from the analysis. The remaining trials were coded as correct if participants responded by lifting the cued finger or as incorrect if they lifted the other finger. Participants who responded incorrectly on more than 10% of trials were excluded from the analysis. Correct and incorrect trials were analyzed separately: only correct trials were included in the RT analysis; incorrect trials were considered in the ER analysis. Before performing the RT analysis, all RTs that were more than 2.5 standard deviations away from the group mean (for correct trials) were removed.

Before conducting each RT and ER analysis, we checked if the data were normally distributed using a Shapiro-Wilk test. When it was not, non-parametric alternatives were used instead of the planned parametric tests. For ANOVAs that yielded a significant interaction, we computed follow-up tests and used the Holm-Bonferroni method to correct for multiple comparisons.

RTs were first analyzed to check for the presence of an imitation effect. To do so, we ran a mixed 2x2 ANOVA with the within-subjects factor Effector Compatibility and the between-subjects factor Stimulus on the data collapsed across all three pilot studies. Next, we computed interference scores to compare the magnitude of the imitation effect between pilot studies. The interference scores were obtained by subtracting each participant's mean RT on compatible trials from their mean RT on incompatible trials. For our planned analysis, we computed four interference scores for each participant: two in response to the experimenter's hand (one following synchronous movement, one following asynchronous movement), and two

in response to their task partner's hand (one following synchronous movement, one following asynchronous movement). These scores were entered into a 2x2 repeated-measures ANOVA with the factors Hand (experimenter vs. partner) and Condition (synchrony vs. asynchrony).

ERs were computed by obtaining the number of errors participants made, indexed by trial type, and calculating the percentage of incorrect trials for each trial type. The resulting percentages were then entered into the analyses. As in the RT analysis, we first analyzed the ERs to check for an imitation effect. To do so, we compared ERs on effector-compatible trials with ERs on effector-incompatible trials. Next, we performed our planned analysis. To do so, we computed difference scores for synchronous and asynchronous coordination blocks. Difference scores were calculated by subtracting the error rate on compatible trials from the error rate on incompatible trials.

IOS Ratings

Participants' IOS ratings for their task partner and the experimenter were compared after synchronous and asynchronous coordination blocks in a repeated-measures 2x2 ANOVA with two-level within-subjects factors Condition and Other. This analysis was performed on the IOS ratings collapsed across all three pilot studies. As only pilot studies 2 and 3 included baseline IOS ratings, analyses which control for participants' baseline ratings of overlap are reported in the supplementary information.

Questionnaire Responses

Participants' responses to items in the final questionnaire which required a rating on a 7-point Likert scale (1-7) were entered into the analysis directly. Participants' ratings of liking (experimenter 1 vs. their task partner), of difficulty of the coordination task (synchronous vs. asynchronous blocks), of enjoyment of the coordination task (synchronous vs. asynchronous blocks), and of difficulty ignoring the depicted finger movement in the automatic imitation task

(experimenter 1's hand vs. task partner's hand) were compared using paired-samples Student's *t* tests. Participants' general ratings of enjoyment of the coordination task, difficulty of the coordination task, difficulty of the automatic imitation task, and of difficulty ignoring the depicted finger movement in the automatic imitation task were compared between pilot studies using a one-way ANOVA with the three-level between-subjects factor Pilot.

Results

Coordination Data

Participants' movements were significantly more coordinated in synchronous compared to asynchronous coordination blocks in all pilot studies (all $p \leq .008$; see supplementary information), suggesting that our manipulation was successful.

Automatic Imitation Data

Two participants made mistakes on more than 10% of trials and were therefore excluded from the analysis. One participant did not pass the recognition test and was therefore excluded from the analysis. In the remaining participants' data sets, 72 trials were removed, as they were coded as invalid or because they timed out (2.03% of trials). 37 trials (1.04% of trials) were coded as incorrect and excluded from the RT analysis. These trials were used for the ER analysis. 120 of the correct trials (3.38% of trials) were removed, as they were more than 2.5 standard deviations away from the group mean (0.53s). The remaining correct trials were included in the RT analysis.

RTs

Before testing for a modulation of imitative tendencies, we ran a paired Student's *t* test comparing participants' RTs in response to compatible and incompatible trials across all three pilot studies to test for the presence of an imitation effect. Compatible trials are those in which

participants are prompted to respond with their index finger (via the number cue) while being shown an index finger movement, as well as trials in which participants are prompted to respond with their middle finger while being shown a middle finger movement, whereas incompatible trials are those in which participants are shown a middle finger movement together with the number cue prompting an index finger response, as well as trials in which participants are shown an index finger movement together with the number cue prompting a middle finger response. The analysis showed no significant difference between participants' mean RTs in response to compatible ($0.51 \pm 0.08\text{s}$) and incompatible trials ($0.52 \pm 0.08\text{s}$): $t(39) = -0.34$, $p = 0.74$, $d = -0.05$. This result was further supported by a paired-samples Bayesian t test, showing moderate evidence for H_0 : $\text{BF}_{01} = 5.55$. This suggests that there was no overall effect of imitative compatibility across our three pilot studies.

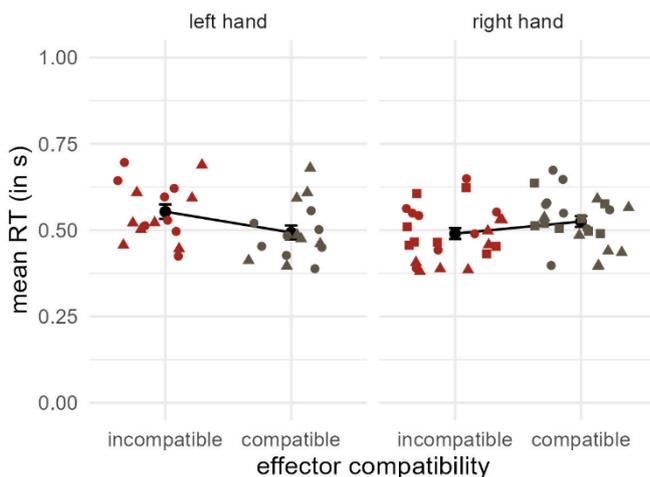
However, the analyses of the individual pilot studies (see supplementary information) suggest that this may be related to the different stimulus images used in pilot studies 1 and 2. We therefore followed up on this result by computing a mixed 2x2 ANOVA with within-subjects factor Effector Compatibility and between-subjects factor Stimulus on the collapsed data from pilot studies 1 and 2. The factor Effector Compatibility refers to the compatibility between observed and cued finger movement; the factor Stimulus refers to the hand participants were shown in the task (left or right). In left hand images, effector compatibility and spatial compatibility were confounded, as the images mirrored participants' hands. For right hand stimuli, effector compatibility and spatial compatibility were in conflict: Trials that were effector-compatible were spatially incompatible and vice versa.

The analysis revealed no main effect of Effector Compatibility $F(1,30) = 3.15$, $p = 0.86$, $\eta^2 = 0.1$, or of Stimulus: $F(1,30) = 0.53$, $p = 0.47$, $\eta_p^2 = 0.02$, but there was a significant interaction between Effector Compatibility and Stimulus: $F(1,30) = 49.41$, $p < .001$,

$\eta_p^2 = 0.62$.² Post-hoc comparisons using Holm correction revealed that participants who were shown images of a left hand were significantly faster on compatible (0.49 ± 0.08 s) than incompatible trials (0.55 ± 0.08 s): $t = -6.23$, $p_{adj} < .001$. However, participants who responded to right hand images were significantly faster on incompatible (0.49 ± 0.08 s) compared to compatible trials (0.52 ± 0.09 s): $t = 3.72$, $p_{adj} = .004$.³ This suggests that while effector compatibility did influence participants' responses overall, participants' RTs were more affected by spatial compatibility (see Figure 6).

Figure 6

Compatibility Effect by Stimulus Group (RTs)



Note. Mean RTs in response to effector-compatible (dark gray) compared to effector-incompatible (dark red) trials in each Stimulus group. The left panel depicts mean RTs for participants who responded to left-hand images; the right panel depicts mean RTs for participants who responded to right-hand images. Each dot represents an individual

² This was supported by a Bayesian ANOVA, which showed that the best model was one that included both factors and the interaction term: $P_{M|data}=1$, and that this model was 27497.26 times more likely to be true than the null model.

³ The outcome remained the same when the data from pilot study 3 were included in the analysis: $t=-4.56$, $p_{adj}<.001$.

participant's mean RT. Round dots represent participants in pilot study 1; triangular dots represent participants in pilot study 2; square dots represent participants in pilot study 3. Black dots represent participants' mean RT in that trial type; black brackets represent the standard error.

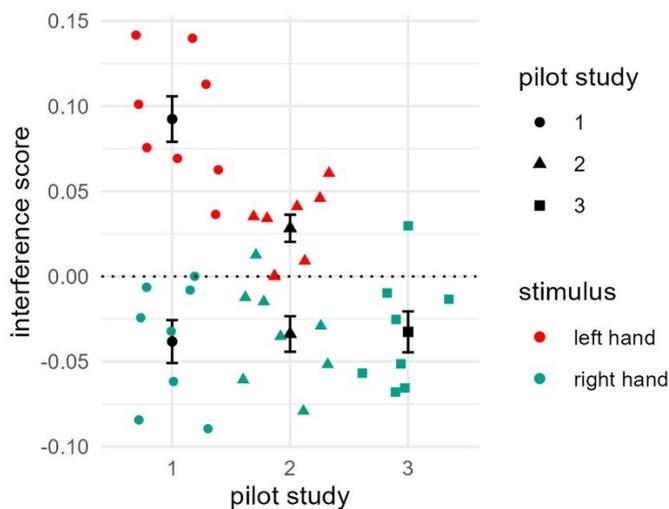
To see whether this effect was comparable across pilot studies, we computed interference scores for each participant and compared these within each Stimulus group. A positive interference score indicates that participants were faster in response to compatible trials than incompatible trials, demonstrating the presence of an effector compatibility effect. A negative interference score shows that participants were faster on incompatible trials compared to compatible ones, suggesting an inversion of the effect. As effector compatibility and spatial compatibility are confounded in the left hand stimuli, positive interference scores in the left hand Stimulus group may also be due to an effect of spatial compatibility. However, in the right hand stimuli, effector compatibility and spatial compatibility were opposed. Negative interference scores in this group therefore point towards the presence of a spatial compatibility effect.

We computed a one-way ANOVA with the three-level between-subjects factor Pilot to compare participants' interference scores for right hand images across pilot studies. The results showed no significant difference between pilot studies: $F(2,21) = 0.07, p = 0.94, \eta^2 = 0.01$ (see Figure 7). This result was further supported by a Bayesian ANOVA, which showed moderate evidence for H_0 : $BF_{01} = 3.97$. This suggests that participants' responses to right hand stimuli did not differ between pilots. To test whether participants in this group were more affected by spatial compatibility or by effector compatibility, we computed a one-sample Student's *t* test comparing the collapsed data to 0. The results showed that participants' interference scores were significantly below 0 (-0.04 ± 0.03): $t(23) = -5.37, p < .001, d = -1.1$. This was further

supported by a Bayesian t test, showing extreme evidence for H_1 : $BF_{10} = 1230$. Together, these results suggest that participants in all three pilot studies were substantially affected by spatial compatibility in their responses to right hand stimuli.

Figure 7

Interference Scores



Note. Interference scores (RT on incompatible trials minus RT on compatible trials) in each pilot study (x axis) by Stimulus group. Each dot represents one participant's mean interference score across all trials. Red dots represent responses to left hand stimuli; green dots represent responses to right hand stimuli. Round dots represent participants in pilot study 1; triangular dots represent participants in pilot study 2; square dots represent participants in pilot study 3. Black dots represent the mean interference score in that group; black brackets depict the standard error.

We also tested whether participants' responses to left hand stimuli differed between pilot studies by computing an independent-samples Student's t test to compare their interference scores between pilot studies 1 and 2. The results revealed that participants in pilot study 1 showed greater interference scores (0.09 ± 0.04) in response to left hand stimuli

compared to participants in pilot study 2 (0.03 ± 0.02): $t(1,14) = 4.12, p = .001, d = 2.06$. These results were further supported by a Bayesian t test, showing strong evidence for H_1 : $BF_{10} = 28.52$. This suggests that participants' responses to left hand stimuli substantially differed between pilot studies 1 and 2, although the stimuli were identical.⁴

Finally, we ran an ANOVA to see whether participants' responses differed after synchronous compared to asynchronous movement blocks, and whether they responded differently to the experimenter's and their task partner's hand. We compared participants' interference scores in response to their task partner's and the experimenter's hand in a 2x2 repeated-measures ANOVA with the factors Model (experimenter vs. partner) and Condition (synchrony vs. asynchrony). The results showed no significant main effect of Model: $F(1,39) = 0.29, p = 0.59, \eta^2_p = 0.01$, or of Condition: $F(1,39) = 0.62, p = .44, \eta^2_p = 0.02$, and the interaction was not significant either: $F(1,39) = 0.0002, p = .99, \eta^2_p = 0.000004$.⁵ This suggests that there was no modulation of imitative tendencies following synchronous compared to asynchronous movement and that participants did not respond differently to their task partner's compared to the experimenter's hand.

ERs

As in the RT analysis, we first tested for an imitation effect by comparing participants' error rates on effector-compatible and effector-incompatible trials. Given the differences between groups in the RT analysis, we computed separate paired Student's t tests for each Stimulus group. The results for the group that responded to left-hand images showed that participants made significantly more errors on effector-incompatible ($6.67 \pm 3.17\%$) compared

⁴ We followed up on this result by testing whether the interference scores in each of these pilot studies differed from 0 using one-sample Student's t test. The analyses confirmed that they did, both in pilot study 1: $t(7)=6.91, p<.001, d=-2.44$ ($BF_{10}=138.24$), and in pilot study 2: $t(7)=3.55, p=.009, d=-1.26$ ($BF_{10}=6.9$).

⁵ This was further supported by a Bayesian ANOVA, which showed that the best model was the null model ($P_{M|data}=0.63$) and that the null model was 3.76 times more likely to be true than the next best model.

to effector-compatible trials ($1.39 \pm 2.18\%$): $t(14) = -4.83, p < .001, d = -1.25$. This analysis was also confirmed by a Bayesian paired t test, showing extreme evidence for H_1 : $BF_{10} = 121.58$.⁶ Conversely, the group that responded to right-hand images made fewer errors on effector-incompatible ($1.8 \pm 2.75\%$) compared to effector-compatible trials ($4.07 \pm 4.79\%$), but this comparison was not statistically significant: $t(21) = 2.07, p = .051, d = 0.44$. The results of a Bayesian paired t test for this group show anecdotal evidence for H_1 : $BF_{10} = 1.33$. Here, the general trend of the ER results mirrors our RT finding: Participants in the left-hand Stimulus group made more errors in response to effector-incompatible trials, whereas participants in the right-hand Stimulus group made more errors in response to effector-compatible trials. This suggests that participants' ERs were also affected by spatial compatibility, similarly to their RTs.

We then tested for a modulation of ERs in response to the task partner's hand in synchronous compared to asynchronous blocks. As the difference scores were not normally distributed ($p = .008$), we computed a Wilcoxon signed-rank test to compare participants' responses to their task partner's hand after synchronous and asynchronous coordination. The results showed no significant difference in participants' ERs in response to their task partner's hand after synchronous (-0.68 ± 7.45) compared to asynchronous coordination (1.35 ± 7.98): $Z = 106.5, p = .2, r = -0.29$.⁷ Conducting separate analyses for each Stimulus group showed that in the group that responded to left hand images, participants' difference scores in response to their task partner's hand were significantly larger in asynchronous (7.22 ± 5.33) compared to synchronous blocks (2.78 ± 5.14): $Z = 16.5, p = .04, r = -0.64$ ($BF_{10} = 3.07$), whereas in the group that responded to right hand images, participants' difference scores in synchronous (-3.03 ± 7.95) and asynchronous blocks (-2.65 ± 6.99) did not differ: $Z = 32, p = .96, r = -0.03$

⁶ For individual analyses of pilot 1 than pilot 2 see supplementary information.

⁷ A Bayesian Wilcoxon signed rank test similarly showed anecdotal evidence for H_0 : $BF_{10} = 0.46$.

($BF_{10} = 0.25$). Together with the results of the RT analysis, this suggests that synchronous coordination did not lead to a stronger tendency to imitate a task partner's movements overall, but that participants' ERs, specifically in response to left hand stimuli, were differentially affected by our manipulation.

IOS Ratings

We computed a repeated-measures 2x2 ANOVA on participants' IOS ratings with the factors Condition (synchrony vs. asynchrony) and Other (experimenter vs. task partner), across all three pilot studies. Contrary to our predictions, we found no significant effect of Condition: $F(1,39) = 0.22, p = .64, \eta_p^2 = 0.006$, or Other: $F(1,39) = 0.49, p = .49, \eta_p^2 = 0.012$, and the interaction was also not significant: $F(1,39) = 0.79, p = .38, \eta_p^2 = 0.02$.⁸ This indicates that participants' perceived self-other overlap with their task partner and the experimenter did not change between synchronous and asynchronous movement blocks.

However, the IOS scale is typically used between-subjects, and when it is used in a within-subject design, it usually does not refer to the same individual more than once. For this reason, we performed an additional analysis of the data from the first block only. Since the data were not normally distributed (all $p \leq .001$), we computed a Mann-Whitney U test comparing participants' IOS ratings for their task partner in the synchrony compared to the asynchrony condition. Participants in the synchrony condition indicated higher overlap with their task partner ($45.3 \pm 42.72\%$) than participants in the asynchrony condition ($17.85 \pm 21.3\%$), but the difference between conditions was not significant: $U = 143.5, p = .12, r = -0.28$ (see Figure 8).⁹

We conducted the same analysis on participants' IOS ratings for the experimenter and found

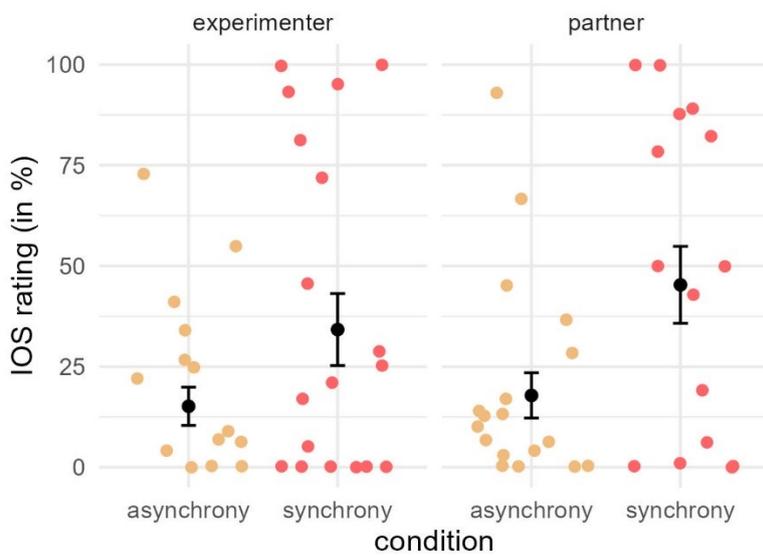
⁸ These results were confirmed by a Bayesian repeated measures ANOVA, which showed moderate evidence for H0 compared to a model including only Other ($BF_{01}=4.21$) or a model including only Condition ($BF_{01}=4.8$), strong evidence for H0 compared to a model that included both factors ($BF_{01}=21.07$) and very strong evidence for H0 compared to a model that included both factors and the interaction term ($BF_{01}=49.42$).

⁹ A Bayesian Mann-Whitney U test revealed anecdotal evidence in support of H0: $BF_{01}=1.57$.

also no significant difference between participants' perceived overlap in the synchrony ($34.2 \pm 39.98\%$) compared to the asynchrony condition ($15.15 \pm 21.3\%$): $U = 160.5$, $p = .27$, $r = -0.2$.¹⁰ This suggests that moving in synchrony with another person did not substantially affect ratings of perceived self-other overlap.

Figure 8

IOS Ratings – Block 1



Note. IOS ratings for the experimenter (left panel) and task partner (right panel) after completing the first condition of the coordination task: asynchronous (orange) or synchronous (pink) movement. Each colored dot represents one participant's rating of perceived overlap in the respective condition: orange dots represent IOS ratings in the asynchrony condition; pink dots represent IOS ratings in the synchrony condition. Black dots represent the mean IOS rating in that condition; black brackets depict the standard error.

¹⁰ A Bayesian Mann-Whitney U test revealed anecdotal evidence in support of H1: $BF_{10}=1.22$.

Final Questionnaire

First, we compared participants' ratings of liking for experimenter 1 and for their task partner. As the data were not normally distributed, we computed a Wilcoxon signed rank test. The results showed that participants rated their liking of the experimenter higher (5.88 ± 1.16) than their liking of their task partner (5.13 ± 1.3): $Z = 193, p < .001, r = 0.84$.¹¹ We tested both ratings against the median value of the scale (4) in two one-sample Wilcoxon signed rank tests and found that participants' ratings of liking for experimenter 1 were significantly higher than the scale average: $Z = 709, p < .001, r = 0.91$, and that the ratings of liking for their task partner were too: $Z = 458.5, p < .001, r = 0.2$. This suggests that that participants liked both their task partner and experimenter 1, but that, overall, participants liked experimenter 1 more than they liked their task partner.

Next, we analyzed participants' ratings of the difficulty of the coordination task. Here, we first computed a paired t test to compare their ratings of difficulty for synchronous and asynchronous coordination blocks. The results demonstrate that participants did not find it significantly more difficult to coordinate their movements in synchronous compared to asynchronous coordination blocks: $t(39) = -0.75, p = .46, d = -0.12$. A Bayesian paired samples t test confirmed that participants' ratings of difficulty of the coordination task did not differ between conditions: $BF_{01} = 4.51$. We then compared participants' overall ratings of the difficulty of the coordination task between pilot studies in a one-way ANOVA with the three-level between-subjects factor Pilot. The results showed that these ratings did not significantly differ between pilots: $F(2,37) = 1.01, p = .38, \eta^2 = 0.05$ ($BF_{01} = 2.69$). Participants' difficulty ratings for the coordination task overall were also significantly below the median scale value

¹¹ A Bayesian Wilcoxon signed-rank test showed extreme evidence for H1: $BF_{10}=180.39$.

of 4: $Z = 56$, $p < .001$, $r = -0.77$ (see Supplementary Table 1). Together, these results suggest that participants did not find the coordination task difficult to complete.

We also analyzed participants' enjoyment of the coordination task. First, we compared their ratings of enjoyment of synchronous compared to asynchronous coordination blocks. As the data were not normally distributed ($p = .04$), we computed a Wilcoxon signed-rank test. The results showed that participants did not enjoy synchronous coordination more than asynchronous coordination: $Z = 192.5$, $p = .27$, $r = -0.22$ ($BF_{01} = 3.8$; see Supplementary Table 2). Next, we compared participants' enjoyment ratings of the coordination task overall between pilot studies using a one-way ANOVA with three-level between-subjects factor Pilot. The results showed that participants' enjoyment of the coordination task did not significantly differ between pilots: $F(2,37) = 1.17$, $p = .32$, $\eta^2 = 0.06$ ($BF_{01} = 2.35$). Participants' enjoyment ratings did not significantly differ from the scale median: $Z = 322$, $p = .14$, $r = 0.2$ ($BF_{01} = 1.5$), suggesting a neutral attitude towards the task (across both conditions).

Finally, we analyzed participants' ratings of the difficulty of the automatic imitation task. First, we compared participants' overall ratings of the difficulty of the task between pilot studies using a one-way ANOVA with three-level between-subjects factor Pilot. The results showed that participants' general ratings of the difficulty of the task did not differ between pilots: $F(2,37) = 0.42$, $p = .66$, $\eta^2 = 0.02$ ($BF_{01} = 4.15$). Next, we performed the same analysis on participants' ratings of how difficult they found it to ignore the depicted finger movement. This analysis also showed no difference between pilots: $F(2,13) = 0.52$, $p = .6$, $\eta^2 = 0.03$ ($BF_{01} = 3.91$). We also analyzed participants' ratings of how difficult it was to ignore the depicted finger movement when it was their task partner's hand compared to the experimenter's hand. As the data were not normally distributed, we used a Wilcoxon signed-rank test.. We did not find a difference between participants' ratings for the experimenter's hand and participants'

ratings for their task partner's hand: $Z = 18.5$, $p = .11$, $r = 0.76$ ($BF_{01} = 1.48$; see Supplementary Table 3). Overall, participants' ratings of difficulty of the automatic imitation task were below average: $Z = 130$, $p < .001$, $r = 0.63$ ($BF_{10} = 74.7$). This suggests that participants did not find the automatic imitation task difficult to complete.

Discussion

Across three pilot studies, we aimed to investigate the effect of interpersonal synchrony on self-other overlap. In each pilot study, participants performed a coordination task, in which they moved either synchronously or asynchronously with a partner. This task was followed by two measures of self-other overlap: an automatic imitation task to quantify participants' imitative tendencies as an index of motor-level self-other overlap, and IOS ratings to measure participants' perceived self-other overlap. We predicted that participants would show higher levels of motor-level and perceived self-other overlap with their partner following synchronous compared to asynchronous movement.

While all three pilot studies used the same metronomes to pace participants in the first 20 seconds of each trial, the movements participants were instructed to perform differed between pilot studies. The coordination task in pilot studies 1 and 2 involved continuous movement: in the first pilot study, participants stood side by side facing a large touchscreen and moved a virtual knob vertically along a track; in the second pilot study, participants stood opposite each other, both facing the touchscreen, and moved a virtual knob horizontally along a track. The coordination task in the third pilot study was discrete: participants sat side by side facing a tablet and tapped on a virtual button. Although different mechanisms are thought to be involved in continuous and discrete forms of synchronized movement (Huys et al., 2008a; Lorås et al., 2012), which could have affected coordination success in our different tasks, we

found our manipulation to be successful in all three pilot studies: participants' movements were more closely coordinated in synchronous compared to asynchronous coordination blocks.

Our first prediction concerned motor-level self-other overlap: we expected participants to show stronger imitative tendencies in response to their task partner's hand. However, as this prediction referred to a modulation of imitative tendencies (as in Cross et al., 2021), we first aimed to establish the existence of an imitation effect. Although it has been argued that automatic imitation is a robust phenomenon (Cracco et al., 2018), we did not find evidence for the presence of an imitation effect in all pilot studies, and when it was present, it was overshadowed by a much larger effect of spatial compatibility. This was reflected in a consistent pattern of results: participants showed an imitation effect when responding to left-hand images (where imitative compatibility and spatial compatibility were confounded), but when responding to right-hand images (where imitative and spatial compatibility were opposed) the effect was inverted.

We anticipated that spatial compatibility would affect participants' responses (and therefore included the between-subjects manipulation of Stimulus in pilots 1 and 2). However, we did not anticipate its magnitude. While the results of pilot 1 can be seen as consistent with an interpretation suggesting the presence of an imitation effect but holding that participants were more affected by spatial compatibility (see Bertenthal et al., 2006), the results of pilot 2 do not support this view. In the latter, we did not find an effect of effector compatibility, and participants' imitative tendencies could be explained by spatial compatibility alone.

This is surprising, as we expected the changes to the coordination task in pilot study 2 compared to pilot study 1 to lead to stronger imitative tendencies. Whereas there was a change in perspective on the task partner's hand between the coordination task and the automatic imitation task in pilot 1, this switch in perspective was eliminated in pilot 2 by changing the

spatial arrangement of the coordination task. We expected this consistency in perspectives to induce a stronger tendency to imitate the task partner's hand. Instead, we found that the effect was diminished.

There could be several reasons for this. It is possible that the novel perspective on the hands in the automatic imitation task (compared to the coordination task) in pilot 1 enhanced the imitation effect, whereas the consistency of perspective between the tasks in pilot 2 may have led to habituation, thereby reducing the effect. It is also possible that the coordination task in pilot 2 highlighted the salience of left-right movements, as participants moved the virtual knob horizontally and not vertically (as in pilot 1). This may have primed participants to be more sensitive to (left-right) spatial aspects of the movement than to the movement effector. Finally, it is possible that the difference in perspective on the task between participants in pilot 2 led to less self-other overlap with the task partner overall – particularly as it was not necessary to take the other person's perspective to successfully complete the coordination task.¹²

To rule out an effect of horizontal compared to vertical movements in the coordination task, we used a tapping paradigm in pilot 3. This task had the added benefit that the movements performed during the coordination task were very similar to the ones participants performed in the automatic imitation task. We expected this to increase participants' imitative tendencies. Pilot study 3 also differed from the other pilot studies in that participants did not wear gloves and they were actually shown pictures of the experimenter's and their task partner' (i.e., the confederate's) hands in the automatic imitation task. We expected this change to further increase participants' motor system activation in the automatic imitation task.

¹² Synchronization of movements in this task referred to the position of the knob on the screen, not in relation to the task partner's body. This was particularly salient in this pilot study, as in order to synchronize their movements, one participant had to move towards their right while the other participant moved to their left. This mismatch may further have highlighted the spatial dimensions of participants' movements.

However, we did not find a stronger imitation effect in pilot study 3 compared to the other two pilot studies. As we only used right-hand images as stimuli in the automatic imitation task in this pilot study, in which spatial compatibility and effector compatibility are opposed, and lack participants' responses to matched left hand images as a contrast, it is not possible to estimate to what extent effector compatibility may have influenced participants' responses. However, the absence of an effect of effector compatibility along with the negative interference scores in pilot study 3, clearly demonstrate that spatial compatibility had a stronger effect on participants' imitative tendencies than effector compatibility.

While the changes to the task in pilot study 3 should have ruled out the possibility of a left-right spatial priming in the coordination task, the overall layout of the task was comparable to pilot study 2. Participants sat facing their task partner, so there was a consistency of perspective on the task partner's hand between the coordination task and the automatic imitation task. This may have led to a habituation effect and thereby reduced the extent to which participants were affected by the depicted finger movements. In pilot study 3, participants' perspectives on the task differed from their task partner's (as in pilot 2), and they were not required to take their task partner's perspective in order to synchronize their movements. This may also have enabled a better self-other distinction.

Overall, these results highlight the role of spatial compatibility in automatic imitation paradigms and emphasize the importance of controlling for it. While the use of both right- and left-hand stimuli has been suggested as a way of addressing this issue (Cracco et al., 2018), other paradigms have been developed that may allow for a more nuanced differentiation between spatial and imitative compatibility (e.g., Catmur and Sowden, 2015). The results presented here also hint at a possible role of habituation to, or novelty of the particular

perspective from which the stimuli are shown in modulating imitative tendencies. Testing this proposal may provide an avenue for future research.

Our prediction regarding motor-level self-other overlap was based on the assumption of an imitation effect and referred to its modulation. Since we did not find robust evidence for the former, a modulation is somewhat difficult to explain. While we did not find any difference between synchronous and asynchronous coordination blocks in the analysis of our RT data or in the analysis of the ER data for participants who responded to right hand images, we did find a difference in the ER data for participants who responded to left hand images. As these participants responded to stimuli, in which effector compatibility and spatial compatibility were confounded, it is not clear which of these effects was affected by our manipulation.

Moreover, it is worth noting that the effect we find in the ER data of the left hand Stimulus group goes against our predictions. We expected participants to experience more self-other overlap with their task partner after moving in synchrony with them, resulting in stronger imitative tendencies in response to their hand in the automatic imitation task following synchronous coordination. However, in our data we find the opposite: Participants showed stronger imitative tendencies for their task partner's hand following asynchronous coordination. A possible explanation for this effect might be that asynchronous coordination was more effortful than synchronous coordination, suggesting higher cognitive load during asynchronous coordination, which may have led to worse executive functioning and stronger imitative tendencies in the automatic imitation task. However, the facts that participants did not indicate that asynchronous coordination was more difficult than synchronous coordination, and the automatic imitation task was not rated as being difficult overall speak against this interpretation of our data.

Furthermore, we chose to focus on participants' RTs over their ERs, primarily because RTs in automatic imitation paradigms have been linked to more covert imitative tendencies, whereas ERs in these tasks have been linked to more overt forms of imitation, and we expected our manipulation to target the former. Despite the medium effect size of our analysis, it seems unlikely that the modulation of ERs that we see in our data was the result of an overt imitation process, as it was restricted only to the errors a subset of our participants made and not present in other participants or in other measured indices of imitation in our study. Together with our small sample size overall, it seems unlikely that this finding reflects a substantial effect.

Overall, our results remain at odds with the findings of Cross et al. (2021), in which imitative tendencies in response to the task partner's hand were stronger following intentional synchronization compared to asynchronous movement. This discrepancy might be related to differences in study design. Cross et al. (2021) used a between-subjects design; participants only completed the automatic imitation task once for each model. While some studies report successful use of automatic imitation tasks in within-subject designs (e.g., Spengler et al., 2010), others do not (e.g., Khemka et al., 2021). Imitative tendencies have also been shown to change over time, as the number of trials increases (Catmur & Heyes, 2011; see Cracco et al., 2018), suggesting that participants' imitative tendencies during later stages of the experiment may differ from their initial tendencies, regardless of any other manipulations.

In addition to this, the automatic imitation task used by Cross et al. (2021) used videos of right-hand finger movements as stimuli, not a sequence of images. Although they do not report testing for an imitation effect, their interference scores, on average, appear to be positive. This indicates that an imitation effect likely was present in their data and makes the modulation they find consistent with their predictions. Although videos have not been shown to lead to stronger imitation effects than sequences of images (see Cracco et al., 2018), it is possible that

features of the stimuli used in Cross et al. (2021) led to a stronger activation of participants' motor system during the task.

It is also possible that the crucial difference was not related to the automatic imitation paradigm but rather to the coordination task. In their study, the coordination task involved manipulating physical objects (wooden knobs on a slider box) rather than performing movements on a screen. Moreover, the task required larger body movements, as participants used both hands to move a wooden knob along a 60cm track, and both participants' tracks were built into the same wooden box. Although participants performed the coordination task on a shared device in the pilot studies reported here, there was no other obvious connection between their tracks. The fact that both knobs were part of the same apparatus in the study by Cross et al. (2021) may have placed a stronger emphasis on the joint nature of the task. It has also been argued that physical exertion might play a role in the effects of interpersonal synchrony (P. Sullivan et al., 2015; Tarr et al., 2015). The joint nature of the task, together with the larger movements, may thus have led to stronger motor activation and co-representation in Cross et al. (2021), thereby enhancing self-other overlap.

While these differences may account for a discrepancy in findings, the results of our three pilot studies reported here suggest that interpersonal synchrony does not have a robust effect on motor-level self-other overlap, and that this is likely not the primary mechanisms mediating its prosocial effects. This interpretation of our findings is not inconsistent with the interpretation put forward by Cross et al. (2021). In their studies, they find that interpersonal synchrony leads to an increase in self-other overlap only when it is brought about intentionally, not when it is incidental. They argue that an interpretation of the prosocial effects of interpersonal synchrony that invokes group membership as its key mechanism is more plausible than motor-level self-other overlap. While our results do not provide empirical

evidence supporting a group membership account, they do question the explanatory power of a proposal that relies primarily on motor-level self-other overlap.

Yet, our study also included a second outcome measure: perceived self-other overlap. We predicted that participants would indicate a higher percentage of overlap between circles that represented themselves and their task partner following synchronous compared to asynchronous coordination. We did not find this when looking at the overall data or at the data from the first block only. However, in our analysis of the first block only, we did see numerically higher ratings of perceived self-other overlap after synchronous compared to asynchronous coordination. Despite not reaching statistical significance, these data point in the same direction as previous results that have been reported in the existing literature on the effects of interpersonal synchrony on affiliation (see Cross et al., 2019) but suggests that the effect may not be as strong as previously assumed.

Although our pilot studies did provide some indication that intentional synchronization with a partner may increase self-other overlap, our results regarding an increase in both motor-level self-other overlap and perceived self-other overlap were inconclusive. The automatic imitation paradigm we used in these studies did not provide a clear-cut measure of self-other overlap, as participants' responses were strongly influenced by the spatial compatibility of the movement in the stimulus images. This was a significant limitation of our pilot studies and makes any modulation of imitative tendencies hard to interpret. For this reason, we made some changes to the design and ran a further experiment, aiming to provide more conclusive findings.

Experiment 1

Experiment 1 followed a between-subjects design (similar to Cross et al., 2021) to avoid any carryover effects. Participants completed the synchronous or the asynchronous condition of a coordination task similar to the one we used in pilot study 3, followed by measures of

motor-level self-other overlap, perceived self-other overlap, prosocial helping behavior, and liking. We measured motor-level self-other overlap using a different automatic imitation paradigm (adapted from Sowden & Catmur, 2015) to control for spatial compatibility in a more stringent way. Our measure of perceived self-other overlap was the same as in pilot study 3 (IOS ratings). We added a prosocial helping task with the aim of validating our task and comparing our results to previous findings (e.g., Cross et al., 2020; Kokal et al., 2011; Valdesolo & DeSteno, 2011). The questionnaire item about liking of the task partner and the experimenter was the same as in the pilot studies. However, as this experiment used a between-subject design, we were able to treat this questionnaire item as a dependent measure of affiliation. We predicted stronger imitative tendencies for the task partner's movements (i.e., more self-other overlap with the task partner), more perceived self-other overlap with the task partner, more helping behavior, and higher ratings of liking of their task partner in the group that moved in synchrony compared to the group that completed the asynchronous condition. The study preregistration can be found at <https://osf.io/p6buy> (last accessed 2023-10-03).

Methods

Participants

Based on the recommendation of at least 25 participants per condition (for a similar design; Gleibs et al., 2016), we aimed for a target sample size of 52, with 26 participants in each group. To reach this target sample size, we recruited 54 participants (38F, 16M, aged 19 to 40 years), two of which had to be excluded, from the Sona System database (<https://www.sona-systems.com/>, last accessed 2023-01-25) of Central European University (CEU) PU. Participants completed the experiment with a confederate who acted as their task partner. Two participants were excluded; one was excluded due to procedural error, and one was excluded because they were friends with the confederate.

All participants were right-handed English-speaking adults with normal or corrected-to-normal vision, no motor impairments, and no history of neurological impairments or diagnoses. The experiment took place in the laboratories of the Social Mind Center of CEU PU in Vienna. The study took approximately one hour to complete, and participants were reimbursed at a rate of 10 euros per hour, with 5 euros added to this rate if the study went overtime. All procedures used in these studies were approved by the Psychological Research Ethics Board (PREBO) at CEU PU Vienna.

Materials and Procedure

After providing informed consent, participants were told that experimenter 2 would take pictures of their hands and that these pictures would later be used as stimuli for one of the tasks. As in the pilot studies, experimenter 2¹³ took pictures of experimenter 1's right hand in resting position, with a raised index finger and with a raised middle finger. They then did the same for experimenter 1's left hand. The same procedure was repeated for the participant's and the confederate's right and left hands. Once all six pictures were taken of each person's hands, experimenter 1 dimmed the lights in the room to avoid reflection of the ceiling lights on the tablet during the coordination task and explained the structure of the experiment (see Figure 9).

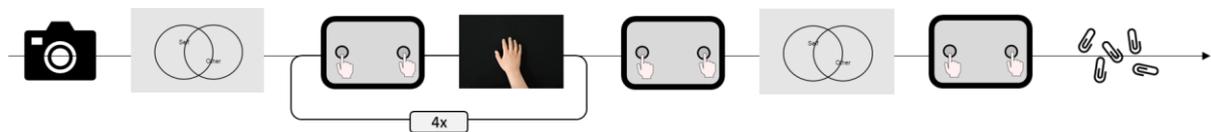
Participants started by completing baseline IOS ratings for the confederate and experimenter 1, respectively, then moved on to the practice phase of the coordination task. While participants completed the practice phase of the coordination task on the tablet with the confederate, experimenter 2 acted as if they were uploading images from the camera onto the laptops (for both the participant and the confederate). After the practice phase of the

¹³ Whereas the same person acted as experimenter 2 throughout all sessions of the pilot studies, this was not possible for this experiment; in some sessions this role was performed by different (female) experimenters. The procedure remained identical in all sessions.

coordination task was complete, participants completed a practice phase of the automatic imitation task. Then participants began the first block of the experiment. Each block of the experiment began with two trials of the coordination task, followed by an automatic imitation task (blocks 1-4), IOS ratings (block 5) or a short helping task and a final questionnaire (block 6). While participants completed the individual tasks, the confederate was instructed to act as if she were performing the same tasks on a laptop. A 65cm long and 55cm high occluder was placed between the tablet and the laptop the confederate was seated in front of, so participants could not see the confederate's screen, and her hand movements were also obstructed from their view.

Figure 9

Procedure



Note. This figure depicts the experimental procedure (from left to right). First, pictures were taken of participants' hands as well as the confederate's and experimenter 1's hands, then participants completed the baseline IOS ratings for their task partner and experimenter 1. This was followed by a practice phase (not depicted in the timeline) of the coordination task and the automatic imitation task, and by four experimental blocks in which participants completed the coordination task using either their left or their right hand, followed by the automatic imitation task with the same hand. After the last block of the automatic imitation task, participants completed a recognition test to see whether they could differentiate their task partner's and experimenter 1's hand. Then they completed two more blocks of the coordination task, one

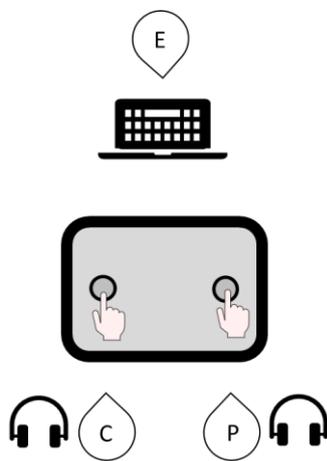
followed by IOS ratings for their task partner and experimenter 1, the other followed by a helping task (collecting paperclips) and a final questionnaire.

Coordination Task

The coordination task followed the same general structure in both conditions. Participants sat next to the confederate facing a 12.9-inch iPad Pro with two buttons displayed on it (see Figure 10). Experimenter 1 sat opposite the participant and the confederate behind a MacBook Pro. Participants performed the coordination task together with the confederate, who was instructed to act as though she were a participant in the study. The coordination task consisted in tapping on a virtual button on the tablet that was placed on the table between participant and confederate.

Figure 10

Coordination Task



Note. Participants (P) performed the coordination task on a 12.9-inch iPad Pro together with a confederate (C). They sat next to each other facing a tablet, and their task was to tap on a virtual knob. Experimenter 1 (E) sat behind a laptop, opposite the participant and the confederate, and controlled the start of each trial. Participants were instructed to move either synchronously or asynchronously with the confederate, depending on which condition they were assigned to. For

the first 20 seconds of each trial, they were paced by a metronome played through a set of headphones, then they continued without the metronome while trying to keep the same pace.

At the beginning of each trial, the participant and the confederate heard a metronome through a set of headphones (either 59bpm or 83bpm) and were instructed to move in time with it. The movements were discrete; participants were instructed to tap on a virtual button each time they heard a tone. The metronome was only present for the first 20 seconds of each trial, then it stopped. Participants were instructed to continue moving at the pace of the metronome once it had stopped until they received a signal indicating that the trial was over.

In the synchronous condition, the participant and the confederate heard identical metronomes (both heard 59bpm or both heard 83bpm). They were additionally instructed to try to move in time with each other once the metronome stopped. In the asynchronous condition, the participant and the confederate heard different metronomes (one heard 59bpm, the other heard 83bpm). Participants in this group were additionally instructed to try to stick to their own pace. As in the pilot study, the emphasis was placed on the second condition-specific part of the instructions to ensure that participants understood the interpersonal nature of the task.

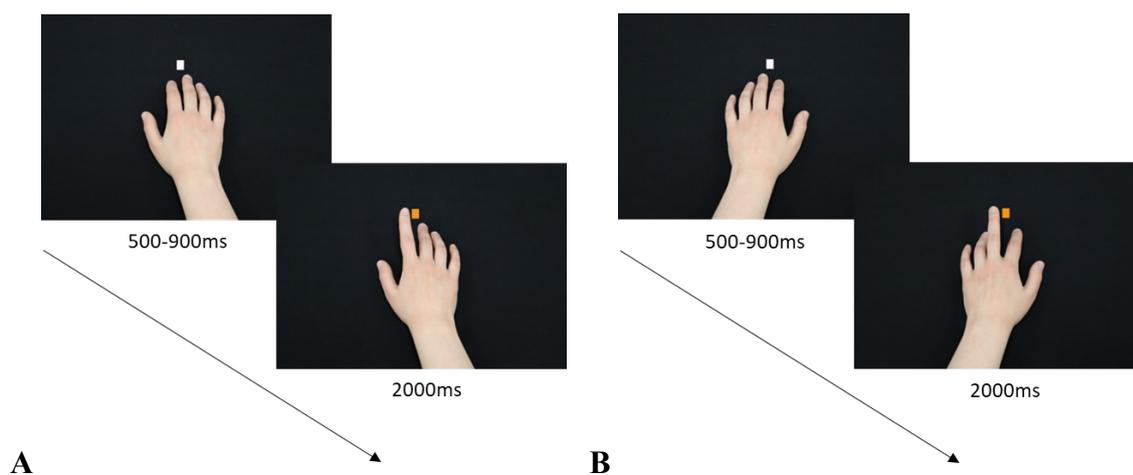
Participants completed four practice trials (two with their right hand, followed by two with their left), which lasted 40 seconds each, before starting the experimental blocks. Experimental blocks consisted of two trials à 60 seconds. Participants used either their left or their right hand in each block; the order was counterbalanced. There were six blocks of the coordination task in total; participants completed three blocks with their right hand and three with their left.

Automatic Imitation Task

We used a two-alternative forced-choice automatic imitation task adapted from Sowden & Catmur (2015), with key release serving as our dependent measure. Participants were shown right- and left-hand images from a first-person perspective, and their task was to respond to the color of a small rectangle that was displayed together with the hand on the screen (see Figure 11). As participants sat next to the confederate in this experiment (and not across from the confederate, as in pilot study 3), their perspective on their task partner's hand remained the same between the coordination task and the automatic imitation task.

Figure 11

Automatic Imitation Task: Trial Structure and Left- vs. Right-Hand Stimuli



Note. Example of right-hand (A) and left-hand stimulus images (B). In the automatic imitation task, participants were shown an image of a hand in resting position (top left image in both A and B) for 500 to 900ms, together with a white rectangle between the index and middle finger. They were then shown an image of the same hand with either a raised index finger (bottom right image in A) or a raised middle finger (bottom right image in B). The second image was displayed together with a colored rectangle (here: orange) for a maximum of 2 seconds, or until a response was made. When using their right hand to respond, participants were instructed to

lift their index finger in response to an orange rectangle and to lift their middle finger in response to a purple rectangle. The depicted finger movement was either compatible (A) or incompatible (B) with the cued response.

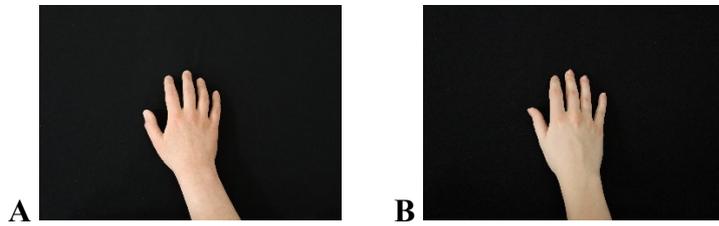
Participants were instructed to press two keys marked with white stickers on their computer keyboard (B and N) with the index and middle finger of their right or left hand while viewing the images, and to release either their index or their middle finger in response to the colored rectangle (i.e., to lift their index finger when the rectangle turned orange or yellow, and to lift their middle finger when it turned purple or blue). The keys were marked with white stickers in this experiment to avoid any conflict with the colors of the rectangle they were instructed to respond to. We also changed the response keys to align participants' hand placement with the location of the depicted index and middle fingers in the stimuli.

In this task, we manipulated the factors Imitative Compatibility (i.e., whether the depicted finger movement was compatible or incompatible with the prompted response), Hand (i.e., whether the depicted hand was congruent or incongruent with the hand participants were using to respond), and Model (i.e., whether participants responded to images of their task partner's or the experimenter's hands). The factors Hand and Model were blocked. Participants completed half of the blocks of the coordination task using their right hand to respond and half using their left; and they responded to images of their task partner's hand in half of the blocks and to images of experimenter 1's hand in the other (see

Figure 12). These factors were crossed, such that participants were presented with each combination of these two factors once.

Figure 12

Automatic Imitation Task: Model



Note. Images of experimenter 1's (A) and the confederate's hand (B). Participants responded to images of experimenter 1's hand in half of the blocks of the automatic imitation task, and they responded to images of their task partner's hand (i.e., the confederate's hand) in the remaining blocks. The images depicted in this figure were used for the recognition test after the final block of the automatic imitation task.¹⁴ The full set of stimuli can be found in the supplementary information.

To manipulate Imitative Compatibility, we varied the finger movement participants were shown together with the colored rectangle. In half of the trials, participants saw a movement that was compatible with the prompted response (e.g., a raised index finger presented together with an orange rectangle, as in Figure 11A). In the other half, they were shown a movement that was incompatible with the one they were prompted to perform (e.g., a raised middle finger together with an orange rectangle, as in Figure 11B). Compatible stimulus images also depicted the same hand as participants were instructed to use (i.e., participants were shown an image of a right hand performing the finger movement when they were using their right hand to respond, as in Figure 11A), whereas incompatible stimulus images depicted the other hand (i.e., participants were shown left-hand images when they were using their right

¹⁴ As recognition rates in the first 50% of our sample were at chance, we slightly modified the stimuli for the remaining participants such that either the experimenter's hand or the confederate's hand was marked, and we added a verbal emphasis on the hands in the instructions (see supplementary information).

hand to respond, as in Figure 11B). These compatible vs. incompatible trial types were presented in random order in each block.

Participants first completed two rounds of practice trials using their right hand, with each round consisting of eight trials, before repeating the same procedure with their left hand. In the first round, participants were shown an image of a hand in resting position together with a white rectangle between the index and middle finger (see top images in Figure 11) for 500 to 900ms. This image was then replaced by an image of the same hand, still in resting position, together with a colored rectangle between the hand's index and middle finger. The second image was shown for a maximum of two seconds and disappeared once one of the two response keys was released.

When participants completed the practice trials with their right hand, the rectangle turned either orange or purple. Participants were instructed to lift their index finger when they saw an orange rectangle and to lift their middle finger when they saw a purple rectangle. When participants completed the practice trials with their left hand, the rectangle turned either yellow or blue. Participants were instructed to lift their index finger when they saw a yellow rectangle and to lift their middle finger when they saw a blue rectangle. This round of practice trials was repeated as often as necessary until participants had learned the correct cue-response mapping and responded correctly on at least 75% of trials.

Participants then moved on to the second round of practice trials, in which they were shown the same image of a hand in resting position together with the white rectangle for 500 to 900ms, then an image of the hand with a colored rectangle, this time with a raised index or middle finger (as in the bottom images in Figure 11), for a maximum of 2 seconds. This round was also repeated as often as necessary until participants responded correctly on at least 75% of trials. After completing the practice phase with their right hand, participants moved on to

complete the practice phase with their left hand. After successful completion of the practice phase with their left hand, participants proceeded to start the first experimental block of the study (starting with the coordination task).

Participants performed four experimental blocks of the automatic imitation task; two in which they used their left hand to respond to the stimulus images (once showing their task partner's hand and once showing the experimenter's hand), and two in which they used their right (once to respond to images of their task partner's hand and once to respond to images of the experimenter's hand). The order of these blocks was counterbalanced and matched with the coordination task, such that participants always used the same hand in the automatic imitation task as they had used in the preceding block of the coordination task, but constrained such that the first two blocks were not performed with the same hand or in response to the same model.

Each block of the automatic imitation task consisted of 24 trials and was preceded by two trials of the coordination task. Compatible and incompatible trials were randomly intermixed in each block, and each trial followed the same structure as in the second round of practice trials: participants were shown an image of a hand in resting position together with a white rectangle for 500 to 900ms, then an image of the same hand with a raised index finger or middle finger together with a colored rectangle for a maximum of 2 seconds (or until participants responded).

At the end of the final block, participants were shown two images of hands and were asked to identify which one was their task partner's or the experimenter's hand (see Figure 12). This question was intended as a manipulation check to make sure that participants were able to match the hand with its identity. Participants were also asked to give a rating of how confident they were that their response was correct. We included this question to gauge how difficult it was for participants to differentiate the confederate's and the experimenter's hand.

Inclusion of Other in Self (IOS) Ratings

For the IOS ratings, we used a continuous version of the IOS scale designed by Aron et al. (1992), as in pilot study 3. Participants were instructed to arrange two circles, labeled ‘Self’ and ‘Other’, in the way that best described their relationship with their task partner, and with experimenter 1. The circles were arranged in such a way that they had an overlap of 50% at the start of each rating. Participants performed one rating of overlap with their task partner and one rating of overlap with experimenter 1 at the beginning of the experiment (after experimenter 2 finished taking pictures of their hands), as a baseline measure of perceived self-other overlap. The IOS ratings were repeated after the fifth block of the coordination task, to measure how synchronous or asynchronous movement affected their perceived self-other overlap. In both cases, the ratings for experimenter 1 and the confederate were performed in random order.

Helping Task

After the final block of the coordination task, the confederate knocked over a container filled with 40 paperclips, which spilled onto the floor between herself and the participant. In doing so, the confederate acted to generate the impression that this happened by accident. The confederate apologized and started collecting the paperclips closest to her chair. She was instructed to only collect 20 of the 40 paperclips. After 10 seconds, experimenter 1 said, “It’s ok. Don’t worry about it, I can take care of the rest.” The confederate then placed the container back on the table. Experimenter 1 took note of the participant’s behavior and counted the number of paperclips in the container at the end of each experimental session.

Final Questionnaire

At the end of the experiment, participants filled out a final questionnaire, as in pilot studies 1-3. They were asked to rate their liking of their task partner and experimenter, their enjoyment of the coordination task, the difficulty of the coordination task, and the difficulty of

the automatic imitation task – in general, with regard to ignoring the depicted finger movement, and with regard to ignoring the depicted finger movement when they responded to the experimenter’s and their task partner’s hand, respectively. All ratings were performed on a scale from 1 to 7. The questionnaire additionally included questions about participants’ experience with music and dance, probed their hypotheses as to what the experiment was about, and gave them the opportunity to provide comments and feedback.

Data Treatment

Coordination Data

To analyze the coordination data, we calculated asynchronies between the participant’s and the confederate’s tap onsets by computing the difference in time between each of the participant’s taps and the confederate’s tap that was closest in time to it, and not more than half an interval of the instructed metronome beat ahead or behind. We then calculated the mean asynchrony for each trial and used these values to calculate the mean asynchrony for each block. From the block means the mean asynchrony for each dyad. These mean asynchronies per dyad were then compared between groups using an independent-measures t test to check that our manipulation was successful and that participants followed the instructions.

We additionally explored the relationship between participants’ ratings of their experience with music and dance, as reported in the final questionnaire, and their performance in the coordination task, as reflected in their mean asynchronies with the confederate, as well as the relationship between the degree of coordination with the confederate in the coordination task and their ratings of self-other overlap with her. The outcome of these analyses is reported in the supplementary information.

Automatic Imitation Data

Our main dependent measures in the automatic imitation task were participants' response times (from the time the colored rectangle was shown) and their error rates (the percentage of trials in which they gave the incorrect response) as indicators of participants' imitative tendencies. Before analyzing participants' response times (RTs) and error rates (ERs), we first excluded trials that were coded as invalid (because participants lifted a finger before the cue image was shown) and trials which timed out (because participants did not respond). We also excluded participants who made mistakes on more than 10% of trials overall.

We then considered the remaining correct and incorrect trials independently. For correct trials, we first computed the group mean and standard deviation, and we excluded all trials in which participants' RTs fell more than two standard deviations away from the group mean. The remaining trials were then entered into the RT analysis. Here, we first computed a mean RT for compatible trials and for incompatible trials, respectively, for each participant to test for an effect of imitative compatibility. To do so, we computed a paired Student's *t* test comparing participants' responses to compatible and incompatible trials. This allowed us to check that the automatic imitation task elicited the predicted effect (i.e., to perform a manipulation check).

For our main analysis, we computed interference scores for each participant's responses to their task partner's and the experimenter's hand. To do so, we subtracted each participant's mean RT on compatible trials from their mean RT on incompatible trials for trials in which they responded to images of their task partner's hand and for trials in which they responded to images of the experimenter's hand, respectively. The interference scores were entered into a mixed 2x2 ANOVA with the between-subjects factor Condition (synchrony vs. asynchrony)

and the within-subjects factor Model (task partner vs. experimenter) to test for a modulation of participant's imitative tendencies.

Incorrect trials were analyzed following the same logic. We first computed each participant's mean ER for compatible and for incompatible trials and tested for a difference between these trial types to check that our manipulation worked. We then computed difference scores for trials in which participants responded to their task partner's hand and for trials in which participants responded to the experimenter's hand by subtracting each participant's mean ER on compatible trials from their mean ER on incompatible trials. The difference scores were then compared between conditions to test for a modulation of imitative tendencies.

Inclusion of Other in Self Ratings

To analyze participants' IOS ratings, we generated deviation scores by subtracting their baseline ratings for their task partner and for the experimenter from their final ratings for their task partner and for the experimenter (after the fifth block of the coordination task). We compared participants' deviation scores for their task partner and for the experimenter between conditions to test for an influence of our manipulation on perceived self-other overlap.

Helping Behavior

We recorded the number of paperclips participants collected after the confederate dropped the container and used this number as an index of participants' willingness to help. To test whether our manipulation affected participants' prosocial tendencies (i.e., their helping behavior), we compared these numbers between conditions.

Questionnaire Responses

As we predicted that our manipulation would affect participants' liking of their task partner (and possibly the experimenter), we compared participants' responses to these two

questionnaire items between blocks. We also analyzed the other questionnaire items; the analysis procedures and results for these items are reported in the supplementary information.

Results

Coordination Data

There was missing data for five participants.¹⁵ In these cases, blocks in which we did not have complete data for both trials were excluded from the analysis.¹⁶ The remaining data was used to calculate the mean asynchrony for each dyad. We used these mean asynchronies to perform a manipulation check and compared them between groups. As the variance in both groups was not equal ($p < .001$), we computed a Mann Whitney U test to compare participants' mean asynchronies between groups. The analysis revealed that participants' asynchronies were significantly smaller in the synchrony condition ($83 \pm 17\text{ms}$) compared to the asynchrony condition ($188 \pm 8\text{ms}$): $W = 676$, $p < .001$, $r = 1$ ($\text{BF}_{10} = 3145.48$). This suggests that our manipulation was successful, and participants' taps were more closely coordinated with the confederate's taps in the synchronous than in the asynchronous group.

Automatic Imitation Data

46 trials (0.92% of all trials) were coded as invalid (i.e., participants lifted their finger before the cue was shown), 14 trials (0.28% of all trials) timed out, and 272 trials (5.45% of all trials) were coded as incorrect. Eight participants (15.38%) made mistakes on more than 10% of trials and were therefore excluded from both the ER and the RT analysis.¹⁷ In addition to this, 168 trials (3.37% of all trials) were excluded from the RT analysis because they were

¹⁵ In each case, there was only data missing from one block. One participant did not complete the second trial in the block due to procedural error. In the other cases, participants completed all trials but due to a technical error their data was not saved correctly.

¹⁶ Here we deviate from our preregistered exclusion criteria by including participants who did not provide complete data sets in our analyses and only excluding parts of their data.

¹⁷ Seven of these participants were in the Synchrony condition.

either two standard deviations above or below the group mean.¹⁸ The remaining correct trials were entered into the RT analysis. Due to the high percentage of non-recognizers in our sample (see below), we did not follow our preregistered exclusion criteria but retained these participants in our sample instead.

Response Times

Before testing for a modulation of imitative tendencies, we checked whether there was an effect of imitative compatibility using a paired-samples Student's *t* test. The analysis showed that participants were significantly faster to respond to compatible trials (0.46 ± 0.05 s) compared to incompatible trials (0.48 ± 0.05 s): $t(43) = -5.11, p < .001, d = -0.77$ ($BF_{10} = 2741.85$), indicating that our manipulation of imitative compatibility was successful.

As participants completed half of the trials using their dominant hand and the other half using their non-dominant hand, we also tested for a difference between participants' mean RTs when using their right compared to their left hand before conducting our main analysis. A paired-samples Student's *t* test revealed that participants were significantly faster to respond with their left hand (0.47 ± 0.05 s) than with their right (0.48 ± 0.05 s): $t(43) = -3.56, p < .001, d = -0.54$ ($BF_{10} = 31.71$). To see whether this difference in RTs also affected participants' imitative tendencies we compared participants' mean interference score for their non-dominant and their dominant hand using a paired-samples Student's *t* test. The analysis showed no significant difference between interference scores when participants responded using their right hand (0.02 ± 0.03 s) compared to their left hand (0.01 ± 0.03): $t(43) = -1.83, p = .08, d = -0.28$ ($BF_{01} = 1.34$). This suggests that although participants were faster to respond when using their non-dominant hand overall, the effect of imitative compatibility did not substantially differ

¹⁸ We repeated all analyses using each participant's individual mean RT as a reference (rather than the group mean). For these analyses, 178 trials (3.57% of all trials) were excluded because they were more than two standard deviations away from the individual participant's mean RT.

between hands. The left- and right-hand data were therefore collapsed in the following analyses.

For our main analysis, we computed a 2x2 mixed ANOVA on participants' mean interference scores with the between-subjects factor Condition (synchronous vs. asynchronous) and the within-subjects factor Model (experimenter vs. partner). The analysis showed no significant main effect of Condition: $F(1,42) = 0.03, p = .87, \eta^2 = 0.0004$, no significant main effect of Model: $F(1,42) = 0.49, p = .49, \eta_p^2 = 0.01$, and no significant interaction between Condition and Model: $F(1,42) = 0.99, p = .32, \eta_p^2 = 0.02$ (see Table 1 for descriptive statistics). These results were further supported by a Bayesian mixed ANOVA, which showed that none of the alternative models provided a better explanation of the data than H0 (all $BF_{01} > 3; P_{H0|data} = 0.57$). This suggests that moving in synchrony with another person did not affect participants' imitative tendencies of their task partner's or the experimenter's hand.

Table 1

Interference Scores by Model and Condition

		Condition		
		Synchrony	Asynchrony	
Model	Partner	0.017 ± 0.022s	0.02 ± 0.029s	0.019 ± 0.026s
	Experimenter	0.018 ± 0.024s	0.013 ± 0.026s	0.015 ± 0.025s
		0.018 ± 0.023s	0.017 ± 0.027s	

Note. Rows in this table depict mean RTs (± SD) for each model (task partner and experimenter); columns show mean RTs (± SD) in each condition (synchrony and asynchrony groups).

We additionally performed an exploratory analysis of participants' mean RTs as well as their interference scores in each block. The outcome of these analyses can be found in the supplementary information.

Error Rates

As in the RT analysis, we first tested for an effect of imitative compatibility in the ER data. As the data were not normally distributed ($p = .03$), we did so using a Wilcoxon signed-rank test. The analysis showed that participants made significantly fewer errors on compatible trials ($1.42 \pm 1.78\%$) compared to incompatible trials ($4.36 \pm 3.62\%$): $W = 37.5$, $p < .001$, $r = -0.86$ ($BF_{10} = 4003.43$), demonstrating that the effect of imitative compatibility was reflected in participants' error rates as well.

We then compared participants' error rates, as well as the magnitude of the imitation effect, between blocks in which they used their dominant hand to respond and blocks in which they used their non-dominant hand to respond, before computing our main analyses. The analysis revealed no significant difference between error rates overall in blocks in which participants responded using their right ($3.6 \pm 3.31\%$) compared to their left hand ($2.98 \pm 3.06\%$): $t(43) = 1.11$, $p = 0.28$, $d = 0.17$ ($BF_{01} = 2.76$). Difference scores (i.e., the difference between participants' ER on compatible and incompatible trials) also did not differ between blocks in which participants responded using their right (0.02 ± 0.03) compared to their left hand (0.02 ± 0.02): $t(43) = -0.85$, $p = .4$, $d = 0.17$ ($BF_{01} = 3.89$). This suggests that error rates in blocks in which participants used their dominant hand and error rates in blocks in which participants used their non-dominant hand were comparable and that the magnitude of the effect of imitative compatibility did not substantially differ between these two kinds of blocks.

For our main analysis, we compared participants' mean difference scores for their task partner and for the experimenter between the synchronous and the asynchronous condition. As

the data were not normally distributed, we did so using Mann-Whitney U tests. The analyses showed that participants in the synchrony group did not show larger difference scores ($2.19 \pm 4.68\%$) than participants in the asynchrony group ($1.33 \pm 4.29\%$) in response to their task partner's hand: $W = 214.5$, $p = .58$, $r = -0.1$ ($BF_{01} = 2.78$). Difference scores for the experimenter's hand also did not differ between the synchronous ($4.17 \pm 5.73\%$) and the asynchronous condition ($4.17 \pm 6.13\%$): $W = 244.5$, $p = .87$, $r = 0.03$ ($BF_{01} = 3.29$). We also compared participants' difference scores in response to the task partner's and the experimenter's hand overall in a Wilcoxon signed-rank test. The results showed that participants' difference scores were significantly higher for the experimenter's hand (4.14 ± 2.55) compared to the task partner's hand (1.73 ± 0.15): $W = 322.5$, $p = .02$, $r = 0.48$ ($BF_{10} = 2.06$). This suggests that participants' imitative tendencies, as reflected in their error rates, were greater in response to the experimenter's hand, compared to their task partner's hand, independent of condition. However, contrary to our predictions, participants' difference scores were not modulated by synchronous movement coordination – neither in response to their task partner's hand nor in response to the experimenter's hand.

An exploratory analysis of participants' overall error rates and interference scores in each block can be found in the supplementary information.

Hand Recognition and Confidence

24 participants (46.15% of all participants) did not recognize their task partner's or the experimenter's hand at test.¹⁹ Overall, participants' recognition rate ($N = 52$, $k = 28$) did not significantly differ from chance performance (50%): $p = .68$ ($BF_{01} = 5.02$). Participants who passed the recognition test were more confident that their response was correct ($70.2 \pm 30.3\%$)

¹⁹ The recognition rate only slightly improved from an initial rate of 50% before modification of the stimuli (and instructions) to a recognition rate of 57.69% (i.e., 42.31% of participants still did not recognize the hand) after the introduction of a mark on the task partner's or the experimenter's hand.

than participants who failed to recognize the hand ($53.1 \pm 30.2\%$): $W = 446.5$, $p = .04$, $r = 0.33$ ($BF_{10} = 1.48$). This suggests that participants failed to recognize their task partner's or the experimenter's hand in the automatic imitation task but that correct identification of the hand was associated with somewhat higher levels of confidence in the correctness of the response.

Inclusion of Other in Self Ratings

As preregistered, we compared deviation scores (i.e., the difference between baseline and final IOS ratings) for the task partner and the experimenter between Synchronous and Asynchronous conditions. Since the data were not normally distributed, we computed Mann-Whitney U tests for deviation scores concerning the task partner and the experimenter, respectively. While participants' IOS ratings for their task partner, on average, showed an increase from baseline IOS ratings in the Synchronous group ($5.65 \pm 21.7\%$) and a decrease from baseline ratings in the Asynchronous group ($-4.92 \pm 18.3\%$), the difference between conditions was not significant: $W = 245.5$, $p = .09$, $r = -0.27$ ($BF_{01} = 1.19$). Deviation scores for the experimenter also did not significantly differ between groups: $W = 302.5$, $p = .68$, $r = 0.16$ ($BF_{01} = 3.06$), with participants showing an increase in IOS ratings, on average, in both the Synchronous ($3.04 \pm 24.1\%$) and the Asynchronous condition ($0.92 \pm 18.7\%$).²⁰ This suggests that our manipulation did not lead to a change in participants' perceived self-other overlap with their task partner or the experimenter.

To see whether participants' absolute ratings of perceived self-other overlap differed between conditions, we additionally conducted exploratory analyses of participants' final IOS ratings for the task partner and for the experimenter, respectively. The analysis of participants' ratings of overlap with their task partner did not reveal a significant difference between the

²⁰ One participant was excluded from this analysis due to an error in the procedure (this participant did not perform the baseline IOS rating for the experimenter before the first block of the coordination task).

Synchronous ($25.31 \pm 25.28\%$) and the Asynchronous group ($17.27 \pm 22.7\%$): $W = 258.5, p = .15, r = -0.24$ ($BF_{01} = 1.51$). Participants' IOS ratings with the experimenter also did not significantly differ between the Synchronous ($22.96 \pm 21.73\%$) and the Asynchronous ($15.35 \pm 20.36\%$) condition: $W = 238.5, p = .1, r = -0.27$ ($BF_{01} = 1.71$). This indicates that participants' ratings of perceived self-other overlap with their task partner and with the experimenter were not substantially affected by synchronous compared to asynchronous coordination.

Helping Behavior

As preregistered, we compared the number of paperclips participants in the Asynchrony group collected with the number of paperclips participants in the Synchrony group collected. Since the data were not normally distributed (all $p < .001$), we used a Mann-Whitney U test. The analysis showed no significant difference between groups: $W = 333, p = .91, r = -0.02$ ($BF_{01} = 3.08$), with participants in the Asynchrony group collecting on average 2.58 of 40 paperclips, and participants in the Synchrony group collecting only 2.

Since helping rates were relatively low overall (only 21.15% of participants collected any paperclips at all), we conducted an additional analysis in which helping behavior was coded as binary, and behaviors such as moving paperclips towards the confederate or picking up the container were included as 'helping' as well. This resulted in a higher helping rate overall (30.77% of participants displayed some form of helping behavior), and a slightly higher helping rate in the Synchrony group (34.61% of participants) compared to the Asynchrony group (26.92% of participants), but the difference between groups was not significant: $X^2(1, N = 52) = 0.36, p = .55$ ($BF_{01} = 2.35$).

Questionnaire Responses

As preregistered, we compared participants' liking of their task partner (the confederate) and of experimenter 1, respectively, in each group. Since the data were not normally distributed, we computed two Mann-Whitney U tests; one for the task partner and one for experimenter 1. The analysis of participants' liking of their task partner showed no significant difference between the Synchronous (5.46 ± 1.14) and the Asynchronous condition (5.04 ± 1.18): $W = 269$, $p = .2$, $r = -0.2$ ($BF_{01} = 1.93$). Similarly, participants' liking of the experimenter also did not differ between the Synchronous (6 ± 1.02) and the Asynchronous group (5.5 ± 1.11): $W = 251.5$, $p = .1$, $r = -0.26$ ($BF_{01} = 1.34$). This suggests that our manipulation did not significantly affect participants' affiliation with their task partner or the experimenter.

In order to see whether participants showed a general preference between the task partner and the experimenter, we also ran a Wilcoxon signed rank test to compare their ratings of liking for the confederate and for the experimenter, collapsed across both conditions. The analysis showed that participants gave significantly higher ratings of liking for the experimenter (5.75 ± 1.08) than for the confederate (5.25 ± 1.17): $W = 307.5$, $p < .001$, $r = 0.75$ ($BF_{10} = 107.11$). This suggests that participants felt more affiliated with experimenter 1 than with the confederate in this study.

We also analyzed participants' ratings of the difficulty and of their enjoyment of the coordination task, and of the perceived difficulty of the automatic imitation task in each group. None of these comparisons reached statistical significance (all $p > .26$); detailed results can be found in the supplementary information.

Discussion

In this experiment, we aimed to investigate how intentional synchronization affects self-other overlap. We did so by using a new measure of motor-level self-other overlap, an automatic imitation task adapted from Sowden & Catmur (2015), as well as other measures that have been used in previous studies on the prosocial effects of interpersonal synchrony: perceived self-other overlap (IOS ratings), helping behavior (collecting dropped objects), and liking. We did not find evidence that intentional synchronization with a task partner influenced any of our outcome measures. Reporting these findings remains important, however, in light of recent debates regarding the replication crisis in psychology (see Open Science Collaboration, 2012; Shrout & Rodgers, 2018) and as the robustness of some of the social effects of interpersonal synchrony have been questioned (see Mogan et al., 2017).

The automatic imitation task we used in this experiment differed from our pilot studies in three ways. First, participants saw their task partner's and the experimenter's hand from a first-person perspective in the stimulus images. We reasoned that this change should increase self-other overlap and lead to greater imitative tendencies due to the anatomical congruency between the depicted hand and participants' own hand used to respond to the stimulus images. Second, participants used both their left and their right hand to respond to the stimulus images. This was necessary in order to implement our third change, which was to show only spatially compatible movements in the stimulus images. As our aim was to quantify motor-level self-other overlap, we tried to control for spatial compatibility as a potential confound by keeping this factor constant across all stimulus images. Participants were thus always shown images in which the movement occurred on the side of the screen that was congruent with the prompted response. For instance, when participants were using their right hand to respond and the color cue prompted a middle finger lift, participants were shown images of either a right hand with

a raised middle finger (i.e., the movement was on the right side of the screen) or an image of a left hand with a raised index finger (where the movement was still on the right side of the screen).

By consequence, imitatively compatible trials were also anatomically compatible, as the depicted hand not only lifted the same finger as the participant, but also used the same hand. While this allowed us to control for spatial compatibility, it also introduced a new potential confound: anatomical compatibility. Whereas we kept this factor constant in our three pilot studies, in our final experiment it co-varied with imitative compatibility (or effector compatibility). We reasoned that this potential confound should not affect the interpretation of our results, as the underlying process is still related to a bodily mapping of one hand onto the other, rather than a simple spatial mapping that forgoes any representation of self and other. It is, however, a factor that should be controlled for in future studies.

Our new implementation of the automatic imitation task was successful in eliciting an imitation effect. As we expected, participants were faster in response to compatible trials and made fewer mistakes on compatible trials. However, we found that participants were also faster to respond when using their left, non-dominant hand compared to their right, dominant hand. Similar effects have been documented in the previous literature (see e.g., Johnstone & Carey, 2016). This did not affect the magnitude of the imitation effect, nor was it visible in participants' error rates, so we do not expect it to influence our results.

Participants' error rates were higher in this experiment than they were in our pilot studies, leading to higher exclusion rates, especially in the synchronous condition. This may also be of theoretical significance, as it seems to suggest that our manipulation may have differentially affected participants' error rates. However, our analyses testing for a modulation of imitative tendencies did not show any difference between conditions. This suggests that our

manipulation did not affect participants' motor-level self-other overlap with their task partner or the experimenter. There may be several reasons for this.

First, participants in this experiment had more trouble recognizing their task partner's or the experimenter's hand at test, even after we added a mark to one person's hand and emphasized in the instructions that they should be able to recognize the two hands. It is not clear why this was the case – the models (i.e., the confederate and experimenter 1) were identical to pilot study 3, the primary difference between the stimuli was the change in perspective on the hand. But the egocentric perspective could imply that participants also did not encode the model's identity as they were completing the automatic imitation task, resulting in no difference in their imitative tendencies. It is also possible that the task itself was not engaging enough, leading participants to zone out and not pay attention to their task partner's hand. As it was possible to complete the coordination task by simply focusing on the metronome and keeping the initial tempo, it wasn't strictly necessary for participants to pay attention to their task partner's movements in order to complete the task (though they were instructed to do so).

Although we did not find a significant difference between conditions in our measure of perceived self-other overlap, it is worth noting that participants' IOS ratings for their task partner, on average, increased following synchronous coordination and decreased following asynchronous coordination. This was not the case in their IOS ratings for the experimenter and may suggest that our manipulation had some influence on participants' perceived self-other overlap with their task partner. However, as the comparison was not statistically significant, this interpretation must be treated with caution. Thus, while our IOS results contradict some findings in the literature (e.g., Hove & Risen, 2009), they are in line with others (e.g., Stupacher et al., 2017) and may suggest that simply synchronizing one's movements with another person

is not sufficient for inducing a strong increase in perceived self-other overlap, or interpersonal connectedness. The IOS ratings also remained low overall, as in the pilot studies, suggesting that participants did not feel particularly connected to either target of the ratings.

This experiment also included a helping task, in which the confederate dropped 40 paperclips (similar to the tasks in Kokal et al., 2011 and in Stupacher et al., 2016), and we observed participants' helping behavior. Unlike previous studies, we did not find a greater willingness to help following intentional synchronization, and helping rates overall remained low across both groups. There could be a number of reasons for this – it is possible that the presence of two experimenters led participants to act differently than they might have under different circumstances, or the fact that they had a task to complete (filling out the final questionnaire) may also have led them to reconsider helping the confederate. It is unclear how this finding should best be interpreted, but these results suggest that our coordination task did not lead to an increase in prosocial behavior.

We also included liking as a dependent measure in this experiment. While we did not find a difference between conditions, it is interesting to note that participants felt more affiliated with the experimenter than with their task partner, overall, similar to our results of pilot study 1 (but not pilot studies 2 and 3; see supplementary information). In both of these studies, participants stood or sat side by side, facing the experimenter. It thus appears that simply facing the experimenter, and thereby having access to facial cues, making eye contact, etc., may have been sufficient to lead to stronger feelings of affiliation with her than a task partner. As we do not have baseline ratings of their liking of the task partner and the experimenter, this interpretation remains speculative, as participants may simply have liked the experimenter more than their task partner, or their responses may have been driven by demand characteristics, as they assumed the experimenter would later analyze their responses.

It is worth noting that completing the experiment with a confederate rather than a naïve task partner could have inadvertently created an artificial situation in which participants might have felt that they were being tricked, or that the differences between the coordination achieved between participants and the confederate in each condition were the result of changes in the confederate's behavior rather than the participants'. We aimed to mitigate these concerns by instructing the confederate to act as though she were a naïve participant and to fulfill the task instructions as though she were participating in the experiment for the first time. This appeared to have been successful, as participants did not reveal any suspicions during the debriefing. However, it remains possible that the confederate's behavior showed subtle differences from the behavior of a naïve participant given her familiarity with the task. Moreover, it is possible that the confederate's familiarity with the experiment resulted in reduced engagement in the joint task, and that this may have negatively affected participants' liking of the confederate or their willingness to help.

Furthermore, the nature of intentionality induced by our synchrony manipulation is worth addressing. Following Cross et al. (2021), we aimed to induce a shared goal of generating synchronized movements by instructing participants to move in time with their task partner. Compared to a situation in which interpersonal synchrony is incidentally achieved by simply giving participants the same metronome to initially pace them and instructing them to stick to the pace of the metronome once it stops (as in the incidental synchronization condition in Cross et al., 2021), interpersonal synchrony in the context of our study was achieved intentionally. Put differently, we aimed to induce a shared intention to produce synchronized movements via our instructions by (1) instilling the intention to move at the same pace in both members of the dyad, (2) prompting both participants to plan their actions in accordance with these intentions, and (3) ensuring common knowledge of the instructions between both members of the dyad

(see Bratman, 1993). Interpersonal synchrony in our study, was thus achieved via a process involving shared intentionality.

However, although studies have found the presence of shared intentions to enhance the effects of interpersonal synchrony (Cross et al., 2021; Reddish et al., 2013), it is unclear whether the presence of a shared intention qualitatively shapes people's experience as well. In particular, the presence of a shared intention may not necessarily entail an experience of sharedness or togetherness beyond the context of the task at hand. Moreover, it is unclear whether the level of shared intentionality differed between conditions in our manipulation. In the asynchrony condition, to induce the goal of generating asynchronous movements – of maintaining a constant pace and not synchronizing with their task partner – by instructing participants to stick to their own pace. It is unclear whether participants represented these instructions as pertaining to a shared goal of collectively producing asynchronous movements, or simply in terms of their individual goal to follow the metronome (and the metronome only). In case of the former, it is possible that differences in terms of how closely coordinated participants' movements were alone simply were not sufficient to give rise to prosocial effects of interpersonal synchrony. In case of the latter, this would suggest that even the combination of differences in terms of how closely coordinated participants' movements were and the level of shared intentionality involved in the task was not sufficient to elicit differences in participants' relation to their task partner.

Overall, the results presented in this chapter do not support the hypothesis that intentional synchronization increases self-other overlap – neither at the motor level, nor in terms of perceived self-other overlap. We also did not find evidence for an increase in prosocial tendencies or affiliation following synchronous compared to asynchronous coordination. These findings question the robustness of the prosocial effects of interpersonal synchrony, and

suggest that other factors, such as the presence of a shared goal or music, may have been the driving force eliciting some of the reported effects. Future work should focus on further teasing apart the processes involved in synchronization to allow for a better understanding of coordination and human social behavior.

Supplementary Information

Pilot 1: Data Treatment and Results

Coordination Task

For each participant's data, we computed the difference in their y axis coordinates as well as the difference in time between each saved value in each trial of the coordination task. We analyzed only the part of the trial in which the metronome was no longer present. After removing duplicate entries and rows in which two different y axis coordinates were recorded for the same timestamp (due to technical error), we computed velocity and acceleration for each timestamp of each trial. We then used the `approx()` method in R to interpolate the data and achieve a consistent sampling rate of 50Hz (one entry every 20ms). After removing null values at the beginning and end of each trial and further smoothing the data using the `loess()` method in R, we mean-centered the data and recomputed velocity and acceleration for each trial.

We then proceeded to compute half period normalized continuous phase values for each trial following Varlet and Richardson (2011). To do so, we first computed the half periods for each trial by computing inflection points of the signal and calculating the duration between them. We then computed normalized velocity for each trial and used this to compute individual phase angles for each participant in each trial. In order to obtain the relative phase angles of the two participants in each trial, we subtracted the signal of participant B from that of

participant A for each timepoint when there was a signal from both participants. We then transformed the relative phase angles into vectors using the R package *circular* (Agostinelli & Lund, 2023) and computed mean vectors for each condition in each dyad.

To compare participants' coordination between conditions, we tested for a difference in mean vector length between synchronous and asynchronous coordination blocks. We used mean vector length, as an indicator of the variability in relative phase between dyads, rather than the mean direction of the vector, as we expected participants to be more closely coordinated in synchronous coordination blocks and more spread out in asynchronous coordination blocks. We compared dyads' mean vector length between conditions using a paired-samples Student's *t* test. The results showed that participants were significantly more closely coordinated in synchronous (0.87 ± 0.11) compared to asynchronous coordination blocks (0.04 ± 0.03): $t(7) = 21.8, p < .001, d = 2.44$ ($BF_{10} = 87,881.61$), suggesting that our manipulation was successful.

We conducted an additional exploratory analysis to test how closely participants stuck to the instructed pace. As the metronomes were not correctly recorded due to technical failure, we used participants' data from the phase during which they could still hear the metronome to determine whether they were instructed to move at the slow or fast metronome.²¹ We then compared their data from the remainder of the trial (when they could not hear the metronome) to the instructed pace in two separate one-sample Student's *t* tests (one for each tempo). The results showed that participants moved at a significantly faster pace compared to the instructed tempo, both in the case of the 59bpm (i.e., one beat every 1.017s) metronome ($0.99 \pm 0.01s$): $t(15) = -13.41, p < .001, d = -3.35$, and in the case of the 83bpm (i.e., one beat every 0.723s)

²¹ In one case it was unclear; this trial was excluded from the analysis.

metronome (0.7 ± 0.01 s): $t(15) = -8.57, p < .001, d = -2.14$. This suggests that participants did not successfully keep the instructed tempo.

We followed up on these results in a 2x2 repeated-measures ANOVA with the within-subjects factors Tempo (fast vs. slow) and Condition (synchrony vs. asynchrony) to test whether participants moved at different paces in the fast compared to the slow trials and whether this differed between conditions. The results revealed a significant main effect of Tempo: $F(1,15) = 9840.49, p < .001, \eta_p^2 = 1$, indicating that participants moved faster when they were instructed to follow the fast tempo compared to the slow tempo. The main effect of Condition was not significant: $F(1,15) = 1.4, p = 0.26, \eta_p^2 = 0.09$, suggesting that participants' paces did not significantly differ between conditions. The interaction between Tempo and Condition also was not significant: $F(1,15) = 0.07, p = 0.8, \eta_p^2 = 0.004$. Together, these results indicate that although participants did not succeed in keeping the instructed pace, they did adapt their pace to the instructed tempo in both the synchrony and the asynchrony condition.

Automatic Imitation Task

The RT results revealed a main effect of Effector Compatibility: $F(1,13) = 8.71, p = .011, \eta_p^2 = 0.4$. Participants were faster to respond to compatible (0.52 ± 0.09 s) compared to incompatible trials (0.55 ± 0.08 s). The main effect of Stimulus was not significant: $F(1,13) = 0.7, p = .42, \eta_p^2 = 0.05$. However, there was a significant interaction between Effector Compatibility and Stimulus: $F(1,13) = 41.95, p < .001, \eta_p^2 = 0.76$. Follow-up comparisons revealed that participants who were shown left-hand images responded faster on effector-compatible (0.47 ± 0.05 s) compared to effector-incompatible trials (0.57 ± 0.09 s): $p_{adj} < .001$, whereas participants who were shown right-hand images responded faster to effector-incompatible (0.53 ± 0.08 s) compared to effector-compatible trials (0.57 ± 0.09 s): $p_{adj} = .05$.

There was no significant modulation of interferences scores based on Condition or Model in either Stimulus group (all $p > .38$).

Similarly to the RT results, the ER results showed that participants' ERs overall were higher in response to effector-incompatible ($4.72 \pm 4.2\%$) compared to effector-compatible trials ($1.11 \pm 1.91\%$): $Z = 7, p = .02$. When considering each Stimulus group separately, participants who responded to right-hand images did not show significantly different ERs in response to effector-compatible ($1.79 \pm 2.53\%$) compared to effector-incompatible trials ($1.79 \pm 3.05\%$): $Z = 3, p = 1$. Participants who responded to left-hand images, however, showed significantly higher ERs in response to effector-incompatible ($7.29 \pm 3.34\%$) compared to effector-compatible trials ($0.52 \pm 0.96\%$): $Z = 0, p = .01$. Participants' difference scores in response to their task partner's hand were not modulated by Condition in either Stimulus group (all $p > .07$).

All participants correctly identified the experimenter's or their task partner's hand, and most participants were also confident that their response was correct ($90.43 \pm 24.28\%$). Most of the participants who were asked to rate their belief that the hand they saw as actually their task partner's or the experimenter's hand indicated that they were fairly confident that this was the case (mean confidence rating: 5.63), suggesting that the deception manipulation was effective.

IOS Ratings

The Wilcoxon signed rank test comparing participants' ratings of perceived overlap with their task partner following synchronous ($40.4 \pm 39.8\%$) and asynchronous coordination blocks ($37.1 \pm 37\%$) did not reveal any significant difference between blocks: $Z = 49, p = .85$.

Final Questionnaire

Participants rated their liking of the experimenter higher (6.1875 ± 0.75) than their liking of their task partner (5.13 ± 1.5): $t(15) = 3.17, p = .006, d = 0.79$. Participants did not find it significantly more difficult to coordinate their movements in synchronous (2.56 ± 1.67) compared to asynchronous coordination blocks (3.31 ± 1.78): $t(15) = 1.38, p = .19, d = 0.35$, and they did not enjoy synchronous coordination blocks (4.25 ± 1.81) more than asynchronous coordination blocks (4.88 ± 1.2): $t(15) = 1.27, p = .22, d = 0.32$.

Pilot 2: Data Treatment and Results

Coordination Task

We applied the same procedure used for the treatment and analysis of participants' y axis coordination data in pilot 1, to participants' x axis coordination data in pilot 2. As the data were not normally distributed, we compared dyads' mean vector length between conditions using a Wilcoxon signed-rank test. The results showed that participants were significantly more closely coordinated in synchronous (0.86 ± 0.18) compared to asynchronous coordination blocks (0.04 ± 0.03): $Z = 36, p = .008, r = 0.38$ ($BF_{10} = 21.45$), suggesting that our manipulation was successful.

As for pilot study 1, we conducted an additional exploratory analysis to test how closely participants stuck to the instructed pace. The metronomes were not correctly recorded in pilot study 2 either due to technical failure, so we used participants' data from the phase during which they could still hear the metronome to determine whether they were instructed to move at the slow or fast metronome.²² We compared participants' data from the remainder of the trial (when they could not hear the metronome) to the instructed pace in two separate one-sample

²² For one dyad, it was unclear (especially in the asynchrony condition). This dyad was excluded from the analysis.

Student's *t* tests (one for each tempo). The results showed that participants moved at a significantly faster pace compared to the instructed tempo in the case of the 59bpm (i.e., one beat every 1.017s) metronome ($0.98 \pm 0.06s$): $t(13) = -2.25, p = .04, d = -0.6$ ($BF_{10} = 1.77$), but not in the case of the 83bpm (i.e., one beat every 0.723s) metronome ($0.72 \pm 0.04s$): $t(13) = -0.48, p = .64, d = -0.13$ ($BF_{01} = 3.35$). This suggests that participants did not succeed in keeping the instructed slow tempo but were more successful at keeping the faster tempo.

As for pilot study 1, we followed up on these results in a 2x2 repeated-measures ANOVA with the within-subjects factors Tempo (fast vs. slow) and Condition (synchrony vs. asynchrony) to test whether participants moved at different paces in the fast compared to the slow trials and whether this differed between conditions. The results revealed a significant main effect of Tempo: $F(1,13) = 616.63, p < .001, \eta_p^2 = 0.98$, indicating that participants moved faster when they were instructed to follow the fast tempo compared to the slow tempo. The main effect of Condition was also significant: $F(1,13) = 22.67, p < .001, \eta_p^2 = 0.64$, suggesting that participants' paces were faster in the synchrony ($0.83 \pm 0.19s$) compared to the asynchrony condition ($0.87 \pm 0.2s$). The interaction between Tempo and Condition was not significant: $F(1,13) = 0.22, p = .65, \eta_p^2 = 0.02$. Together, these results indicate that participants adapted their pace to the instructed tempo, albeit more accurately in the case of the faster tempo, and that their tempo was affected by our manipulation such that participants moved faster in synchronous compared to asynchronous coordination blocks.

Automatic Imitation Task

The RT results revealed no main effect of Effector Compatibility: $F(1,13) = 0.03, p = .87, \eta_p^2 = 0.002$. The main effect of Stimulus also was not significant: $F(1,13) = 0.23, p = .23, \eta_p^2 = 0.11$. However, there was a significant interaction between Effector Compatibility and Stimulus: $F(1,13) = 22.73, p < .001, \eta_p^2 = 0.64$. Follow-up comparisons revealed that

participants who were shown left-hand images responded faster on effector-compatible (0.5 ± 0.1 s) compared to effector-incompatible trials (0.53 ± 0.08 s): $p_{adj} = .01$, whereas participants who were shown right-hand images responded faster to effector-incompatible (0.45 ± 0.07 s) compared to effector-compatible trials (0.48 ± 0.08 s): $p_{adj} = .01$. There was no significant modulation of interferences scores based on Condition or Model in either Stimulus group (all $p > .09$).

Similarly to the RT results, the ER results showed no significant difference in participants' ERs overall between effector-incompatible ($3.06 \pm 3.51\%$) and effector-compatible trials ($2.78 \pm 3.02\%$): $Z = 44$, $p = .63$. When considering each Stimulus group separately, however, participants who responded to right-hand images showed higher ERs in response to effector-compatible ($3.12 \pm 3.34\%$) compared to effector-incompatible trials ($0.52 \pm 0.96\%$), but this difference was not significant: $Z = 25$, $p = .07$. Participants who responded to left-hand images showed higher ERs in response to effector-incompatible ($5.95 \pm 3.05\%$) compared to effector-compatible trials ($2.38 \pm 2.8\%$), but this difference also was not significant: $Z = 3$, $p = .07$. Participants' difference scores in response to their task partner's hand were not modulated by Condition in either Stimulus group (all $p > .21$).

All participants correctly identified the experimenter's or their task partner's hand, and most participants were also confident that their response was correct ($85.34 \pm 21.35\%$). Most of the participants indicated that they were fairly confident that the hand they saw as actually their task partner's or the experimenter's hand (mean confidence rating: 5.38), suggesting that the deception manipulation was effective.

IOS Ratings

Participants' IOS ratings for their task partner did not differ between synchronous ($24.7 \pm 33.8\%$) and asynchronous coordination blocks ($24.5 \pm 33.8\%$): $Z = 44$, $p = .94$. The difference

between the baseline ratings of overlap with their task partner and their ratings after synchronous (+2 ± 33.4%) compared to asynchronous coordination (+2.19 ± 33.9%) also did not significantly differ: $Z = 32.5, p = 1$.

Final Questionnaire

Participants did not rate their liking of the experimenter significantly higher (5.75 ± 1.53) than their liking of their task partner (5.13 ± 1.26): $t(15) = 1.78, p = .1, d = 0.44$. Participants did not find it significantly more difficult to coordinate their movements in synchronous (3 ± 1.46) compared to asynchronous coordination blocks (1.56 ± 1.26): $t(15) = -0.75, p = .47, d = -0.19$, and they did not enjoy synchronous coordination blocks (4.19 ± 1.8) more than asynchronous coordination blocks (4.56 ± 1.71): $t(15) = 0.67, p = .51, d = 0.17$.

Pilot 3: Data Treatment and Results

Coordination Task

To perform a manipulation check in pilot study 3, we computed mean asynchronies for each participant and compared these between synchronous and asynchronous coordination blocks. Mean asynchronies were computed by calculating the difference in time between the onset of each of the confederate's taps and the onset of the participant's tap that was closest in time to it (but not more than half an interval of the metronome the participant heard at the beginning of the trial before or after the confederate's tap). We then computed the mean asynchrony for each trial, and from these means a mean asynchrony for synchronous and asynchronous coordination blocks, respectively.

The paired Student's t test comparing mean asynchronies in synchronous and asynchronous coordination blocks showed a significant difference between blocks: $t(7) = 32.38, p < .001, d = 11.45$. Participants' taps were more closely coordinated with the

confederate's taps in synchronous (83 ± 11 ms difference) than in asynchronous coordination blocks (214 ± 3 ms difference), suggesting that the manipulation was successful in inducing more closely coordinated movements in synchronous compared to asynchronous blocks.

As in pilot studies 1 and 2, we compared participants' pace to the instructed tempo in a one-sample Student's *t* test for each metronome. The analysis showed that participants moved at a significantly faster pace compared to the instructed tempo, both in the case of the 59bpm (i.e., one beat every 1.017s) metronome (0.98 ± 0.04 s): $t(7) = 73.7, p < .001, d = 26.06$, and in the case of the 83bpm (i.e., one beat every 0.723s) metronome (0.72 ± 0.02 s): $t(7) = 102.82, p < .001, d = 36.35$. This suggests that participants did not succeed in keeping the instructed tempo in the case of both the fast and the slow metronome, moving faster than the instructed tempo.

As in pilots 1 and 2, we followed up on these results in a 2x2 repeated-measures ANOVA with the within-subjects factors Tempo (fast vs. slow) and Condition (synchrony vs. asynchrony) to test whether participants moved at different paces in the fast compared to the slow trials and whether this differed between conditions. The results showed a significant main effect of Tempo: $F(1,7) = 1.593, 690.76, p < .001, \eta_p^2 = 0.99$, indicating that participants adapted their pace to the instructed tempo. The main effect of Condition was also significant: $F(1,7) = 9.85, p = .02, \eta_p^2 = 0.59$, suggesting that participants' paces were faster in the synchrony (0.79 ± 0.15 s) compared to the asynchrony condition (0.87 ± 0.18 s). The interaction between Tempo and Condition was not significant: $F(1,7) = 1.39, p = .28, \eta_p^2 = 0.17$. Together, these results indicate that participants adapted their pace to the instructed tempo and that their tempo was affected by our manipulation such that participants moved faster in synchronous compared to asynchronous coordination blocks.

Automatic Imitation Task

The Wilcoxon signed rank test comparing RTs on effector-compatible trials (0.54 ± 0.05 s) to RTs on effector-incompatible trials (0.51 ± 0.07 s) did not reveal a significant difference: $Z = 24, p = .11$. There was no significant modulation of interferences scores based on Condition or Model (all $p > .61$). The ER results similarly showed no significant difference in participants' ERs in response to effector-compatible ($7.44 \pm 6.35\%$) and effector-incompatible trials ($3.27 \pm 3.37\%$): $Z = 23, p = .15$, and participants' difference scores in response to their task partner's hand did not differ between synchronous and asynchronous coordination blocks: $Z = 4, p = .41$. Seven participants (87.5%) correctly identified the experimenter's or their task partner's hand, and most participants were confident that their response was correct (mean confidence rating: $71.42 \pm 33.02\%$).

IOS Ratings

There was no significant difference in participants' IOS ratings for their task partner between synchronous ($22.8 \pm 29\%$) and asynchronous coordination blocks ($21 \pm 29.7\%$): $Z = 4, p = .42$. The 2x2 repeated-measures ANOVA on the deviation scores (i.e., the difference between baseline IOS ratings and participants' IOS ratings after each block of the coordination task) with factors Condition (synchrony vs. asynchrony) and Other (experimenter vs. partner) did not reveal any significant differences (all $p > .34$).

Final Questionnaire

Participants did not rate their liking of the experimenter significantly higher (5.5 ± 0.93) than their liking of their task partner (5.13 ± 1.13): $t(7) = 2.05, p = .08, d = 0.73$. Participants did not find it significantly more difficult to coordinate their movements in synchronous (3.13 ± 1.89) compared to asynchronous coordination blocks (3.63 ± 1.41): $t(7) = 0.64, p = .54, d =$

0.23, and they did not enjoy synchronous coordination blocks (4.75 ± 1.39) more than asynchronous coordination blocks (3.88 ± 1.25): $t(7) = -1.37, p = .21, d = -0.48$.

Pilots 1-3: Supplementary Tables

Final Questionnaire

Supplementary Table 1

Difficulty Ratings of the Coordination Task

	Difficulty of the Task (M ± SD)	Difficulty in Synchronous Blocks (M ± SD)	Difficulty in Asynchronous Blocks (M ± SD)
Pilot 1 (n = 16)	3.19 ± 1.52	2.56 ± 1.67	3.31 ± 1.78
Pilot 2 (n = 16)	2.56 ± 1.21	3 ± 1.46	2.56 ± 1.26
Pilot 3 (n = 8)	3.25 ± 1.58	3.13 ± 1.89	3.63 ± 1.41
	2.95 ± 1.41	2.85 ± 1.61	3.08 ± 1.54

Note. Participants rated the difficulty of the coordination task overall, of synchronous coordination blocks, and of asynchronous coordination blocks on a 7-point Likert scale ranging from “very easy“ (1) to “very difficult” (7). Their ratings did not differ between pilots or conditions.

Supplementary Table 2

Enjoyment Ratings of the Coordination Task

	Enjoyment of the Task (M ± SD)	Enjoyment in Synchronous Blocks (M ± SD)	Enjoyment in Asynchronous Blocks (M ± SD)
Pilot 1 (n = 16)	4.81 ± 1.28	4.25 ± 1.81	4.88 ± 1.2
Pilot 2 (n = 16)	4 ± 1.75	4.19 ± 1.8	4.56 ± 1.71
Pilot 3 (n = 8)	4.5 ± 1.41	4.75 ± 1.39	3.88 ± 1.25
	4.43 ± 1.52	4.33 ± 1.7	4.55 ± 1.45

Note. Participants rated their enjoyment of the coordination task overall, of synchronous coordination blocks, and of asynchronous coordination blocks on a 7-point Likert scale ranging from “not at all” (1) to “very much” (7). Their ratings did not differ between pilots or conditions.

Supplementary Table 3

Difficulty Ratings of the Automatic Imitation Task

	Difficulty of the Task (M ± SD)	Difficulty Ignoring Finger (M ± SD)	Difficulty Ignoring E's Finger (M ± SD)	Difficulty Ignoring P's Finger (M ± SD)
Pilot 1	3 ± 1.63	3.13 ± 1.82	2.63 ± 1.59	2.75 ± 1.7
Pilot 2	2.88 ± 1.63	2.94 ± 1.44	2.75 ± 1.44	2.94 ± 1.48
Pilot 3	3.5 ± 1.63	3.63 ± 1.19	3.5 ± 1.31	3.75 ± 1.16
	3.05 ± 1.57	3.15 ± 1.55	2.85 ± 1.48	3.03 ± 1.53

Note. Participants rated the difficulty of the automatic imitation task overall, as well as how difficult it was to ignore the depicted finger movement overall, when they were responding to the experimenter's hand, and when they were responding to the partner's hand. Ratings were performed on a 7-point Likert scale ranging from "very easy" (1) to "very difficult" (7) and did not differ between pilots, conditions, or models (experimenter vs. partner).

Experiment 1: Methods and Results

Automatic Imitation Task Stimuli

During the practice phase, all participants saw the same stimulus images (see Supplementary Figure 1). The images were of experimenter 2's hand (participants were not informed about this and we did not expect this to influence their responses in the task). Then the first 26 participants in our sample were shown the images of the confederate's and experimenter 1's hand depicted in Supplementary Figure 2. However, as recognition rates were at chance and participants reported having difficulty identifying the images, we slightly modified the stimuli for the final 26 participants, such that these participants were shown the same images but with the addition of a mark on either experimenter 1's (see Supplementary Figure 3A) or the confederate's hand (see Supplementary Figure 3B). This modification was further supported by an added verbal emphasis in the instructions (before the first experimental block) on the fact that participants would be responding to the images that were taken at the beginning of the session in the experimental blocks (in the automatic imitation task), and that they should be able to recognize the two hands.

Supplementary Figure 1

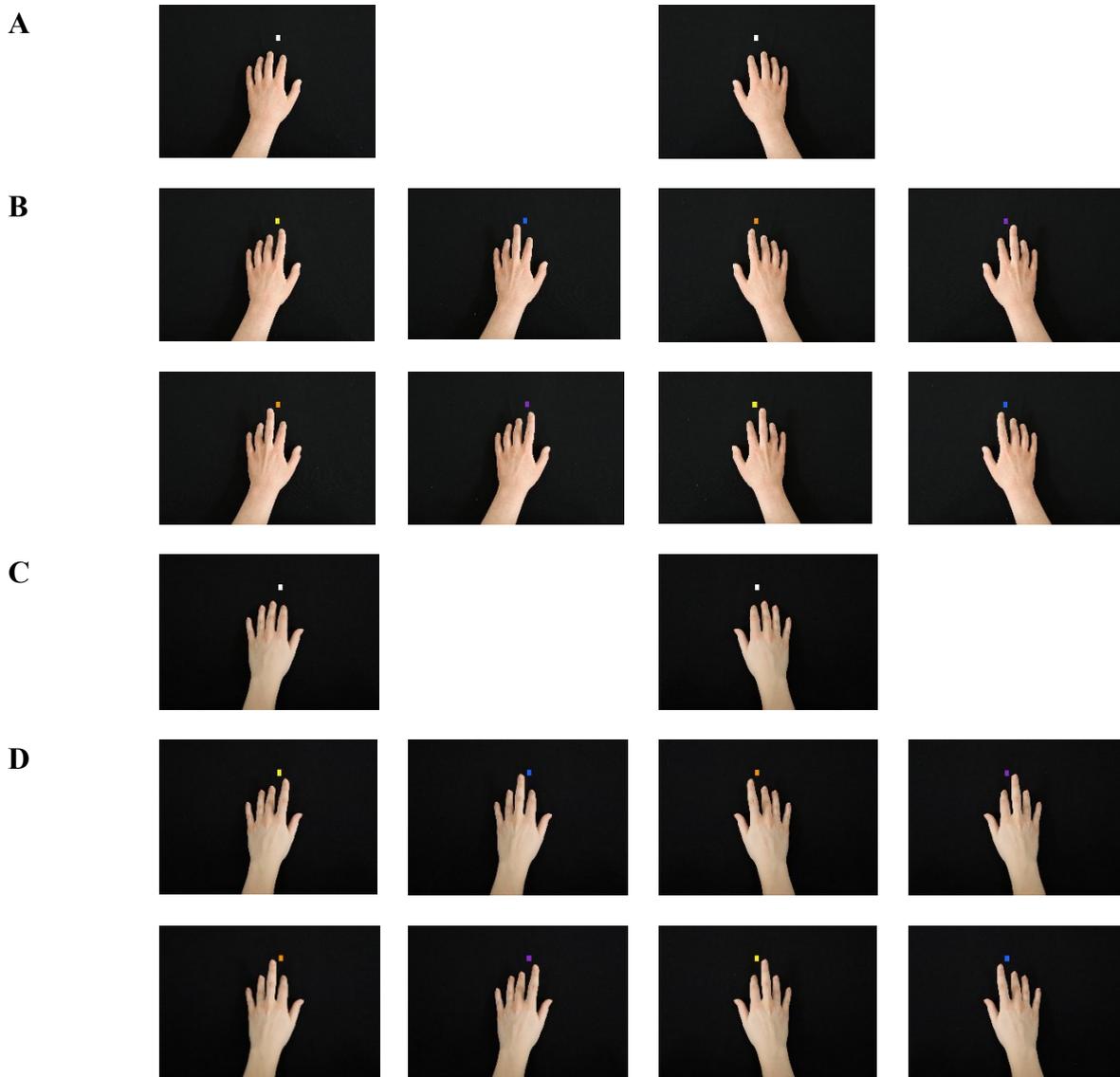
Automatic Imitation Task: Stimuli – Practice Phase



Note. All practice trials started with an image of either a left hand (two columns on the left) or of a right hand (two rightmost columns) in resting position, displayed together with a white rectangle (A). In the first round of practice trials, this image was replaced by an image of the same hand, still in resting position, together with a colored rectangle (B). During this round, participants learned the correct cue-response mapping. In the second round of practice trials, participants were shown images of a hand with a raised index or middle finger, together with a colored rectangle (C). The depicted hand movements were either compatible (top row in C) or incompatible (bottom row in C) with the cued response.

Supplementary Figure 2

Automatic Imitation Task: Stimuli – Experimental Blocks

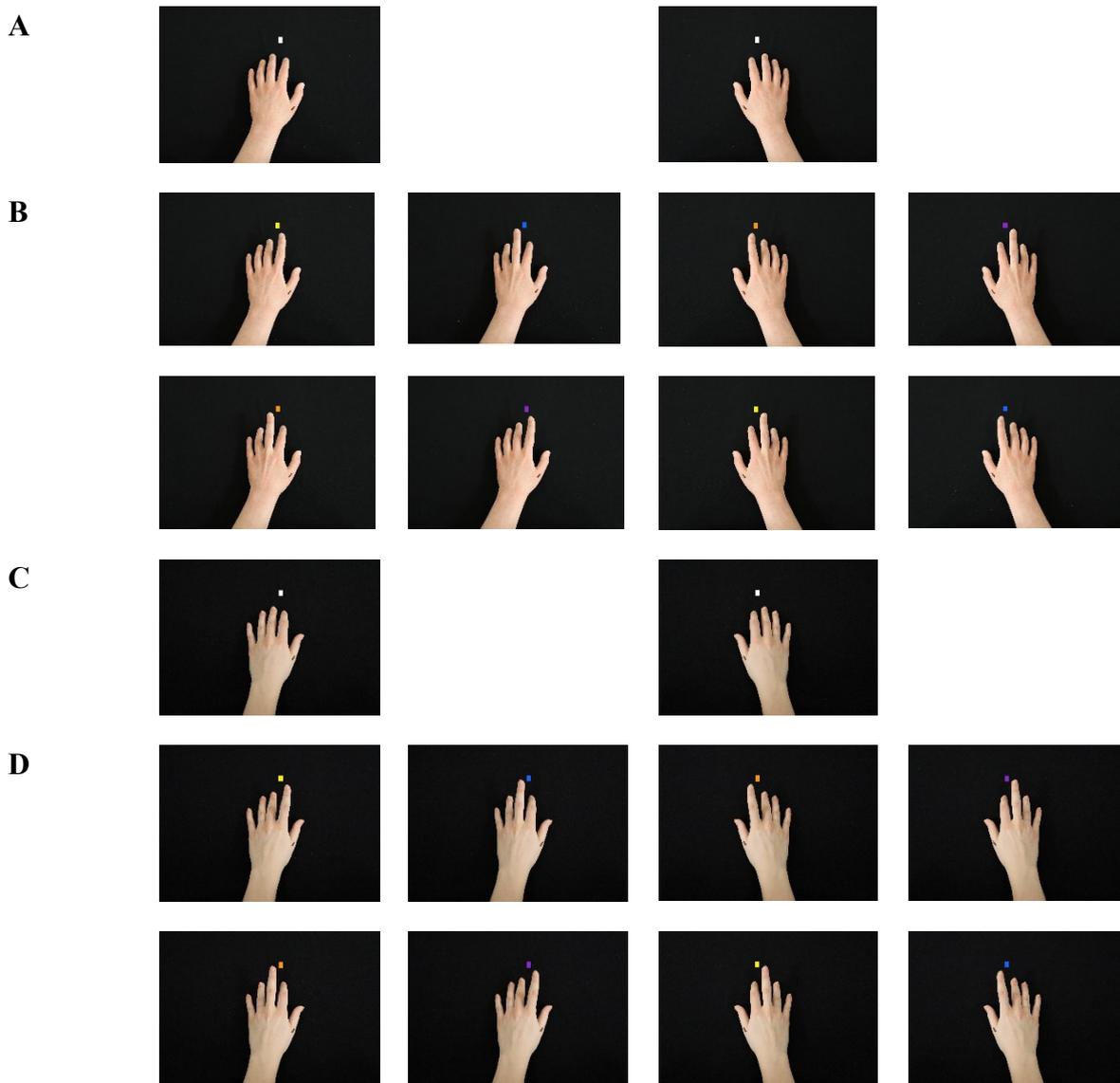


Note. Experimental trials started with an image of either a left hand (two columns on the left) or of a right hand (two rightmost columns) in resting position, displayed together with a white rectangle. In half of the blocks, the images showed experimenter 1's hand (A); in the other half, the images were of the confederate's hand (C). In each trial, the image of a hand in resting position was replaced by an image depicting a hand with a raised index or middle finger, together with a colored rectangle (B for the experimenter's hand; D for the confederate's). The

depicted hand movements were either compatible (top row in B and D) or incompatible (bottom row in B and D) with the cued response.

Supplementary Figure 3

Automatic Imitation Task: Stimuli – Experimental Blocks (Marked)



Note. The last 26 participants were shown stimuli that included a mark on either the experimenter's (A-B) or the confederate's (C-D) hand. For half of these participants, experimenter 1 drew a mark on her hand before the start of the experimental session – these participants saw the stimuli depicted in A and B, paired with the images of the confederate's

hand shown in Supplementary Figure 2. For the other half of these participants, the confederate drew a mark on her hand before the start of the experimental session – these participants saw the stimuli depicted in C and D, paired with images of experimenter 1's hand as shown in Supplementary Figure 2.

Automatic Imitation Data

As the data were not normally distributed, we compared participants' mean RTs in each block using a Friedman test with the four-level within-subject factor Block. The analysis showed that there was a difference in participants' RTs between blocks: $X^2(3) = 17.92, p < .001$ ($BF_{10} = 5.31$). Conover's post hoc comparisons revealed a significant difference in participants' mean RTs between block 1 ($0.49 \pm 0.06s$) and block 2 ($.48 \pm 0.05s$): $p_{adj} = .04$, between block 1 and block 3 ($0.47 \pm 0.06s$): $p_{adj} < .001$, and between block 1 and block 4 ($0.47 \pm 0.04s$): $p_{adj} = .04$. The other contrasts were not significant (all $p_{adj} > .49$). We additionally compared participants' interference scores in each block using a repeated-measures one-way ANOVA with the four-level within-subjects factor Block. The analysis showed that there was no significant difference between participants' interference scores between blocks: $F(3,129) = 1.21, p = .31, \eta_p^2 = 0.03$ ($BF_{01} = 7.21$). Together, these analyses suggest that while participants' reaction times were slower in the first block of the automatic imitation task compared to subsequent blocks of the task, their interference scores did not differ between blocks.

We performed the same analyses on the ER data. The Friedman test showed no significant difference between participants' ERs overall between blocks: $X^2(3) = 1.01, p = .8$ ($BF_{01} = 16.68$). As the difference scores were not normally distributed, we computed a Friedman test with the four-level within-subject factor Block to compare participants' difference scores between blocks. The analysis showed no significant difference in participants' difference scores between blocks: $X^2(3) = 3.81, p = .28$ ($BF_{01} = 10.33$). Together

with the outcome of the RT analysis, these results indicate that while participants' responses in this task became faster overall after the first block of the task, their error rates and imitative tendencies did not differ between blocks.

Questionnaire Responses

Participants' ratings of the difficulty of the coordination task did not significantly differ between groups: $W = 365.5$, $p = .61$, $r = 0.08$ ($BF_{01} = 3.18$), with participants in the Synchrony condition rating it as slightly easier (2.81 ± 1.47) than participants in the Asynchrony condition (3.04 ± 1.54). The ratings of enjoyment of the coordination task were not significantly higher in the Synchrony group (4.31 ± 1.83) compared to the Asynchrony group (4.23 ± 1.48): $W = 336$, $p = .98$, $r = -0.01$ ($BF_{01} = 3.55$). This suggests that participants in both groups found the coordination task to be similarly difficult and enjoyable.

Participants rated the overall difficulty of the automatic imitation task similarly in both conditions: $W = 306.5$, $p = .56$, $r = -0.09$ ($BF_{01} = 3.09$), with participants in the Synchrony condition rating it as slightly more difficult (3.54 ± 1.63) than participants in the Asynchrony condition (3.27 ± 1.43). Ratings of how difficult it was to ignore the depicted finger movement were also higher in the Synchrony group (3.54 ± 1.61) than in the Asynchrony group (3 ± 1.52), but this difference was not significant: $W = 277$, $p = .26$, $r = -0.18$ ($BF_{01} = 1.99$). These results demonstrate that participants in both groups perceived the automatic imitation task to be similarly difficult, both in general and in terms of participants' perceived difficulty of ignoring the depicted finger movement.

When asked how difficult it was to ignore the depicted finger movement when participants were responding to images of their task partner's hand, participants did not provide significantly different ratings for their task partner's hand in the Synchronous (2.92 ± 1.32)

compared the Asynchronous condition (3.08 ± 1.5): $W = 346$, $p = .69$, $r = 0.07$ ($BF_{01} = 3.07$).²³ Similarly, participants' ratings of how difficult it was to ignore the finger movements when they were responding to images of the experimenter's hand did not significantly differ between the Synchronous (3.08 ± 1.41) and the Asynchronous group (3.08 ± 1.41): $W = 321.5$, $p = .87$, $r = 0.03$ ($BF_{01} = 3.52$).²⁴

Relation between Experience with Music and Dance and Coordination Performance

To explore the relationship between participants' experience with music and dance and their degree of coordination with the confederate, we aggregated participants' ratings of their experience with music and their ratings of their experience with dance to form a collapsed score ranging from 2 (no experience with music or dance) to 14 (very much experience with music and dance). These collapsed scores were correlated with participants' absolute asynchronies with their task partner during the coordination task. We conducted separate analyses for the synchrony and the asynchrony condition, respectively.

The Pearson correlation between participants' collapsed ratings of their music and dance experience and their mean absolute asynchrony during the coordination task for participants in the synchrony group revealed a significant positive correlation between participants' experience with music and dance and their mean asynchrony during the coordination task: $r = 0.41$, $p = .04$ ($BF_{10} = 1.87$; see Supplementary Figure 4, left). We followed up on these results and tested for a correlation between participants' mean asynchronies and their ratings of their experience with music and dance, respectively, to see

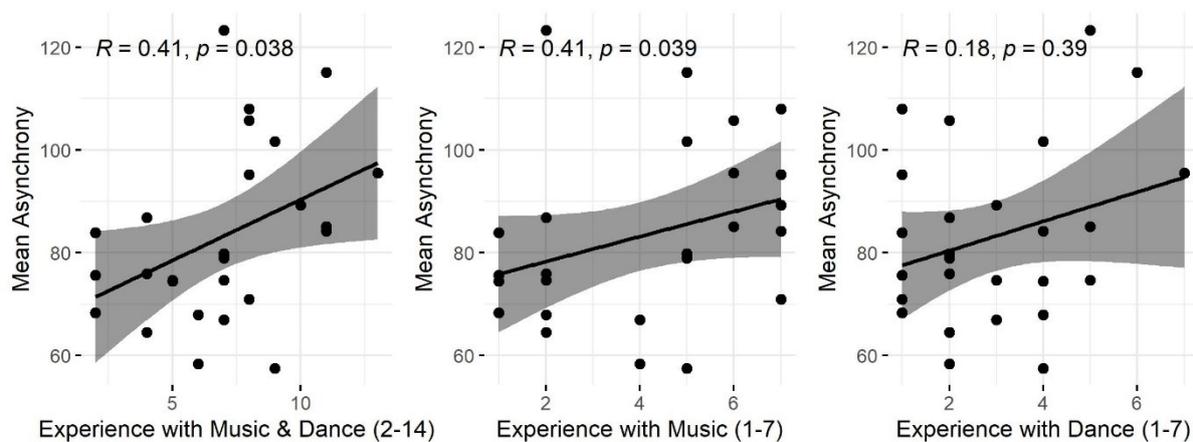
²³ One participant was excluded from this analysis because they did not provide ratings for this questionnaire item, as they did not understand that their task partner's and the experimenter's hand were depicted on the images in the automatic imitation task.

²⁴ Two participants were excluded from this analysis because they did not provide ratings for this questionnaire item, as one did not understand that their task partner's and the experimenter's hand were depicted on the images in the automatic imitation task and one reported that they could not distinguish the experimenter's hand from their own hand.

whether this effect was driven by one of the two items in particular. As the data for both of these questionnaire items were not normally distributed ($p \leq .02$), we did so using Spearman correlations.²⁵ The analyses showed a significant positive correlation between participants' mean asynchronies and their ratings of their experience with music: $\rho = 0.41$, $p = .04$ ($BF_{10} = 1.35$), but the correlation between participants' mean asynchronies and their ratings of their experience with dance was not significant: $\rho = 0.18$, $p = .39$ ($BF_{01} = 2.83$; see Supplementary Figure 4). This suggests that the more experience participants had with music, in particular, the less successful they were in coordinating their movements with the confederate and the higher their mean asynchronies were in the synchrony condition of the coordination task.

Supplementary Figure 4

Relation between Music and Dance Experience and Synchronous Tapping



Note. This figure depicts the relationship between participants' experience with music and dance (x axis) and their mean asynchrony during the coordination task (y axis) in the synchrony condition. The leftmost plot shows the relationship between participants' collapsed ratings of their experience with music and dance and their mean asynchronies in the coordination task.

²⁵ Whereas reported Bayes factors for Pearson correlations refer to their Bayesian equivalent, as implemented in JASP, the Bayes factors reported for Spearman correlations reflect the Bayesian equivalent of Kendall's tau.

The other two plots depict the relationship between participants' experience with music (center plot) and dance (right plot) and their mean asynchronies in the coordination task. The black line in each plot shows the regression line, the shading around it represents the confidence interval, and the black dots depict individual participants' data. The correlation coefficient and p-value are included in the top left corner of each plot.

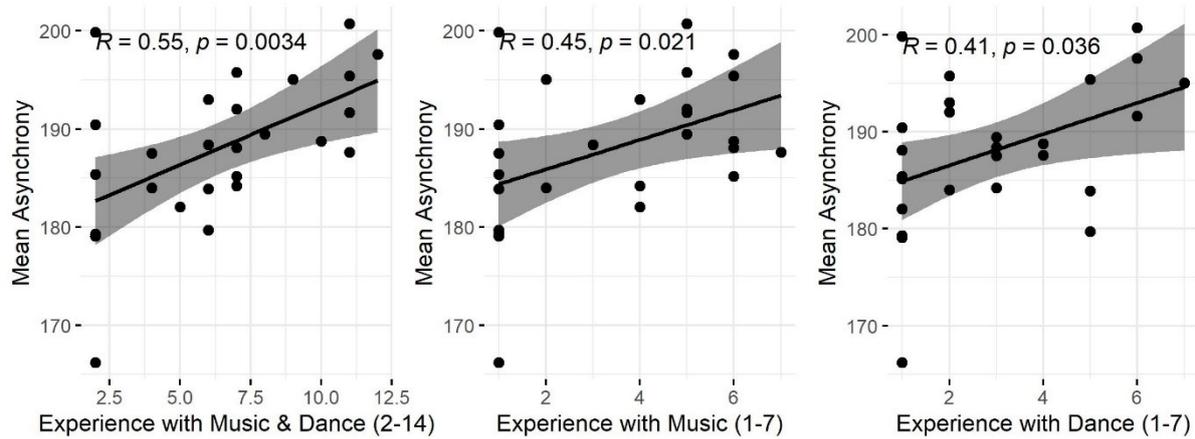
As participants' ratings of their experience with music and dance were not normally distributed in the asynchrony group ($p = .03$), we computed a Spearman correlation to test for a relationship between participants' experience with music and dance and how well participants succeeded in tapping asynchronously with the confederate in this group. The analysis showed a significant positive correlation between participants' experience with music and dance and their mean asynchrony during the coordination task: $\rho = 0.55$, $p = .003$ ($BF_{10} = 30.17$), suggesting that the more experience participants had with music and dance, the more successful they were in coordinating their movements with the confederate in the asynchrony condition of the coordination task (see Supplementary Figure 5, left).

We followed up on these results and tested for a correlation between participants' mean asynchronies and their ratings of their experience with music and dance, respectively, to see whether this effect was driven by one of the two items in particular. As the data for both of these questionnaire items were not normally distributed ($p < .003$), we did so using Spearman correlations. The analyses showed a significant positive correlation between participants' mean asynchronies and their ratings of their experience with music: $\rho = 0.45$, $p = .02$ ($BF_{10} = 3.3$), and between participants' mean asynchronies and their ratings of their experience with dance: $\rho = 0.41$, $p = .04$ ($BF_{10} = 2.43$; see Supplementary Figure 5). This suggests that the outcome of the collapsed analysis was not driven by either of these items in particular and that participants

in the asynchrony condition showed higher asynchronies both with increased experience with music and with increased experience with dance.

Supplementary Figure 5

Relation between Music and Dance Experience and Asynchronous Tapping



Note. This figure depicts the relationship between participants' experience with music and dance (x axis) and their mean asynchrony during the coordination task (y axis) in the asynchrony condition. The leftmost plot shows the relationship between participants' collapsed ratings of their experience with music and dance and their mean asynchronies in the coordination task. The other two plots depict the relationship between participants' experience with music (center plot) and dance (right plot) and their mean asynchronies in the coordination task. The black line in each plot shows the regression line, the shading around it represents the confidence interval, and the black dots depict individual participants' data. The correlation coefficient and p-value are included in the top left corner of each plot.

Relation between Degree of Coordination and Overlap with Task Partner

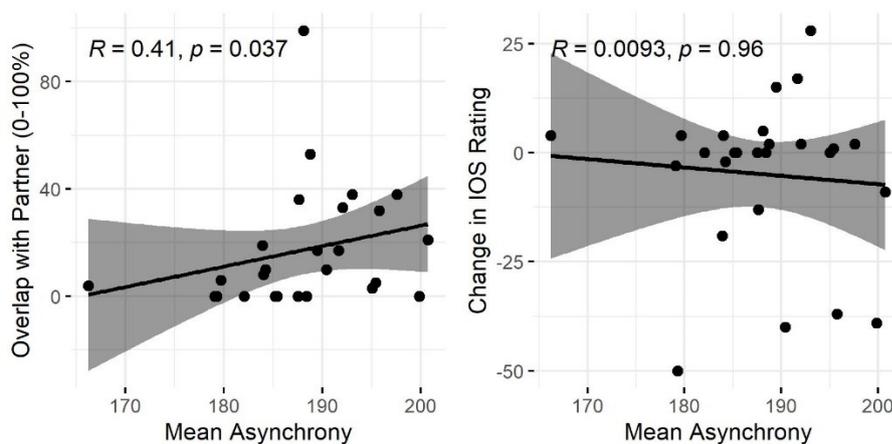
To explore the relationship between participants' degree of coordination with the confederate and participants' ratings of overlap with her, we also computed correlation

analyses to test for a relationship between participants' mean asynchronies during the coordination task and their subsequent IOS ratings for their task partner. We conducted separate analyses for the synchrony and the asynchrony condition, respectively.

As the data were not normally distributed (all $p \leq .007$), we computed Spearman correlations to test for a relationship between participants' mean asynchronies during the coordination task and their IOS ratings for their task partner. For participants in the synchrony condition, the correlation was not significant: $\rho = -0.11$, $p = .6$ ($BF_{01} = 3.6$). However, for participants in the asynchrony condition, the analysis revealed a significant positive correlation between participants' mean asynchronies during the coordination task and their final IOS ratings for their task partner: $\rho = 0.41$, $p = .04$ ($BF_{10} = 1.84$; see Supplementary Figure 6, right). This suggests that, for participants in the asynchrony group, the higher participants' mean asynchrony with the confederate was during the coordination task, the higher the perceived self-other overlap with her was at the end of the session.

Supplementary Figure 6

Relation between Asynchronous Tapping and Perceived Overlap with the Task Partner



Note. This figure depicts the relationship between participants' mean asynchronies during the coordination task (x axis) and their IOS ratings (y axis, left plot) or the change in their IOS

ratings (y axis, right plot) in the asynchrony group. The plots show the regression line, the confidence interval (shaded area), and individual participants' data (black dots). The correlation coefficient and p-value are included in the top left corner of each plot.

We also computed the same analyses on participants' deviation scores to see whether the degree of coordination with the confederate was related to the magnitude of their change in perceived overlap in each group. As the deviation scores also were not normally distributed ($p \leq .02$), we computed Spearman correlations to test for a relationship between participants' mean asynchronies during the coordination task and their deviation scores (i.e., the change in their IOS ratings for their task partner between the baseline IOS rating at the start of the experimental session and the final IOS rating after the 5th block of the coordination task). The analyses did not show a significant correlation between participants' mean asynchronies and their deviation scores in the synchrony group: $\rho = -0.26$, $p = .2$ ($BF_{01} = 1.7$), or in the asynchrony group: $\rho = 0.01$, $p = .96$ ($BF_{01} = 3.83$; see Supplementary Figure 6, left), suggesting that the degree of coordination with the confederate did not significantly affect their perceived self-other overlap with her.

CHAPTER THREE: DOES INTERPERSONAL SYNCHRONY INCREASE SELF-OTHER OVERLAP?

The present chapter continues the line of questioning opened up in Chapter 2 by asking whether moving in synchrony with another person increases self-other overlap and additionally investigating the role of auditory feedback. We addressed this question regarding both motor-level and perceived self-other overlap in two experiments, while also including a novel measure of interpersonal connectedness during interpersonal synchrony. This allowed us to additionally explore whether moving in synchrony with another person has more immediate, short-lived effects on people's relationships during the interaction, which might not stand the test of time.

According to the model put forward by Hove (2008), moving in synchrony with another person should increase self-other overlap via two processes. First, the matching of the perceived sensory consequences of another person's actions with the predicted sensory consequences of one's own actions is thought to reduce one's ability to distinguish between self and other, and thereby increase self-other overlap. Second, the experienced priming of one's own motor system when observing another person's actions, combined with a lack of motor inhibition of the primed action, is thought to further blur the boundary between self and other, and thereby increase self-other overlap. Similar conclusions can also be reached based on the framework proposed by ideomotor theory (Greenwald, 1970), the theory of common coding for perception and action (Prinz, 1990), or the theory of event coding (Hommel et al., 2001), which all assume that there are shared representation for perceptions and actions. Thus, seeing another person act will activate the motor representations required to perform the action in the observer – which, if not inhibited, may lead to the performance of the observed action.

While much of the literature on interpersonal synchrony has focused on contrasting it with asynchronous movement, there is also evidence that moving in synchrony with another person specifically increases people's prosocial behavior compared to no movement at all (Mogan et al., 2017). Moreover, in-phase synchrony has also been shown to boost memory compared to anti-phase synchrony and solitary movement (Macrae et al., 2008), and it has more recently been linked to greater perceived self-other overlap compared to uncoordinated movement and no movement (Lang et al., 2017; Biswas & Brass, 2024). Together, these studies suggest that the prosocial effects of interpersonal synchrony are brought about specifically by an increase in self-other overlap.

In contrast to the literature on interpersonal synchrony, the results on effects of asynchronous coordination are somewhat more mixed. While asynchronous movement has been found to lead to more prosocial behavior than interpersonal synchrony overall (Mogan et al., 2017), its relation to different types of control conditions is less clear. Some studies have found evidence for a difference between prosocial behavior following interpersonal synchrony compared to asynchronous movement, but not between prosocial behavior following asynchronous movement compared to no interaction at all (e.g., Lang et al., 2017). Others did not find a difference between social closeness following asynchronous movement and interpersonal synchrony (Tarr et al., 2016), despite finding a difference between social closeness following partially compared to fully synchronized movements, suggesting that the operationalization of asynchrony may also play an important role. The latter study also found a significant decrease in people's pain thresholds following partial synchrony, whereas pain thresholds significantly increased following interpersonal synchrony, but only documented a nonsignificant decrease in pain thresholds following asynchronous movement. Together, this suggests that the social effects of interpersonal asynchrony are not clear-cut.

However, studies in the domain of automatic imitation and the representation of self and other suggest that moving asynchronously with another person may act as a kind of training in self-other distinction, and thereby reduce self-other overlap. In particular, Santiesteban et al. (2012) showed that imitation inhibition training can enhance visual perspective taking in adults. Similarly, Yue et al. (2024) found that imitation inhibition training reduced the extent to which people attributed other people's actions to themselves in the context of observer-inflation effects. This suggests that if moving asynchronously with another person similarly increases self-other distinction, then effects of interpersonal synchrony that were shown in contrast with asynchronous movement may additionally have been enhanced by a decrease in self-other overlap following interpersonal asynchrony.

In this study, we tested how moving (a-)synchronously with another person influences self-other overlap, both in terms of motor-level self-other overlap, using an automatic imitation task (Sowden & Catmur, 2015), and in terms of perceived self-other overlap with the task partner, using the Inclusion of Other in Self (IOS) scale (Aron et al., 1992). We did so by administering both measures at the start of the experimental session to record information about participants' baseline self-other overlap with their partner, and then repeating the measures after synchronous and asynchronous coordination. In addition to this, we included a continuous measure of connectedness during the coordination task with the aim to test for more immediate effects while coordinating with their task partner that may diminish shortly after the interaction has taken place.

As in Chapter 2, we expected that moving in synchrony with another person would lead to increased motor-level self-other overlap, indexed by stronger imitative tendencies, as well as perceived self-other overlap, as indexed by higher ratings of overlap on the IOS scale, compared to asynchronous movement. Moreover, we expected a change from participants'

baseline responses, such that moving in synchrony with another person would lead to an increase in self-other overlap, and moving asynchronously would lead to a decrease in self-other overlap, in both motor-level and perceived self-other overlap. Regarding participants' continuous ratings of connectedness with their task partner, we expected participants to feel more connected to their task partner during the coordination task when moving in synchrony with them compared to moving asynchronously.

Experiment 1

Methods

Participants

We recruited 34 right-handed, English-speaking adults (19F, 15M, $M_{age} = 24.85$ years) with normal or corrected-to-normal vision and no motor impairments to take part in this study.²⁶ Participants completed the study in pairs. We used a Bayesian sequential testing design (Schönbrodt & Wagenmakers, 2018a) with a minimum sample size of 20 (following Schönbrodt et al., 2017). 14 participants were excluded because they did not meet our preregistered inclusion criteria for the coordination task (see Data Treatment section below). Once our minimum sample size was reached, we computed our main analysis and stopped data collection, as the criteria for our stopping rule ($BF > 3$ in support of either H_1 or H_0) was reached. All study procedures were preregistered prior to data collection on the Open Science Framework (OSF): <https://osf.io/av75c>.

Participants were recruited via the Sona Systems database at Central European University PU. The study took place in the laboratories of the Social Mind Center at CEU in Vienna and lasted 75 minutes. Participants received 15€ for their participation; the rate

²⁶ One participant slipped through the pre-screening and turned out to be left-handed.

increased by 5€ when the session exceeded 90 minutes. All study procedures were approved by the Psychological Research Ethics Board (PREBO) at CEU.

Design

This study used a within-subject design. Participants completed a coordination task in two conditions: Synchrony and Asynchrony. The order of conditions was counterbalanced across participants. This constituted our main experimental manipulation.

Participants completed Inclusion of Other in Self (IOS) ratings (Aron et al., 1992) and an automatic imitation task (Sowden & Catmur, 2015) as dependent measures of self-other overlap. They completed each of these measures in three conditions: Baseline (before the first block of the coordination task), Synchrony, and Asynchrony. The three conditions were blocked: participants first completed the Baseline condition, then they completed either the Synchrony or and the Asynchrony condition, before completing the other of these two.

The automatic imitation task additionally included the factors Hand, with the levels Left and Right; and Imitative Compatibility, with the factors Compatible and Incompatible. The factor Hand was blocked; participants used either their Left hand or their Right hand to respond in each block. The order of blocks was counterbalanced between participants and constrained such that each level of this factor was completed once per condition before any level is repeated. The factor Imitative Compatibility was block randomized; Compatible and Incompatible trials were presented in random order in each block.

Materials and Procedure

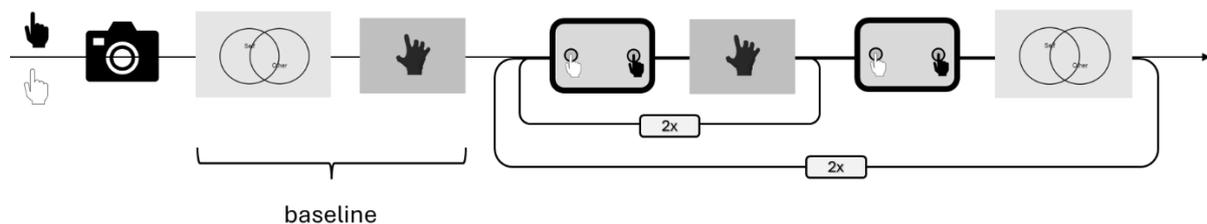
Overall Sequence of Tasks

After providing informed consent, participants were assigned a pair of black or white gloves to wear for the duration of the experiment. Then they were given an overview of the

experimental procedure (see Figure 13) and told that the experimenter would take pictures of their hands and that these pictures would later be used as stimuli for one of the tasks. We did this to ensure that participants would believe that they were seeing their task partner's hand in the automatic imitation task. However, the stimuli were ready-made images generated in Adobe Photoshop based on one image of a gloved hand, and were identical across dyads.

Figure 13

Procedure – Experiment 1



Note. This figure depicts the experimental procedure (from left to right). First, each participant was assigned a pair of gloves to wear (either black or white), and pictures were taken of both participants' hands. Then participants completed the baseline IOS ratings for their task partner, as well as an automatic imitation task as a baseline measure of their imitative tendencies. This was followed by a block of the coordination task and the automatic imitation task, once using their left hand, once using their right hand, and a block of the coordination task followed by IOS ratings in one condition (e.g., synchrony). Then this latter part of the procedure was repeated in the other condition (e.g., asynchrony).

The experimenter then took pictures of each participant's right and left hand in resting position, with a raised index finger and with a raised middle finger, respectively. Once all six pictures were taken of each participant's hands, the experimenter pretended to transfer the images from the camera to the individual laptops, then handed one to each participant for them to start the baseline phase of the experiment. Participants completed baseline IOS ratings for

their task partner, as well as an automatic imitation task with both their left and right hand to measure their baseline imitative tendencies.

Then participants began the first block of the experiment. Each of the six blocks of the experiment began with two trials of a coordination task, followed by the automatic imitation task (blocks 1, 2, 4, and 5), and IOS ratings (block 3 and 6). After the last block of the experiment, participants additionally completed a final questionnaire. There was an occluder between each participant's laptop and the tablet on which participants performed the coordination task together, so participants could not see each other's screens or hand movements while performing the individual tasks. The first three blocks were performed in one condition (e.g., synchrony), the last three blocks were performed in the other (e.g., asynchrony).

Coordination Task

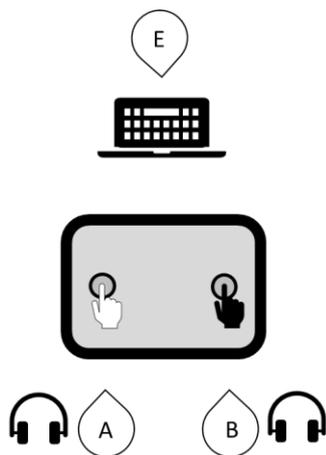
The coordination task followed the same general structure in both conditions. Participants sat next to the confederate facing a 12.9-inch iPad Pro with two buttons displayed on it (see Figure 14). The experimenter sat opposite them behind a MacBook Pro. Participants performed the coordination task together by tapping on a virtual button on the tablet that was placed on the table between them, producing a tone with each tap.

At the beginning of each trial, each participant heard a metronome (played by a drum) through a set of headphones (either 59bpm or 83bpm) and was instructed to move in time with it. Participants were instructed to tap on a virtual button each time they heard a tone. Whenever they tapped on the button, their action also produced a tone (a piano note: C). The metronome was only present for the first 20 seconds of each trial, then it stopped. While the metronome was present, participants only heard the metronome and their own taps, not the other participant's taps. Once the metronome stopped, participants began hearing their task partner's

taps as well (also a piano note: E). Participants were instructed to continue moving at the pace of the metronome once it had stopped until they received a signal indicating that the trial was over.

Figure 14

Coordination Task – Experiment 1



Note. Participants (A and B) performed the coordination task together on a 12.9-inch iPad wearing either white or black colored gloves. They sat next to each other facing a tablet, and their task was to tap on a virtual knob. The experimenter (E) sat opposite the participants and controlled the start of each trial. Participants were instructed to move either synchronously or asynchronously with the confederate, depending on the condition. For the first 20 seconds of each trial, they were paced by a metronome played through a set of headphones, then they continued without the metronome while trying to keep the same pace.

In the synchronous condition, both participants heard identical metronomes (both heard 59bpm or both heard 83bpm), and they were additionally instructed to try to move in time with each other once the metronome stopped. In the asynchronous condition, each participant heard a different metronome (one heard 59bpm, the other heard 83bpm), and they were additionally instructed to try to stick to their own pace.

Participants completed four practice trials (two with their right hand, followed by two with their left), before starting the first experimental block of each condition.²⁷ Experimental blocks consisted of two trials à 60 seconds. Both participants used the same hand (either their left or their right hand) in each block; the order was counterbalanced. There were six blocks of the coordination task in total; participants completed three blocks with their right hand and three with their left. The on- and offsets of their taps were recorded during the task.

Ratings of Connectedness

While performing the coordination task, participants were additionally instructed to continuously rate their feeling of connectedness with their task partner using M-Audio EX-P expression pedals with their feet (see Figure 15). They were told to keep their foot on the pedal for the duration of the trial and to apply pressure, or adjust the angle of the pedal, using their foot. The more connected they felt, the more pressure they should apply using the front part of their foot. Participants were encouraged to try out the pedals before the first practice trials of the coordination task and given feedback by the experimenter about the amount of pressure they applied and the range of the scale.

The scale ranged from 0 to 127. Pedals were connected to the MacBook Pro operated by the experimenter via Akai Professional MAX25 MIDI controllers, and the data was saved via the same Max MSP script which also recorded the coordination data. However, due to technical failure, the data could not be analyzed in Experiment 1 of the present chapter and are only reported for Experiment 2.

²⁷ Practice trials lasted 40 seconds for the first 20 participants. As half of these failed to meet our inclusion criterion for the coordination task, we increased the duration of the practice trials to match the real trials (i.e., 60s) for all subsequent participants.

Figure 15

Ratings of Connectedness



Note. Participants operated an expression pedal with their foot during the coordination task to indicate how connected they felt to their task partner. They could continuously adjust the pedal by changing the angle of their foot to indicate any level of connectedness from not feeling connected at all (A) to feeling maximally connected (B) to their task partner.

Automatic Imitation Task

We used a two-alternative forced-choice automatic imitation task adapted from Sowden & Catmur (2015), with key release serving as our dependent measure. Participants were shown right- and left-hand images of a hand wearing a glove matching the one that their task partner was wearing from a first-person perspective, and their task was to respond to geometrical shape that was displayed together with the hand on the screen (see Figure 16).²⁸ As participants sat next to each other in this experiment, their perspective on the hand remained the same between the coordination task and the automatic imitation task.

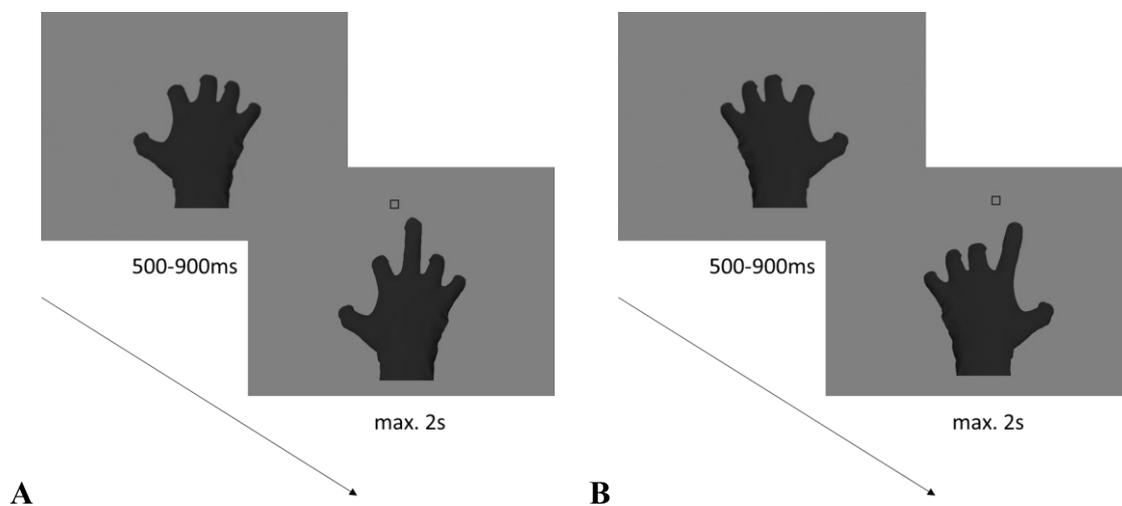
Participants were instructed to press two keys marked with red stickers on their computer keyboard (B and N) with the index and middle finger of their right or left hand while

²⁸ The full set of stimulus images (Figure S1) and example stimuli for each symbolic cue (Figure S2) can be found in the supplementary information.

viewing the images, and to release either their index or their middle finger in response to the geometrical shape (i.e., to lift their index finger when they saw a circle, and to lift their middle finger when they saw a square). The keys were marked with red stickers in this experiment to avoid any conflict with the colors of the gloves participants were wearing, and participants responded to shapes, so that color only served as a cue to hand identity.

Figure 16

Automatic Imitation Task – Experiment 1: Trial Structure



Note. Example of right-hand (A) and left-hand stimulus images (B). In the automatic imitation task, participants were shown an image of a hand in resting position (top left image in both A and B) for 500 to 900ms. They were then shown an image of the same hand with either a raised middle finger (bottom right image in A) or a raised index finger (bottom right image in B). The second image was displayed together with a geometrical shape (here: a square) for a maximum of 2 seconds, or until a response was made. The depicted finger movement was either compatible (A) or incompatible (B) with the cued response (here: a middle finger lift).

To manipulate Compatibility, we varied the finger movement participants were shown together with the geometrical shape. In half of the trials, participants saw a movement that was

compatible with the prompted response (e.g., a raised middle finger presented together with a square, as in Figure 16A). In the other half, they were shown a movement that was incompatible with the one they were prompted to perform (e.g., a raised index finger together with a square, as in Figure 16B). Compatible stimulus images also depicted the same hand as participants were instructed to use (i.e., participants were shown an image of a right hand performing the finger movement when they were using their right hand to respond, as in Figure 16A), whereas incompatible stimulus images depicted the other hand (i.e., participants were shown left-hand images when they were using their right hand to respond, as in Figure 16B).

Participants first completed two rounds of practice trials using their right hand, with each round consisting of eight trials, before completing a block of baseline trials with their right hand and repeating the same procedure with their left hand.²⁹ Baseline blocks of the automatic imitation task had the same structure as experimental blocks. Each block of the automatic imitation task consisted of 24 trials. Compatible and incompatible trials were randomly intermixed in each block, and each trial followed the same structure as in the second round of practice trials: participants were shown an image of a hand in resting position for 500 to 900ms, then an image of the same hand with a raised index finger or middle finger together with a geometrical shape for a maximum of 2 seconds (or until participants responded). Experimental blocks of the automatic imitation task were preceded by two trials of the coordination task, using the same hand as for the automatic imitation task. We recorded participants' responses during this task (response times and whether the response was correct).

Inclusion of Other in Self Ratings

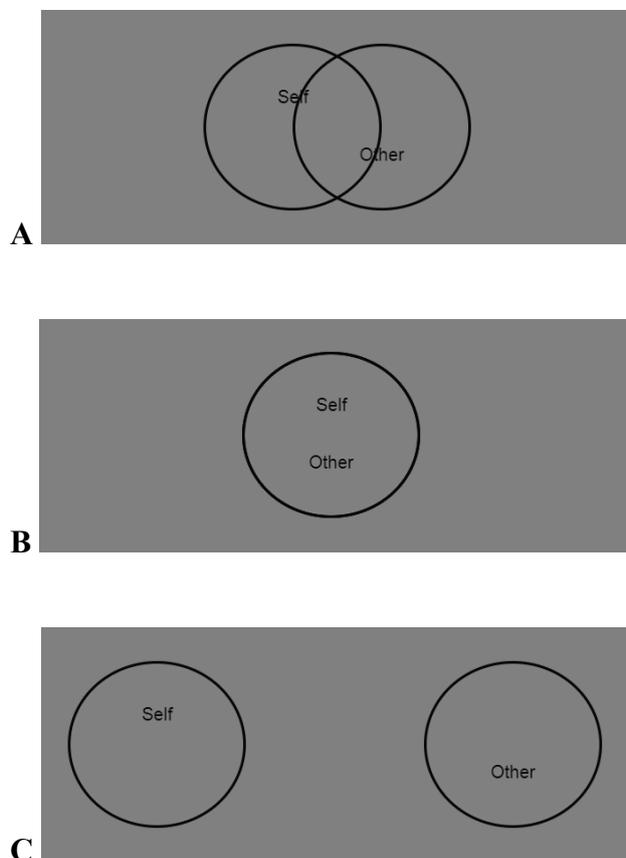
For the IOS ratings, we used a continuous version of the IOS scale designed by Aron et al. (1992). Participants were instructed to arrange two circles, labeled 'Self' and 'Other', in

²⁹ A more detailed description of the practice phase can be found in the supplementary information.

the way that best described their relationship with their task partner, and with experimenter 1. The circles were arranged in such a way that they had an overlap of 50% at the start of each rating. Participants could move the circle labeled ‘Self’ left and right using the arrows on their computer keyboard. In doing so, the circle labeled ‘Other’ automatically moved accordingly as well, either closer to the other circle or further away from it.

Figure 17

Inclusion of Other in Self Ratings



Note. This figure depicts the two circles participants were asked to arrange so they depict their relationship with their task partner in the IOS ratings. Circles were displayed at +50% overlap at the start of each rating (A) and could be arranged anywhere between a maximum of +100% overlap (B) and a minimum of -100% overlap (C).

The circles could be moved together to a maximum of complete overlap (see Figure 17B), and they could be moved apart to a maximum of -100% overlap (i.e., so the distance between the two circles was equivalent to one circle's diameter; see Figure 17C). Participants were familiarized with the range of the scale at the start of the experimental session, as they were instructed first to move the circles closer together so they overlap completely, then to move them as far apart as possible, before they were asked to complete the baseline IOS ratings with respect to their task partner. The circles reset to their starting position of 50% overlap before each new rating began to keep a consistent anchor across ratings.

Participants performed a rating of overlap with their task partner at the beginning of the experiment as a baseline measure of perceived self-other overlap, and another after the 3rd and 6th block of the coordination task, respectively, to measure how synchronous and asynchronous movement affected participants' perceived self-other overlap. We recorded the percentage of overlap between the two circles (-100 to +100%) in each rating.

Final Questionnaire

At the end of the experiment, participants filled out a final questionnaire. They were asked to rate their enjoyment of the coordination task, the difficulty of the coordination task, and the difficulty of the automatic imitation task. All ratings were performed on a scale from 1 to 7. The questionnaire additionally included questions about participants' interpretation of the IOS scale, about how connected they felt to their task partner during different phases of the experiment, and about their experience with music and dance. It also probed their hypotheses as to what the experiment was about and gave them the opportunity to provide comments and feedback.

Data Treatment

For our main analyses, we used paired-samples two-sided Bayesian t tests to compare H_1 (condition 1 \neq condition 2) to H_0 (condition 1 = condition 2) using a Cauchy prior. As preregistered, we interpret a $BF > 3$ in support of either H_1 or H_0 as substantial evidence for the existence of an effect (in case H_1 is supported) or no effect (if the BF is in support of H_0).

Coordination Data

We excluded participants who failed to follow the task instructions. More specifically, we excluded participants who spent more than 60% of trials in the asynchrony condition in a stable coordination pattern of in- or anti-phase synchrony (0° or 180°), as well as participants who spent less than 40% of trials in the synchrony condition in a stable coordination pattern of in-phase synchrony (0°).

To calculate participants' asynchronies, we first computed participant A's mean inter-tap-interval (ITIs) and used this to define the range of ITIs to consider, excluding any intervals that were longer than the mean ITI plus 2.5 standard deviations as well as any intervals that were shorter than the mean ITI minus 2.5 standard deviations. We then computed a new mean ITI after these exclusions, which served as an anchor for our calculation of asynchronies. To calculate the asynchronies, we considered each of participant A's taps and recorded the timestamps of participant B's taps that fell either within half of participant A's final mean ITI before or after participant A's tap. Of these, the one closest in time to participant A's tap was used to calculate the asynchrony for that tap. This was repeated for each of A's taps.

We computed percentages of trials spent in in-phase synchrony by computing the mean inter-tap-interval (ITI) for each trial and then evaluating the percentage of absolute asynchronies in each trial which were smaller than 10% of the mean ITI (e.g., for a mean ITI of 1000ms, we computed the percentage of absolute asynchronies that were smaller than

100ms). We computed percentages of trials spent in anti-phase synchrony by computing the mean ITI for each trial and then evaluating the percentage of absolute asynchronies in each trial which were 40 and 60% of the ITI (e.g., for a mean ITI of 1000ms, we computed the percentage of absolute asynchronies that were between 400 and 600ms).³⁰

For the included dyads, we computed their mean absolute asynchronies for each trial and used these to compute the mean absolute asynchrony for each condition and compared these using a paired-samples two-sided Bayesian t test to test whether our manipulation was successful in inducing more closely coordinated taps in the synchrony condition compared to the asynchrony condition.

Ratings of Connectedness

Due to a technical error, the initial position of the expression pedal at the start of each trial was not recorded, and new values were only saved once participants changed the position of the pedal, thus making any analyses of these incomplete data hard to interpret. We therefore do not report our analyses of the expression pedal data for this experiment.

Automatic Imitation Data

Trials in which participants did not respond to the cue image (9 trials; i.e., 0.31% of all trials) or in which they responded before the cue image appeared (6 trials; i.e., 0.21% of all trials) were excluded from the analysis. For the remaining trials, we first computed whether participants responded correctly or incorrectly in each trial. Participants who made mistakes in more than 10% of trials were excluded from the analysis. Correct and incorrect trials were then considered independently. No participant made mistakes in more than 10% of trials; all participants' data were entered into the analyses.

³⁰ This calculation was computed after each trial during the data collection process after the first 20 participants in order to give feedback to participants about whether the trial had to be repeated or not.

For correct trials, we computed participants' response times (RTs). RTs slower than the group mean plus 2.5 standard deviations were excluded from the analysis (76 trials; 2.74% of correct trials). Using the remaining data, we first computed participants' mean RTs for compatible and incompatible trials and ran a paired-samples two-sided Bayesian t test to test for an effect of imitative compatibility. Next, we generated interference scores for each condition by subtracting each participant's mean RT on compatible trials from their mean RT on incompatible trials in each condition. We used these for our main analysis, in which we compared participants' interference scores in the synchrony condition to their interference scores in the asynchrony condition using a paired-samples two-sided Bayesian t test. To test our secondary predictions, we additionally used two paired-samples two-sided Bayesian t tests to compare participants' interference scores in the synchrony condition and in the asynchrony condition, respectively, to baseline.

For incorrect trials, we computed participants' error rates (ERs). Using these, we first computed participants' mean ERs for compatible and incompatible trials and ran a paired-samples two-sided Bayesian t test to perform a manipulation check and test for an effect of imitative compatibility. Next, we generated difference scores for each condition by subtracting each participant's mean ER on compatible trials from their mean ER on incompatible trials in each condition. We used these for our main analysis, in which we compared participants' difference scores in the synchrony condition to their difference scores in the asynchrony condition using a paired-samples two-sided Bayesian t test. To test our secondary predictions, we additionally used two paired-samples two-sided Bayesian t tests to compare participants' difference scores in the synchrony condition and in the asynchrony condition, respectively, to baseline.

Inclusion of Other in Self Ratings

For our main analysis, participants' IOS ratings were compared between conditions using a Bayesian paired-samples t test. For our secondary analyses, participants' IOS ratings in each condition (synchrony, asynchrony) were compared to their baseline IOS ratings in two Bayesian paired-samples t tests to test for an in- or decrease following our manipulation.

Questionnaire Data

Participants' ratings of difficulty and of enjoyment of the coordination task were compared between conditions using two Bayesian paired-samples t tests.

Results

Coordination Data

Seven dyads were excluded because they did not meet our inclusion criteria for the coordination task. The remaining ten dyads spent less than 60% of asynchrony trials in a stable coordination pattern ($38.75 \pm 6.58\%$) and more than 40% of synchrony trials in in-phase synchrony ($64.71 \pm 11.26\%$). The analysis comparing these dyads' mean absolute asynchronies in the synchrony and asynchrony condition revealed that participants' taps were much more closely coordinated in the synchrony condition ($64.22 \pm 14.08\text{ms}$) than in the asynchrony condition ($170.11 \pm 26.87\text{ms}$): $BF_{10} = 4139.79$, suggesting that our manipulation was successful.

We additionally computed an exploratory analysis of participants' inter-tap-intervals (ITIs) to test whether participants successfully kept the predefined tempo. To do so, we computed separate Bayesian one-sample t tests on participants' mean ITIs in trials when they were paced with the slower metronome (target interval: 1017ms) and trials when they were paced with the faster metronome (target interval: 723ms), respectively. The analysis showed

that participants' mean ITIs were substantially faster than the target tempo both for the fast tempo ($666 \pm 44\text{ms}$): $\text{BF}_{10} = 1694.5$, and for the slow tempo ($812 \pm 72\text{ms}$): $\text{BF}_{10} = 90350000$.

We followed up on this result by looking at each condition separately. The analysis of participants' ITIs when they were completing the synchrony condition showed that participants substantially sped up once they could no longer hear the metronome and instead could only hear their own and their task partner's taps, both for the faster ($611 \pm 57\text{ms}$): $\text{BF}_{10} = 297458.92$, and for the slower target tempo ($741 \pm 91\text{ms}$): $\text{BF}_{10} = 252000000$. The analysis of participants' ITIs when they were completing the asynchrony condition, however, showed moderate evidence for no difference between participants' ITIs and the target tempo for the faster metronome ($720 \pm 54\text{ms}$): $\text{BF}_{01} = 418$, but extreme evidence for a difference between participants' mean ITIs and the target tempo for the slower metronome ($855 \pm 113\text{ms}$): $\text{BF}_{10} = 519.98$. When comparing participants' mean ITIs in each condition in two Bayesian paired samples t tests, we found that mean ITIs differed between conditions for both the faster ($\text{BF}_{10} = 16902.88$) and the slower metronome ($\text{BF}_{10} = 116.49$). Together, these analyses suggest that although participants were not successful at keeping the target tempo overall, they were more successful in doing so in the asynchrony condition than in the synchrony condition.

Automatic Imitation Data

Response Times

We did not find evidence for an effect of imitative compatibility in the analysis of participants' mean RTs. The paired-samples Bayesian t test comparing participants' RTs in response to compatible and incompatible trials showed moderate evidence for H_0 : $\text{BF}_{01} = 3.79$, suggesting that participants' RTs were not substantially faster in compatible ($0.54 \pm 0.06\text{s}$) compared to incompatible trials ($0.54 \pm 0.06\text{s}$). The analysis showed similar results when

considering blocks in which participants used their left hand ($BF_{01} = 1.47$) and blocks in which they used their right hand ($BF_{01} = 3.68$) separately.

The paired-samples Bayesian t test comparing participants' interference scores in the synchrony and asynchrony condition similarly showed moderate evidence for H_0 : $BF_{01} = 3.45$, suggesting that participants' interference scores did not substantially differ between the synchrony (0.01 ± 0.02) and the asynchrony condition ($0.002 \pm 0.04s$). The interference scores also did not substantially differ between baseline ($-0.003 \pm 0.04s$) and the synchrony ($BF_{01} = 2.42$) or the asynchrony condition ($BF_{01} = 3.68$).

To exclude carryover or order effects, we computed a Bayesian mixed ANOVA with the three-level within-subject factor Condition (baseline, synchrony, asynchrony) and the two-level between-subject factor Order (synchrony first, asynchrony first). The analysis revealed that the null model, which assumes no effect of either factor or their interaction, was the best model of the data, 2.6 times better than the next best model, which included only the factor Order. We additionally computed a Bayesian repeated-measures ANOVA with the three-level within-subject factor Condition (baseline, synchrony, asynchrony) and the two-level within-subject factor Hand (right, left) to test whether there was a difference between blocks when participants used their dominant or their non-dominant hand. The analysis revealed that the null model was the best model of the data ($P_{M|data} = 0.54$), 1.82 times better than the next best model, including only the factor Hand.

Error Rates

The results of the ER analysis mirrored those of the RTs: We did not find evidence for an effect of imitative compatibility in the analysis of participants' mean error rates. The paired-samples Bayesian t test comparing participants' ERs in response to compatible and incompatible trials showed moderate evidence for H_0 : $BF_{01} = 4.3$, suggesting that participants'

ERs were not substantially lower in compatible (0.03 ± 0.03) compared to incompatible trials (0.03 ± 0.03). This was the case, at least anecdotally, when considering the left ($BF_{01} = 1.8$) and right hand ($BF_{01} = 3.12$) individually too, and a Bayesian repeated-measures ANOVA with the two-level within-subjects factors Imitative Compatibility (compatible, incompatible) and Hand (left, right) also showed that the null model was the best model of the data ($P_{M|data} = 0.51$), 2.61 times more likely than the next best model, which included only the factor Hand. As these results indicate that there likely was no effect of imitative compatibility present in participants' error rate data, we did not test for a modulation of this effect.

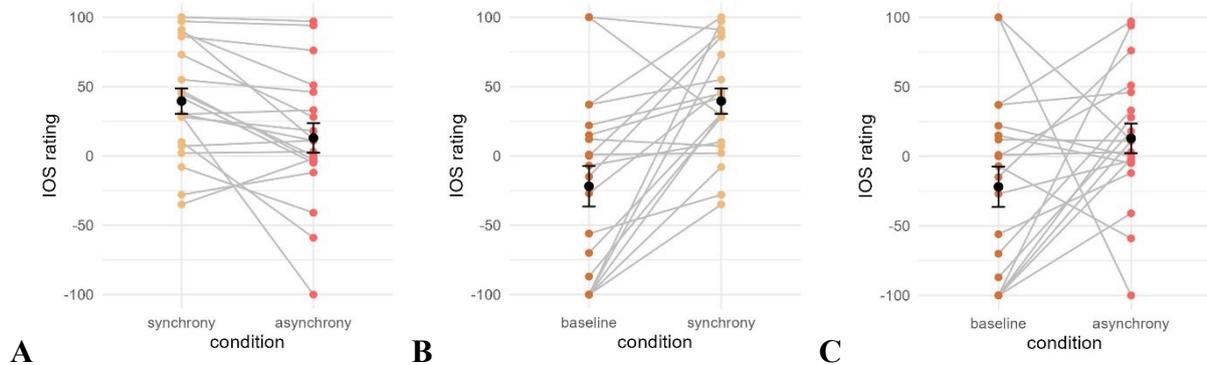
Inclusion of Other in Self Ratings

The paired-samples Bayesian *t* test comparing participants' IOS ratings in the synchrony and asynchrony condition showed moderate evidence for H_1 : $BF_{10} = 9.05$, with participants indicating more overlap between the circles after completing the synchrony condition (39.5 ± 40.91) compared to the asynchrony condition (12.9 ± 47.76 ; see Figure 18A). Compared to participants' baseline ratings (-21.9 ± 65.14), participants' IOS ratings substantially increased following the synchrony condition ($BF_{10} = 67.7$; see Figure 18B), but not the asynchrony condition ($BF_{01} = 1.14$; see Figure 18C).³¹

³¹ One participant indicated in the questionnaire that he interpreted the IOS scale in terms of “geographical proximity” the first time he filled it out. Excluding this participant did not substantially change the outcome of the analysis: Compared to participants' baseline ratings (-22.26 ± 66.9), participants' IOS ratings substantially increased following the synchrony condition ($BF_{10} = 34.21$), but not the asynchrony condition ($BF_{01} = 1.5$).

Figure 18

IOS Ratings – Experiment 1



Note. This figure depicts participants' IOS ratings (y-axis, ranging from -100 to 100%) in each condition (x-axis). Panel A depicts participants' IOS ratings in the synchrony and asynchrony condition. The other two panels compare participants' baseline IOS ratings to their IOS ratings following the synchrony condition (B) and the asynchrony condition (C), respectively. Colored dots depict each participant's IOS rating. Yellow dots depict ratings in the synchrony condition; pink dots depict ratings in the asynchrony condition; orange dots depict baseline ratings. Grey lines connect individual participant's ratings between conditions. The darker black dot depicts the mean IOS rating in that condition, and the black error bar represents the standard error.

Questionnaire Data

Overall, participants rated the difficulty of the coordination task as medium, with 40% of participants rating it a 4 on a scale of 1-7, and the remaining participants being evenly split between higher and lower ratings ($M = 3.75$, $SD = 1.65$). Participants indicated that they enjoyed the coordination task, with 68% of participants' ratings being greater than 4 out of 7 ($M = 4.68$, $SD = 1.57$). They rated the automatic imitation task as rather easy, with a mean rating of 2.95 out of 7 ($SD = 1.87$).

The paired-samples two-sided Bayesian t test showed moderate evidence for a difference in participants' ratings of the difficulty of the coordination task: $BF_{10} = 6.71$, with participants rating the synchrony condition to be easier (2.9 ± 1.62) than the asynchrony condition (4.45 ± 1.64). The paired-samples two-sided Bayesian t test showed anecdotal evidence for a difference in participants' ratings of their enjoyment of the coordination task: $BF_{10} = 1.35$, suggesting that participants may have enjoyed the synchrony condition (4.9 ± 1.85) somewhat more than the asynchrony condition (4.05 ± 1.51).

Exploratory correlation analyses showed that participants' ratings of the difficulty of the coordination task overall were negatively correlated with their enjoyment of the coordination task overall ($r = -0.52$, $BF_{10} = 3.04$), their enjoyment of the synchrony condition ($r = -0.62$, $BF_{10} = 11.65$), and their enjoyment of the asynchrony condition ($r = -0.54$, $BF_{10} = 4.05$). This also held for the relationship between their ratings of difficulty of the asynchrony condition and their enjoyment of the asynchrony condition ($r = -0.6$, $BF_{10} = 9.11$), but not for the relationship between their ratings of difficulty of the synchrony condition and their enjoyment of the synchrony condition specifically ($r = -0.26$, $BF_{01} = 2.09$).

Discussion

The inclusion of auditory feedback for participants' own and their partner's taps may have made the task more difficult to complete compared to previous versions (see Chapter 2), given the high rate of exclusions and the necessity of introducing an online inclusion criterion, implemented as participants were completing the task, for each trial (which still proved challenging for participants to meet). Interestingly, this appears to be reflected in the link between participants' enjoyment of the task and the perceived difficulty of the task, with participants who found the coordination task easier overall finding it more enjoyable, and participants who found the task more difficult overall finding it less enjoyable. However, the

introduction of auditory feedback, along with the more stringent inclusion criteria, may also have enhanced the effect of interpersonal synchrony on participants' experience of perceived self-other overlap.

In contrast to our findings in Experiment 1 in Chapter 2, participants in the present study showed greater perceived self-other overlap after tapping in synchrony with their task partner than after asynchronous tapping, and they showed a substantial increase in perceived self-other overlap following interpersonal synchrony compared to baseline. It is likely that the auditory feedback made the task partner's taps harder to ignore and made the difference between participants' own taps and their partner's more salient. At the same time, participants were forced to stick to the task instructions as closely as they could, since trials that did not meet the inclusion criteria had to be repeated. It may thus be either of these factors independently or the combination of auditory feedback and more stringent following of the instructions that enhanced this effect.

Despite having successfully elicited an imitation effect with this paradigm in Chapter 2, the changes in stimuli in the present chapter may have caused it to disappear. The stimuli in the present chapter differ from the ones used in Experiment 1 in Chapter 2 in two important ways: The present study used symbolic cues rather than color cues to prompt participants' responses. However, more importantly, the present stimuli depicted gloved hands rather than 'real' hands, which may have been less successful in priming participants' motor system and which furthermore may also inadvertently have highlighted the difference between self and other in a more salient way, as the depicted hand was wearing a different colored glove than the participant themselves. As participants' responses were also slower overall in this experiment compared to the previous experiment, this lends further support to the idea that their response was less automatic in the current experiment and that participants may have

relied on more symbolic processing of the stimuli, leading to a diminished effect of imitative compatibility.

As we did not successfully elicit an effect of imitative compatibility in this experiment, we were not able to meaningfully interpret our results from the automatic imitation task in light of our synchrony manipulation. We therefore aimed to remedy this by conducting a second experiment, in which we return to the automatic imitation task used in Experiment 1 in Chapter 2. In doing so, we aimed to replicate the effect of imitative compatibility that we found in Chapter 2, as this would allow us to interpret evidence for or against a modulation of this effect in a more meaningful way.

Experiment 2

Methods

The methods for Experiment 2 were identical to Experiment 1, with the following exceptions: Participants were not assigned gloves to wear, and we did not take pictures of their hands. They were not told that the hand they responded to in the automatic imitation task was their task partner's. Instead, they were told that the hand they see represents their task partner's hand. As the depicted hand was female, we recruited only female participants to take part in Experiment 2, and the experimenter was asked to rate the similarity between each participant's hand and the stimulus hand at the beginning of each session. We also modified the stopping rule, adding that we would not continue data collection once our minimum sample was reached if we found evidence against the presence of an effect of imitative compatibility, as this would render any potential modulation uninterpretable. Finally, we fixed the error in our script collecting the ratings of connectedness participants made during the coordination task using the expression pedals with their feet such that the initial position of the expression pedal was automatically recorded at the start of each trial.

Participants

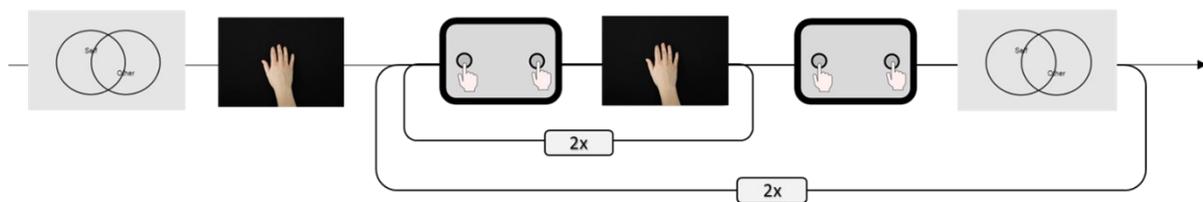
We recruited 20 right-handed, female, English-speaking adults ($M_{\text{age}} = 26$ years) with normal or corrected-to-normal vision and no motor impairments to take part in this study. Once our minimum sample size was reached, we computed our main analysis and stopped data collection, as the criteria for our stopping rule ($BF > 3$ in support of either H_1 or H_0) was reached. All study procedures were preregistered prior to data collection on the Open Science Framework (OSF): <https://osf.io/q24sy> (last accessed 2025-06-30).

Materials and Procedure

After providing informed consent, participants immediately started the baseline phase of the experiment (see Figure 19). The remaining overall structure was the same as in Experiment 1; the relevant differences are highlighted below.

Figure 19

Procedure – Experiment 2



Note. This figure depicts the experimental procedure (from left to right). First, participants completed the baseline IOS ratings for their task partner, as well as an automatic imitation task, in which they responded to a hand they were told represent their task partner's hand, as a baseline measure of their imitative tendencies. As in Experiment 1, this was followed by a block of the coordination task and the automatic imitation task, once using their left hand, once using their right hand, and a further block of the coordination task followed by IOS ratings in

one condition (e.g., synchrony), before this latter part of the procedure was repeated in the other condition (e.g., asynchrony).

Coordination Task

We used the same procedure as in Experiment 1, with practice trials matched in duration to real trials and the experimenter providing feedback about participants' performance after each trial and repeating trials up to three times where necessary.

Ratings of Connectedness

The script was modified such that the initial position of the expression pedal at the start of each trial was automatically recorded and could be used to determine participants' ratings of connectedness in the initial phases of each trial.

Automatic Imitation Task

As in Experiment 1, we used a two-alternative forced-choice automatic imitation task adapted from Sowden & Catmur (2015), with key release serving as our dependent measure. We used the stimuli generated for Experiment 1 in Chapter 2 for this experiment. Of the two sets of stimuli generated for Experiment 1 (Chapter 2), we opted for the images of the confederate's hand, as we found the images of the confederate's hand to elicit a greater effect of imitative compatibility than the images of the experimenter's hand. We chose the images without the added mark because we reasoned that they would seem more natural than those with the added mark.³²

Participants thus saw images of a hand depicted against a black background (see Chapter 2, Supplementary Figure 2) displayed together with a colored rectangle. As in Experiment 1 of Chapter 2, participants were instructed to respond to the colored rectangle by

³² Detailed analysis outcomes can be found in the supplementary information.

lifting either their index finger or their middle finger. Unlike in Experiment 1 of Chapter 2, the color mapping was counterbalanced between participants in the present experiment, such that half of the participants responded to the combination purple-orange with their right hand and to yellow-blue with their left hand, and the other half of participants responded to the purple-orange combination with their left hand and to yellow-blue with their right hand. We opted for this change to account for color mapping as a potential confounding factor affecting participants' responses when using their left (nondominant) compared to their right (dominant) hand.

In addition to this, we also modified the instructions such that participants were told that the hand they would respond to represents their task partner's hand. As we aimed to measure participants' self-other overlap with their task partner in this task, but did not use images of their task partner's hands, this seemed like the most transparent (and credible) way to generate a link between the depicted hand in this task and their task partner's hand.

Data Treatment

For our main analyses, we used paired-samples two-sided Bayesian t tests to compare H_1 (condition 1 \neq condition 2) to H_0 (condition 1 = condition 2) using a Cauchy prior. As preregistered, we interpret a $BF > 3$ in support of either H_1 or H_0 as substantial evidence for the existence of an effect (in case H_1 is supported) or no effect (if the BF is in support of H_0).

Ratings of Connectedness

The data from the expression pedals were cleaned for each participant in each trial, such that missing values were inserted and existing values were interpolated to reach a consistent sampling rate of 50 Hertz (one entry every 20ms). The files were then cut to include only the parts of the trial when the metronome was no longer present. For these data, the mean rating

was computed and used to compute a mean rating for each condition. The resulting means were then entered into the analysis.

Automatic Imitation Data

Six trials (0.21% of all trials) were excluded because participants did not respond to the cue image, and 26 trials (0.9% of all trials) were excluded because participants responded before the cue image appeared. 77 trials (2.8% of correct trials) were slower than the group mean plus 2.5 standard deviations and were therefore excluded from the analysis. The remaining correct trials were entered into the response time analysis. No participant made mistakes in more than 10% of trials; all incorrect trials were entered into the error rate analysis.

Results

Coordination Data

All dyads met our inclusion criteria for the coordination task, spending less than 60% of asynchrony trials in a stable coordination pattern ($41.3 \pm 5.8\%$) and more than 40% of synchrony trials in in-phase synchrony ($72.1 \pm 11.2\%$). The analysis comparing these dyads' mean absolute asynchronies in the synchrony and asynchrony condition revealed that participants' taps were much more closely coordinated in the synchrony condition ($53 \pm 11\text{ms}$) than in the asynchrony condition ($173 \pm 13\text{ms}$): $BF_{10} = 862827.67$, suggesting that our manipulation was successful.

We additionally computed an exploratory analysis of participants' inter-tap-intervals (ITIs) to test whether participants succeeded in keeping the predefined tempo. To do so, we computed separate Bayesian one-sample t tests on participants' mean ITIs in trials when they were paced with the slower metronome (target interval: 1017ms) and trials when they were paced with the faster metronome (target interval: 723ms), respectively. The analysis showed

that participants' mean ITIs were substantially faster than the target tempo both for the fast metronome ($661 \pm 49\text{ms}$): $BF_{10} = 1298.9$, and for the slow metronome ($892 \pm 109\text{ms}$): $BF_{10} = 435.73$.

We followed up on this result by looking at each condition separately. The analysis of participants' ITIs when they were completing the synchrony condition showed that participants substantially sped up once they could no longer hear the metronome and instead could only hear their own and their task partner's taps, both for the faster ($608 \pm 98\text{ms}$): $BF_{10} = 594.19$, and for the slower target tempo ($809 \pm 140\text{ms}$): $BF_{10} = 8044.75$. The analysis of participants' ITIs when they were completing the asynchrony condition, however, showed moderate evidence for no difference between participants' ITIs and the target tempo for the faster metronome ($714 \pm 59\text{ms}$): $BF_{01} = 3.57$, and anecdotal evidence for no difference between participants' mean ITIs and the target tempo for the slower metronome ($974 \pm 106\text{ms}$): $BF_{01} = 1.1$. When comparing participants' mean ITIs in each condition in two Bayesian paired samples t tests, we found that mean ITIs differed between conditions for both the faster ($BF_{10} = 25.68$) and the slower metronome ($BF_{10} = 3835.86$). Together, these analyses suggest that although participants were not successful at keeping the target tempo overall, they were more successful in doing so in the asynchrony condition than in the synchrony condition.

Ratings of Connectedness

The paired-samples Bayesian t test comparing participants' ratings of connectedness using the expression pedals during the coordination task in the synchrony and asynchrony condition showed strong evidence for H_1 : $BF_{10} = 28.03$, indicating that participants' ratings of connectedness with their task partner were substantially higher in the synchrony (71.01 ± 34.69) compared to the asynchrony condition (39.82 ± 26.64). This was also reflected in participants' responses to the questionnaire items about how connected they felt to their task

partner in the different phases of the experiment, with the majority of participants (60%) explicitly stating that they felt more connected to their task partner when they were trying to tap at the same pace than when they were instructed to stick to their own pace.

Automatic Imitation Data

Response Times

The paired-samples Bayesian t test comparing participants' RTs in response to compatible and incompatible trials showed moderate evidence for H_1 : $BF_{10} = 8.78$, indicating that our manipulation was successful in eliciting an effect of imitative compatibility, with participants responding faster to compatible ($0.49 \pm 0.05s$) compared to incompatible trials ($0.5 \pm 0.05s$). Looking at participants' responses when using their dominant and non-dominant hand separately, we found similar results when participants used their non-dominant hand ($BF_{10} = 6.82$), but anecdotal evidence for H_0 when participants used their dominant hand ($BF_{01} = 1.09$). We followed up on this result using a Bayesian repeated measures ANOVA with the two-level within-subjects factors Imitative Compatibility (compatible, incompatible) and Hand (right, left). The analysis did not show evidence for an interaction between imitative compatibility and hand: $BF_{01} = 1.37$ for the comparison between the null model and the full model. The best model of the data was one that included only the factor Imitative Compatibility ($P_{M|data} = 0.54$), 5.17 times more likely than the null model.³³

Next, we computed our main analysis and ran a paired-samples Bayesian t test comparing participants' interference scores in the synchrony and asynchrony condition. The analysis revealed moderate evidence for H_0 : $BF_{01} = 3.34$, suggesting that participants' interference scores did not substantially differ between the synchrony (0.01 ± 0.03) and the

³³ We ran additional analyses to test whether the different color mappings influenced participants' responses in this task; these analyses can be found in the supplementary information.

asynchrony condition (0.01 ± 0.03 s). The interference scores also did not substantially differ between baseline (0.02 ± 0.04 s) and the synchrony ($BF_{01} = 1.49$) or the asynchrony condition ($BF_{01} = 2.54$), suggesting that acting in (a)synchrony did not lead to a modulation of imitative tendencies.

To exclude carryover or order effects, we computed a Bayesian mixed ANOVA with the three-level within-subject factor Condition (baseline, synchrony, asynchrony) and the two-level between-subject factor Order (synchrony first, asynchrony first). The analysis revealed that the null model was the best model of the data ($P_{M|data} = 0.37$), 1.64 times better than the full model. We additionally computed a Bayesian repeated-measures ANOVA with the three-level within-subject factor Condition (baseline, synchrony, asynchrony) and the two-level within-subject factor Hand (right, left) to test whether there was a difference between blocks when participants used their dominant or their non-dominant hand. The analysis revealed that the null model was the best model of the data ($P_{M|data} = 0.54$), 2.38 times better than the next best model, including only the factor Condition.

Error Rates

The paired-samples Bayesian Wilcoxon signed rank test comparing participants' ERs in response to compatible and incompatible trials showed moderate evidence for H_0 : $BF_{01} = 3.95$, suggesting that participants did not make more mistakes in response to incompatible (0.04 ± 0.03) compared to compatible trials (0.04 ± 0.04). This was the case when considering the left ($BF_{01} = 3.85$) and right hand ($BF_{01} = 3.73$) individually too, and a Bayesian repeated-measures ANOVA with the two-level within-subjects factors Imitative Compatibility (compatible, incompatible) and Hand (left, right) also showed that the null model was the best model of the data ($P_{M|data} = 0.58$), 3.26 times more likely than the next best model, which included only the factor Imitative Compatibility. As these results indicate that there was no

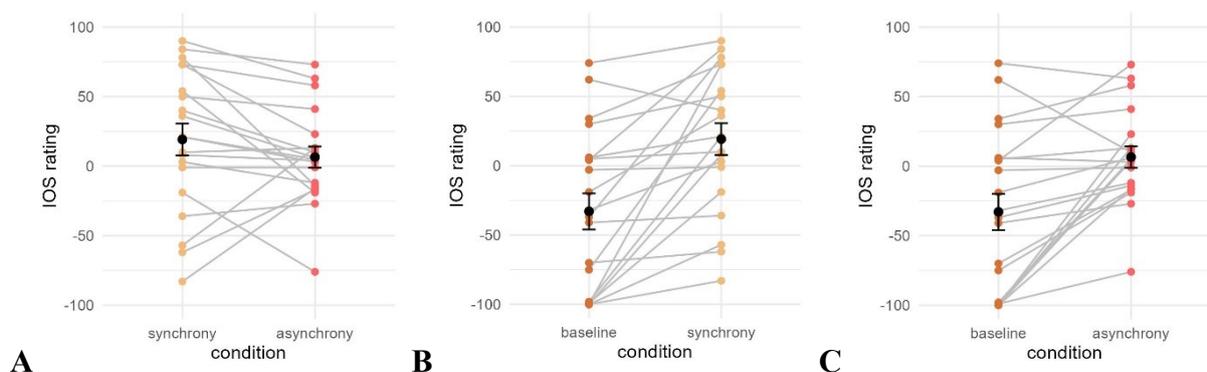
effect of imitative compatibility present in participants' error rate data, we did not test for a modulation of this effect.

Inclusion of Other in Self Ratings

The paired-samples Bayesian t test comparing participants' IOS ratings in the synchrony and asynchrony condition showed anecdotal evidence for H_0 : $BF_{01} = 1.88$, suggesting that participants' ratings of perceived self-other overlap with their task partner did not substantially differ following synchronous (19.15 ± 51.08) compared to asynchronous tapping (6.5 ± 34.26 ; see Figure 20A). Compared to participants' baseline ratings (-32.95 ± 58.09), participants' IOS ratings substantially increased following synchronous tapping ($BF_{10} = 100.65$; see Figure 20B), and asynchronous tapping ($BF_{10} = 28.62$; see Figure 20C).³⁴

Figure 20

IOS Ratings – Experiment 2



Note. This figure depicts participants' IOS ratings (y-axis, ranging from -100 to 100%) in each condition (x-axis). Panel A depicts participants' IOS ratings in the synchrony and asynchrony

³⁴ One participant indicated in the questionnaire that she interpreted the IOS scale in terms of "physical distance". Excluding this participant did not substantially change the outcome of the analysis: The Bayesian paired t test comparing participants' IOS ratings between the synchrony (19.74 ± 52.41) and the asynchrony condition (6.63 ± 35.2) still showed anecdotal evidence for H_0 : $BF_{01}=1.89$, and participants' IOS ratings substantially increased from baseline (-29.42 ± 57.44) following the synchrony condition ($BF_{10} = 47.97$), and the asynchrony condition ($BF_{10} = 14.84$).

condition. The other two panels compare participants' baseline IOS ratings to their IOS ratings following the synchrony condition (B) and the asynchrony condition (C), respectively. Colored dots depict each participant's IOS rating. Yellow dots depict ratings in the synchrony condition; pink dots depict ratings in the asynchrony condition; orange dots depict baseline ratings. Grey lines connect individual participant's ratings between conditions. The darker black dot depicts the mean IOS rating in that condition, and the black error bar represents the standard error.

Questionnaire Data

Overall, participants rated the difficulty of the coordination task as medium to high, with 25% of participants rating it a 4 on a scale of 1-7 and most participants (40%) rating it a 5 ($M = 4.2$, $SD = 1.28$). Participants' enjoyment of the coordination task was medium to low, with 20% of participants rating it a 4 on a scale of 1-7 and most participants (25%) rating it a 3 ($M = 3.9$, $SD = 1.62$). They rated the automatic imitation task as rather easy, with most participants (35%) rating it a 1 out of 7 ($M = 2.7$, $SD = 1.98$).

The paired-samples two-sided Bayesian t test showed moderate evidence for no difference in participants' ratings of the difficulty of the coordination task in the synchrony (4.05 ± 1.76) compared to the asynchrony condition (4.55 ± 1.73): $BF_{01} = 3.35$, suggesting that participants did not find one condition substantially more difficult than the other. The paired-samples two-sided Bayesian t test showed anecdotal evidence for no difference in participants' ratings of their enjoyment of the coordination task in the synchrony (4.4 ± 1.76) compared to the asynchrony condition (3.8 ± 1.67): $BF_{01} = 2.62$, suggesting that one condition also was not substantially more enjoyable than the other.

We computed exploratory correlational analyses between participants' ratings of difficulty and enjoyment of the coordination task overall and in each condition. The only supported relationship was the negative correlation between participants' ratings of the

difficulty of the coordination task when they were instructed to tap in time with their partner and their enjoyment of the coordination task in this condition: $r = -0.69$, $BF_{10} = 50.01$.

Discussion

As in Experiment 1, participants were able to successfully complete the coordination task, though it remained challenging for some. Interestingly, although participants in Experiment 2 were more successful at completing the coordination task than participants in Experiment 1, with no dyads having to be excluded, their ratings of the difficulty of the task were numerically higher, and their ratings of enjoyment were numerically lower than the ratings by participants in Experiment 1.

The coordination task elicited an increase in perceived self-other overlap. However, contrary to our predictions the increase was present in both conditions, and the two conditions did not elicit substantially different ratings of self-other overlap. This is consistent with the findings of Biswas & Brass (2024) in that they also found numerically higher IOS ratings following asynchronous movement compared to no movement, though they do not report if this comparison was significant and they still find the ratings following synchronous movement to be higher than those following asynchronous movement. It also appears consistent with the findings of no significant difference between social closeness following asynchronous movement and interpersonal synchrony by Tarr et al. (2016), although their results show a nonsignificant but negative change in pain threshold following interpersonal asynchrony. In a similar vein, Rabinowitch & Meltzoff (2017) found that four-year-old children showed more generous allocation of resources after swinging either synchronously or asynchronously with a peer, compared to children who did not complete a joint movement task. The authors interpret their findings as suggestive that children may have recategorized their task partner as an ‘in-

group' peer simply by performing the movement task together. It is possible that a similar process may have been at work in the present experiment.

Unlike Experiment 1, the present experiment was successful in eliciting an imitation effect regarding participants' RTs, though not their ERs. While covert and overt automatic imitation have been shown to be correlated in some contexts (Cracco et al., 2024), it has been argued that RTs provide an index of covert imitative tendencies (Cracco & Brass, 2019) and therefore provide the more relevant dependent measure for the current study. Interestingly, the effect of imitative compatibility seemed smaller in the present experiment than in Experiment 1 in Chapter 2, despite using the same stimuli. The larger sample size in Experiment 1 in Chapter 2, combined with the greater familiarity with the depicted hands, may explain this difference.

Whereas the present experiment found evidence for an effect of imitative compatibility, our data do not provide support for a modulation of this effect through acting in synchrony. More specifically, we found moderate evidence for no difference between participants' imitative tendencies following the synchrony compared to the asynchrony condition, and anecdotal evidence for no difference to baseline in either condition. As the evidence regarding changes to baseline was anecdotal, this suggests that the analysis may have been underpowered. However, the direction of this effect, along with the moderate evidence for no difference between conditions, suggests that moving in synchrony with another person does not modulate people's imitative tendencies or lead to a substantial increase compared to baseline.

It is worth noting that participants' ratings of connectedness during the coordination task were substantially higher during the synchrony condition than during the asynchrony condition. This suggests that participants felt more connected to one another while tapping in time with each other than when they were instructed to stick to their own pace, thus indicating

that interpersonal synchrony affects interpersonal connectedness. However, since we do not see clear effects of interpersonal synchrony compared to asynchrony in either of our other dependent measures, administered after the coordination task, this suggests that but that the feeling of connectedness elicited by interpersonal synchrony may have been short lived, present during the interaction but no longer as salient once confronted with a new task.

General Discussion

Across two experiments, we aimed to test whether moving in synchrony with another person would increase both motor-level and perceived self-other overlap. To do so, we recorded baseline measures of participants' motor-level and perceived self-other overlap with their task partner, then repeated the same measures again after participants completed a synchronous and an asynchronous coordination task, respectively. Our results showed evidence for an increase in perceived but not motor-level self-other overlap, suggesting that effects of interpersonal synchrony likely are not primarily mediated by an increase in motor-level self-other overlap, but may rely on a social recategorization according to changes in the levels of perceived self-other overlap instead.

As in Chapter 2, we found the manipulation of interpersonal synchrony in the present chapter to be successful in eliciting more closely coordinated movements in the synchrony condition and less closely coordinated movements in the asynchrony condition. The inclusion of auditory feedback in the present chapter made the task more challenging for participants to complete, with participants rating the difficulty of the coordination task to be higher in both experiments presented in this chapter compared to Experiment 1 in Chapter 2. This change in the coordination task likely also enhanced our manipulation by adding an additional cue to interpersonal (a-)synchrony that participants could use during the task to coordinate their taps and that also may have had a more lasting effect on participants' relationship with one another.

It is worth noting that our analyses of participants' tapping tempo with respect to the metronome showed that participants tapped faster than the predefined tempo overall both after hearing the fast metronome and after hearing the slow metronome. However, both Experiment 1 and Experiment 2 showed that participants were better at sticking to the predefined tempo in the asynchrony condition, particularly for the faster metronome. While this could simply be due to the emphasis on sticking to their own pace in the instructions for asynchronous coordination blocks, it may also be the result of better self-other distinction during interpersonal asynchrony that allowed participants to better distinguish their own taps from the other person's and adapt their behavior accordingly.

While Experiment 1 in the present chapter was not successful in eliciting an automatic imitation effect, there are several possible reasons for this. Focusing on differences between self and other has been shown to decrease imitative tendencies (Genschow et al., 2021). Although this was not our intention in the creation of the stimuli, it is possible that the depiction of hands wearing different colored gloves served not only as a cue to hand identity, but also highlighted the difference between the participant and the owner of the hands depicted in automatic imitation task stimuli. It is also possible that the gloves alone may have led the hands to seem less realistic and thereby decreased people's imitative tendencies (see also De Souter et al., 2021), or that using geometrical shapes as cues led participants to engage in more symbolic processing and to ignore the hand movements more easily (see also Pilot Study, Chapter 2).

We were, however, successful in eliciting an effect of imitative compatibility in Experiment 2. Yet, we still did not find a modulation of imitative tendencies after interpersonal (a)synchrony. This may have been due to the fact that in Experiment 2, similar to Experiment 1, participants were not shown pictures of their task partner's hands but instead were told that

the depicted hand should represent their task partner's hand. However, it is not clear that this should be crucial for an increase in self-other overlap to occur and for there thus to be an increase in people's imitative tendencies following interpersonal synchrony, or a modulation of participants' imitative tendencies following synchronous compared to asynchronous coordination. Moreover, a previous study found evidence for such a modulation using pictures that were not of participants' hands (Cross, Atherton, et al., 2021a). Yet our present results, together with those of the previous chapter, provide evidence to the contrary, demonstrating that interpersonal synchrony does not modulate people's imitative tendencies.

Whereas neither the Pilot Study nor Experiment 1 in Chapter 2 showed evidence for a modulation of perceived self-other overlap following interpersonal synchrony, both experiments in the present chapter showed a substantial increase in perceived self-other overlap following synchronous coordination. Experiment 1 found evidence for a modulation of perceived self-other overlap following interpersonal synchrony compared to asynchrony; however, this finding did not replicate in Experiment 2. Experiment 2 instead found evidence for an increase in perceived self-other overlap also following asynchronous coordination, which was not present in Experiment 1. Despite the differing findings for the asynchrony condition, these results suggest that interpersonal coordination leads to an increase in perceived self-other overlap, and that this effect may be enhanced through the presence of auditory feedback.

Moreover, participants' ratings of connectedness in Experiment 2 showed that they felt more connected when tapping in synchrony with one another than when they were tapping asynchronously. While we were not able to analyze participants' ratings in Experiment 1, it is possible that simply asking participants to indicate how connected they felt during the coordination task may have made their connectedness more salient in both conditions of the

task, giving rise to numerically higher ratings of perceived self-other overlap after asynchronous movement compared to baseline in Experiment 1 and substantially higher ratings in Experiment 2. It may thus have been the presence of auditory feedback combined with both more precise fulfillment of the instructions and the greater salience of interpersonal connectedness during the coordination task in the present study that contributed to us finding an increase in perceived self-other overlap following interpersonal synchrony.

Tying these results back to the theoretical assumptions introduced at the start of this chapter, we do not find support for the idea that interpersonal synchrony increases motor-level self-other overlap. However, our results do suggest that interpersonal synchrony leads to an increase in perceived self-other overlap, providing support for accounts that posit a social recategorization in terms of group membership as the key process underlying the social effects of interpersonal synchrony (e.g., Cross et al., 2019; Rabinowitch, 2023). The discrepancy between our results reported in the present chapter and those of Chapter 2, however, suggest that this mechanism does not hold in the face of synchronous movement alone but instead appears to rely on multiple sources of evidence, or multimodal cues to group membership, to produce a reliable effect.

Supplementary Information

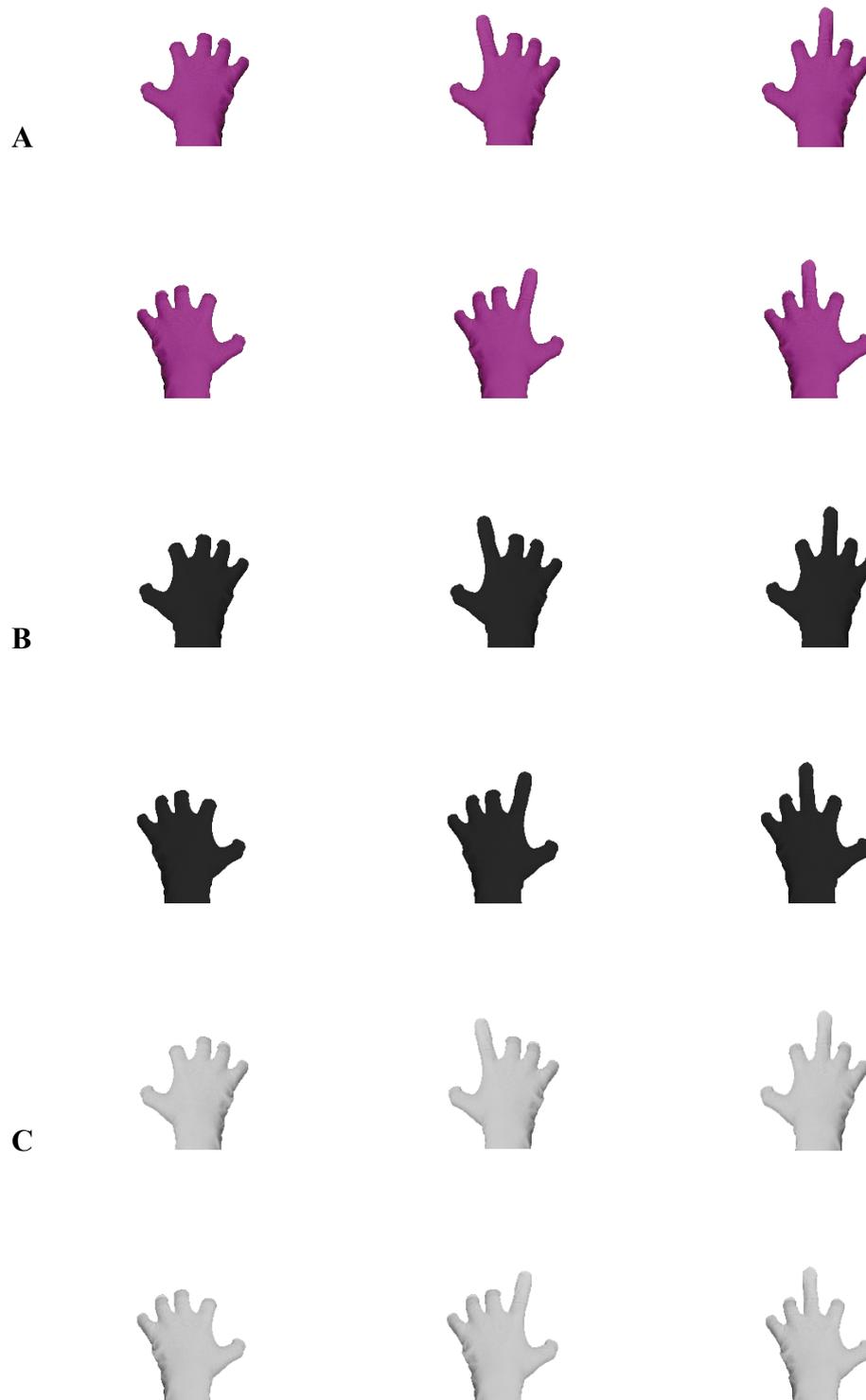
Experiment 1

Automatic Imitation Task

Stimuli

Supplementary Figure 7

Automatic Imitation Task – Stimulus Images

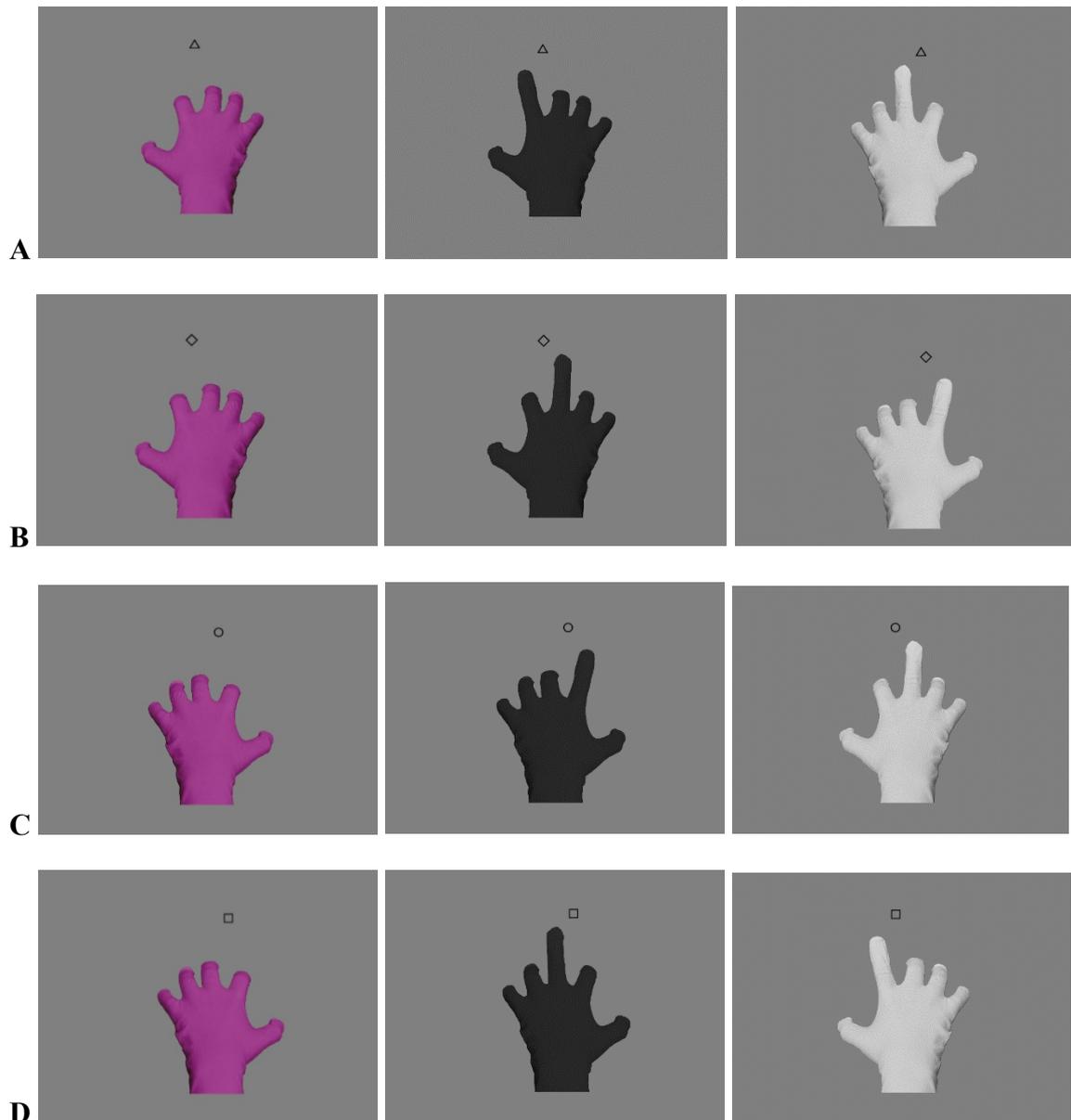


Note. This figure shows the images of gloved hands participants were shown in the automatic imitation task. All participants saw images of hands wearing a pink glove (A) during the

practice phase. During baseline trials and experimental blocks, participants saw a hand wearing the same colored glove as their task partner – either black (B) or white (C). All trials began with a hand in resting position (leftmost image in each row). In the first round of practice trials participants were then shown the same hand, still in resting position, together with a symbolic cue. In all other trials, participants the initial image of a hand in resting position was followed by an image of the same hand with a raised index finger (middle image in each row) or with a raised middle finger (rightmost image in each row) together with a symbolic cue. Participants were shown right- (top row for each glove color) and left-hand images (bottom row for each glove color) randomly intermixed within each block.

Supplementary Figure 8

Automatic Imitation Task – Example Stimuli



Note. This figure depicts example stimuli for each symbolic cue. Participants were shown the images of gloved hands depicted in Supplementary Figure 7 together with a symbolic cue: a triangle (A), a diamond (B), a circle (C), or a square (D). Symbols were paired such that in a given block, participants would either be instructed to lift their index finger in response to a triangle (A) and to lift their middle finger in response to a diamond (B), or they would be instructed to lift their index finger in response to a circle (C) and to lift their middle finger in

response to a square (D). One set of symbols was assigned to blocks when participants used their right hand to respond; the other was assigned to blocks when participants used their left hand to respond. The symbol-hand mapping was counterbalanced between participants.

Practice Trials

In the first round of practice trials, participants were shown an image of a hand wearing a pink glove in resting position for 500 to 900ms. This image was then replaced by an image of the same hand, still in resting position, together with a geometrical shape between the hand's index and middle finger. The second image was shown for a maximum of two seconds and disappeared once one of the two response keys was released. This round of practice trials was repeated as often as necessary until participants had learned the correct cue-response mapping and responded correctly on at least 75% of trials.

Participants then moved on to the second round of practice trials, in which they were shown the same image of a hand in resting position for 500 to 900ms, then an image of the hand with a geometrical shape, this time with a raised index or middle finger, for a maximum of 2 seconds. This round was also repeated as often as necessary until participants responded correctly on at least 75% of trials. After completing the practice phase with their right hand, participants moved on to complete a block of baseline trials with their right hand, then repeated the same procedure with their left hand.

Experiment 2

Automatic Imitation Task

Stimuli

As there the Bayesian paired-samples t tests showed evidence for the presence of an effect of imitative compatibility in participants' responses to images of both the confederate's

($BF_{10} = 1317.8$) and the experimenter's hands ($BF_{10} = 130.79$), and a comparison of interference scores showed no difference between the two sets of stimuli ($BF_{10} = 0.23$), we opted to use the stimuli depicting the confederate's hands, as the evidence for the presence of an imitation effect was found to be stronger, and the chances participants would encounter the individual whose hands they were responding to were lower.

Results

To test whether the different color mappings influence participants' responses in this task, we ran a Bayesian mixed ANOVA with the two-level within-subjects factor Imitative Compatibility (compatible, incompatible) and the two-level between-subjects factor Color Mapping (mapping 1, mapping 2). The analysis revealed that the best model of the data was one that included only the factor Imitative Compatibility ($P_{M|data} = 0.48$), 8.25 times more likely than the null model. We repeated this analysis with the additional two-level within-subjects factor Hand (right, left) and found that the model that included only the factor Imitative Compatibility remained the best model of the data ($P_{M|data} = 0.18$), 5.29 times more likely than the null model, suggesting that although the color mapping and the hand participants used to respond may have influenced participants' responses, imitative compatibility had a greater influence over their response times.

CHAPTER FOUR: THE SUBJECTIVE EXPERIENCE OF INTERPERSONAL COORDINATION

The present chapter addresses the topic of self-other relations during interpersonal synchrony from a different angle. The literature on interpersonal synchrony has documented effects such as increased affiliation (Hove & Risen, 2009), greater connectedness, trust, and feelings of being on the same team (Wiltermuth & Heath, 2009), or increased social bonding (Tarr et al., 2015). However, most lab studies aiming to test for these effects have relied on simple questionnaire items to measure their outcome of interest. Whereas Chapters 2 and 3 aimed to clarify the underlying processes using a more implicit approach, the present chapter aims to tackle the question of experience head-on. Instead of assessing people's experience during interpersonal synchrony using questionnaire items, we conducted in-depth interviews about people's experience of interpersonal coordination in order to gain a better understanding of the experience of interpersonal synchrony.

One strand of the literature on interpersonal synchrony has focused on its role in fostering group cohesion and in supporting social bonding (Launay et al., 2016; Tarr et al., 2014, 2015). This body of work has drawn on literature on music and dance (Ehrenreich, 2006; Freeman, 1998; McNeill, 1997; Savage et al., 2021), as well as collective rituals (Fischer et al., 2013; Durkheim, 2008; Haidt et al., 2008). They have posited that one common factor these practices share is the presence of synchrony and that this may be one of the key mechanisms driving their prosocial effects as well as the selection and retention of these practices from an evolutionary perspective (Reddish et al., 2013). Accounts of interpersonal synchrony that link it to social bonding have been linked to underlying neurohormonal mechanisms involving the endogenous opioid system (Cohen et al., 2009; Tarr et al., 2014) and tested by assessing changes in people's subjective pain thresholds as an indicator of endorphin activity, for

instance in the context of rowing (Cohen et al., 2009; P. Sullivan & and Rickers, 2013), treadmill running (P. J. Sullivan et al., 2014), drumming (P. Sullivan et al., 2015), walking (Lang et al., 2017), dancing (Tarr et al., 2015, 2016), or in a VR context (Weinstein et al., 2016), with most, but not all (P. Sullivan et al., 2015) studies finding a positive effect.

Other authors have emphasized the role of shared intentionality in social bonding (Kirschner & Tomasello, 2010a), and studies have found that interpersonal synchrony and shared intentions interact in fostering cooperative behaviors (e.g., Reddish et al., 2013; Cross, Atherton, et al., 2021). An alternative explanation for these findings has invoked the idea of social categorization and group identity (Tajfel, 2010; Tajfel et al., 1971), suggesting that interpersonal synchrony acts as a cue to group membership (Cirelli, 2018a; Cross et al., 2019b) and thereby elicits increased prosocial behavior towards in-group members, or in-group favoritism. Similarly, it has also been argued that music and dance serve an important function of coalition signaling (Hagen & Bryant, 2003), thereby facilitating social organization and coordination within social groups. The idea that the effects of interpersonal synchrony are directed towards the individuals involved in the interaction and members of their social group rather than general has found empirical support in both adults (e.g., Cross et al., 2020; Tarr et al., 2015) and infants (Cirelli et al., 2016), and studies have also found differing effects of interpersonal synchrony with in- and out-group members (Cross, Turgeon, et al., 2021; Good et al., 2017; Qian et al., 2020; Tunçgenç & Cohen, 2016a). However, other studies have found evidence for more generalized effects of interpersonal synchrony (e.g., Reddish et al., 2014, 2016), suggesting that the evidence is not clear cut (see Cross et al., 2020 for a discussion).

Importantly, it has been suggested that different types of movements may recruit different mechanisms involved in eliciting the prosocial effects of interpersonal synchrony. For instance, processes involving the endogenous opioid system may only be activated when

movements are sufficiently strenuous (Lang et al., 2017). Moreover, whereas interpersonal synchrony is thought to lead to an increase in positive affect (Ehrenreich, 2006), and there is some empirical evidence to support this view (e.g., Tschacher et al., 2014), other studies do not find this effect (e.g., Wiltermuth & Heath, 2009). This discrepancy in the literature has been linked to differences between laboratory contexts and ‘naturally occurring’ interpersonal synchrony in the contexts of rituals or music and dance (Mogan et al., 2017b), suggesting further that contextual factors, as well as the presence of music (Stupacher, Wood, et al., 2017), likely influence the mechanisms at play.

Furthermore, continuous and discrete movements are thought to rely on distinct timing control mechanisms (Huys et al., 2008b). Whereas the timing of continuous movement emerges from the movement dynamics themselves and has been argued not to require an external reference once a tempo has been set, the timing of discrete movements requires external time keeping even once the tempo has been established (Ivry et al., 2002; Spencer et al., 2003). This may imply different mechanisms for discrete and continuous interpersonal synchronization, with continuous coordination conceivably being more likely to perpetuate itself and discrete coordination relying on more constant monitoring of coordination success. Prosocial effects of interpersonal synchrony have been documented for both continuous (e.g., Cross, Atherton, et al., 2021; Tarr et al., 2016; Valdesolo et al., 2010) and discrete synchronization tasks (e.g., Good et al., 2017; Hove & Risen, 2009; Launay et al., 2013; see Chapter 2). However, it is possible that synchronization paradigms involving discrete movements rely more heavily on external timing cues, such as a metronome or action effects (sounds or visual effects) produced by one’s own or other people’s movements, than those involving continuous movement, and that these two types of coordination may elicit qualitatively different experiences.

The present study aimed to contribute to a better understanding of the experience of interpersonal synchrony in different coordination contexts. In particular, we aimed to explore how people experienced themselves in relation to their task partner during synchronous compared to asynchronous coordination. Moreover, we aimed to explore potential differences in people's experience of continuous and discrete coordination tasks. To do so, we used the micro-phenomenological interview (MPI) method (Petitmengin, 2006) to collect in-depth descriptions of people's experience of different types of interpersonal coordination – during synchronous and asynchronous coordination and using continuous and discrete paradigms. The MPI method allows a trained interviewer to guide participants in generating faithful descriptions of their subjective experience without requiring a trained interviewee. This method allows for an in-depth exploration of an evoked or previously lived experience without disrupting the unfolding of the experience itself. In doing so, both the diachronic structure of the experience, i.e., its temporal unfolding, and synchronic structure of the experience, i.e., its simultaneous aspects, are explored, making this method well-suited for the examination of short experiential episodes that unfold over time.

Methods

Participants

We recruited four dyads (eight participants, 5F, 3M, aged 22-30 years) to take part in the study. Participants were right-handed fluent English-speakers with normal or corrected-to-normal vision who were interested in exploring their experience. They had no motor impairments, and no history of neurological impairments or diagnoses. They were assigned a task partner for the duration of the study and completed all four sessions of the study together with their assigned partner. Participants were recruited via the Sona Systems database at Central European University PU. The study took place in the laboratories of the Social Mind

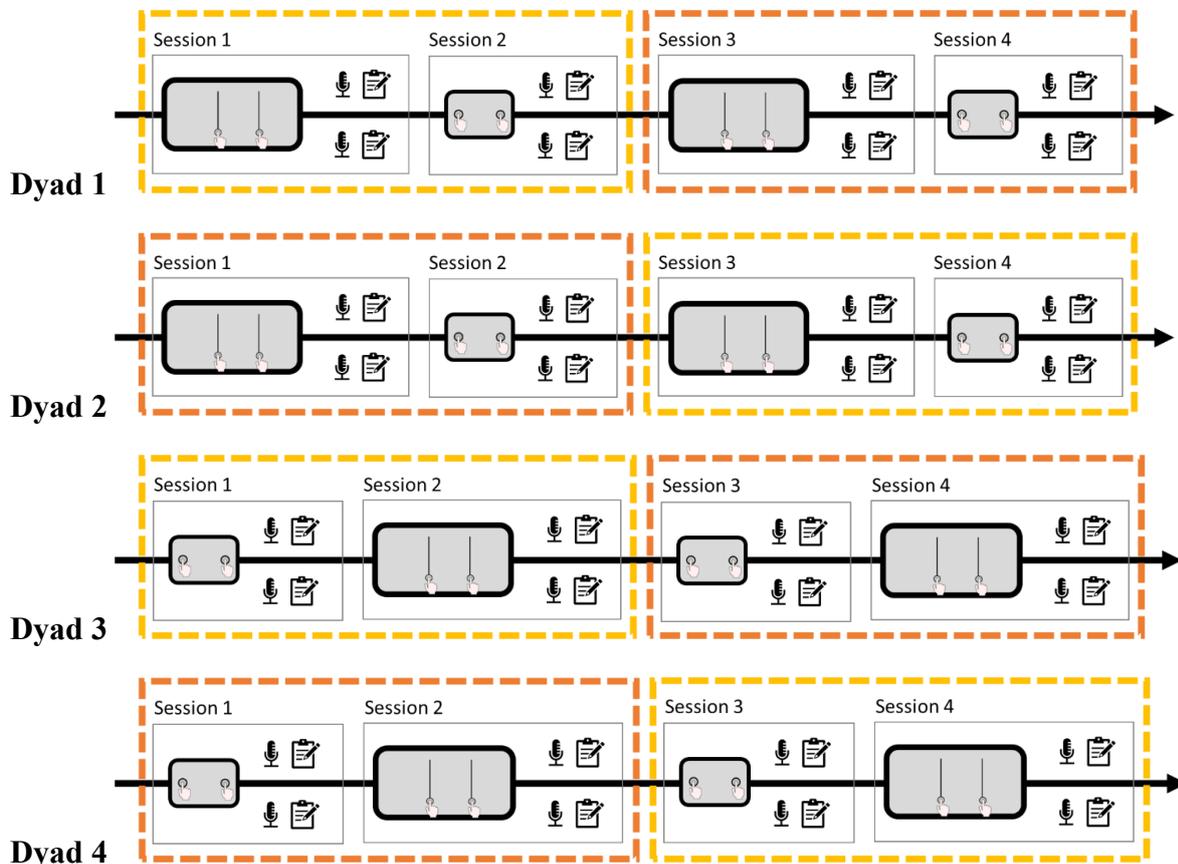
Center at CEU in Vienna and consisted of four one-hour sessions that all took place within the same week, at least an hour apart. Participants received 40€ for their participation (10€ for each session); the rate for a session increased by 5€ if the session lasted longer than 60 minutes. All study procedures were approved by the Psychological Research Ethics Board (PREBO) at CEU.

Materials and Procedure

Each dyad took part in four sessions. Participants completed each session with the same partner (i.e., the pairings remained constant for the duration of the study). Each session lasted approximately one hour and consisted of a coordination task and subsequent individual interviews (see Figure 21).

Figure 21

Procedure



Note. This figure illustrates the study procedure for each dyad. Each row represents the procedure for one dyad, from left to right, following the arrow. Each dyad took part in four sessions; each dashed grey box represents one session. Each session consisted of a coordination task, which was either continuous (represented by the larger rectangle with two vertical tracks, as in session 1 in the top row) or discrete (represented by the smaller rectangle with two buttons, as in session 2 in the top row), and an interview with each participant (represented by the microphone and clipboard in each session). Dashed colored boxes represent conditions; yellow represents synchronous coordination; orange represents asynchronous coordination.

Coordination Task

Participants completed a different coordination task in each session (see Figure 21). The tasks all followed the same structure: Participants were asked to coordinate their movements such that they were both moving at the same tempo (synchrony condition) or such that they were not moving at the same tempo (asynchrony condition) for a duration of one and a half minutes. The beginning and end of each trial were marked by a sound (which participants were familiarized with before the start of the first trial), and participants were given a chance to try out the apparatus before the start of the trial. Unlike Chapters 2 and 3, participants were not given a metronome to pace them but could choose to follow any pace, as long as it was either the same as (synchrony) or different from (asynchrony) their task partner's.

In two sessions, participants completed a sliding task together: They stood next to each other in front of a large touchscreen displaying two virtual knobs on a 30cm long track (see Chapter 2, pilot 1). Their task was to slide the knob up and down the track for one and a half minutes. We recorded the timestamped coordinates of their finger on the slider during this time. In the other two sessions, participants completed a tapping task together: They sat next to each other facing an iPad, which displayed two virtual buttons (see Chapter 2, experiment 1). Their

task was to tap on the buttons for one and a half minutes. We recorded the timestamped onsets of their taps on the button during this time. Both tasks were completed once in the synchronous condition and once in the asynchronous condition. Participants completed both tasks in one condition (e.g., synchrony) first before moving on to the other condition (e.g., asynchrony). Each dyad completed the tasks in a different order (see Figure 21).

Micro-phenomenological Interviews

The interviews were conducted following the guidelines for micro-phenomenological interviews (Petitmengin, 2006). Participants were informed that the aim of the interview was to understand what their experience was like during these different tasks, and that the focus would be on how they related to their task partner in each session. The structure of the interview was then explained to them, and after giving informed consent, participants were asked to provide an overview of their experience during the coordination task.

The interviewer then elicited a more detailed description of the experience using targeted questions, while helping the participant remain in touch with their experience. The latter was achieved by regularly repeating back what the participant had already described, by asking concrete questions about the particular experience, and by using pointers (such as gestures or generic terms) when an element being described was still vague in the participant's description (see Petitmengin, 2006, pp. 239-240). Once the interviewee was in touch with their experience, the questions aimed to be content-neutral but directive, in order to elicit a more precise description of the experience without imposing any particular interpretation. At the end of the interview, the interviewer summarized the experience and asked the participant if there was anything else they would like to add to or change about the description, then asked whether they felt that the final description was faithful to their experience and whether it was complete, following the validation criteria proposed by Bitbol & Petitmengin (2013).

All interviews were conducted by the same interviewer in English. The interviewer had completed a micro-phenomenological interview training (<https://www.microphenomenology.com/interview>, last accessed 2024-01-02), in which she was introduced to the method by an expert in the field and had ample opportunity to practice and receive feedback regarding her application of the key principles of the micro-phenomenological interview technique. In addition to this training, she had previously gained experience with this technique as both an interviewer and interviewee through her involvement in different research projects, including a comparison of two methods of empirical phenomenology (Bass-Krueger et al., 2024). While this experience and training do not rule out the presence of biases during the interview process, they serve to ensure a solid grasp of the method in order to mitigate bias and elicit faithful experiential reports by raising awareness of the risk of introducing bias into the interviewee's descriptions and supporting the early recognition of bias in the resulting experiential reports.

The interviewer was also present during the coordination task which preceded the interview. Each interview lasted for approximately 20 minutes. Participants were allowed to talk between sessions but were asked to refrain from discussing any content related to the study until they had completed all four sessions. They were fully debriefed after the final session.

Data Treatment

Coordination Data

Discrete and continuous coordination data was treated differently, as in Chapter 2. For the tapping task (discrete data), we computed participants' mean inter-tap-intervals (ITIs) by subtracting the onset of the previous tap from the current tap onset and compared these ITIs between conditions using a Bayesian Wilcoxon signed rank test to establish whether participants' average tapping speeds differed between conditions. Next, we computed each

dyad's asynchronies as described in Chapter 3 and compared both the absolute unsigned asynchronies and the signed asynchronies between conditions using Bayesian Wilcoxon signed rank tests to test whether dyads were more closely coordinated in the synchrony compared to the asynchrony condition.

For the sliding task, we interpolated the data using the `approx()` method in R to achieve a consistent sampling rate of 50Hz (one entry every 20ms), and computed velocity and acceleration for each timestamp. We compared participants' mean velocity between conditions using a Bayesian Wilcoxon signed rank test to establish whether participants' average sliding speeds differed between conditions. We then mean-centered the data and computed half period normalized continuous phase values as for pilot 2 in Chapter 2. To obtain the relative phase angles of the two participants in each trial, we subtracted the signal of participant B from that of participant A for each timepoint when there was a signal from both participants. We then transformed the relative phase angles into vectors using the R package `circular` (Agostinelli & Lund, 2023) and computed mean vectors for each trial. We compared dyads' mean vector length between conditions using a Bayesian Wilcoxon signed rank test to test for a difference in the variability of relative phase between conditions.

Interview Data

The interviews were automatically transcribed using Amberscript software, then manually checked against the audio recording and corrected by the interviewer. The resulting transcripts were analyzed according to the guidelines for the analysis of micro-phenomenological interviews (Petitmengin et al., 2019; Valenzuela-Moguillansky & Vásquez-Rosati, 2019a).³⁵ First, descriptive utterances were extracted from the transcripts. That is,

³⁵ The interviewer completed the micro-phenomenological interview analysis training (<https://www.microphenomenology.com/analysis>, last accessed 2024-01-02).

utterances that contained descriptions of the particular experience which the interviewee was asked to describe were retained, whereas general statements, judgments, comments, and reflections were not included in the analysis.

The reliability of the extracted utterances was also evaluated with respect to preceding questions by the interviewer and with respect to verbal and paraverbal clues on the part of the interviewee. Responses to questions that were not content-neutral, in which the interviewee takes on the imposed content, were deemed unreliable. However, responses to such questions, which rejected the imposed content were deemed reliable and included in the analysis. Verbal clues on the part of the interviewee that were taken to be indicative of a reliable description of their experience included the use of short and simple language, present tense or simple past tense, and the description of concrete details; paraverbal clues to the reliability of the description included the slowing of speech, and the presence of silences and hesitations (see Petitmengin et al., 2019, pp. 698).

The remaining reliable utterances were then reorganized to follow the chronology of the experience and grouped into moments. The criteria for each grouping were made explicit, then the moments were further grouped into associations of utterances that form a descriptive category, or ‘diachronic unit’ (Valenzuela-Moguillansky & Vásquez-Rosati, 2019, p. 128). The resulting diachronic units then received labels and were organized into phases and sub-phases. The criteria for each phase and sub-phase were spelled out, and the specific diachronic structure of the experience described in each interview was visualized. The specific diachronic structures from each interview were subsequently compared – first within each dyad, then across dyads – in order to extract a generic diachronic structure of the experience of interpersonal coordination, which could then be compared between tasks and conditions. We did not perform a full analysis of all synchronic aspects of participants’ experience, as this was not the primary

focus of our investigation. We did, however, include noteworthy synchronic aspects in our analysis of the diachronic structure of participants' experience.

Results

Coordination Data

Discrete Coordination

Participants' mean ITIs in the tapping task did not substantially differ between the synchrony ($514 \pm 157\text{ms}$) and the asynchrony condition ($593 \pm 177\text{ms}$): $BF_{01} = 2.37$.³⁶ Their mean absolute asynchronies also did not substantially differ between the synchrony ($63 \pm 21\text{ms}$) and the asynchrony condition ($86 \pm 21\text{ms}$): $BF_{01} = 1.22$, nor did their mean signed asynchronies ($+8 \pm 46\text{ms}$ in the synchrony condition; $-7 \pm 13\text{ms}$ in the asynchrony condition): $BF_{01} = 1.68$ (see

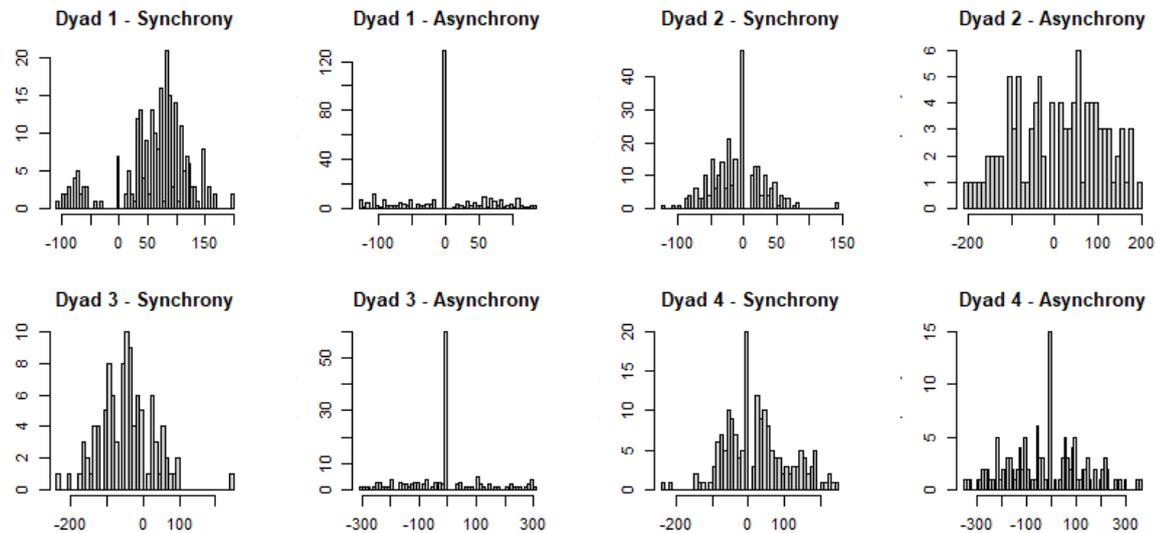
³⁶ For participants' mean ITI in each trial see Supplementary Table 4.

Figure 22).³⁷ Given the small samples sizes for these tests (8 participants for the analysis for ITIs, 4 dyads for the analysis of asynchronies), inconclusive results are unsurprising. Numerically, trials in the synchrony condition showed smaller absolute asynchronies than asynchronous trials. However, visual inspection of the data does not support the idea that participants' taps were more closely coordinated during synchronous compared to asynchronous trials of the tapping task.

³⁷ For participants' mean asynchronies in each trial see Supplementary Table 5.

Figure 22

Tapping Task: Signed Asynchronies in each Trial



Note. This figure depicts participants' signed asynchronies in each trial of the tapping task. Each histogram represents one trial. Labels above the histograms indicate which dyad and which condition the histogram depicts. Each histogram displays the frequency (y-axis) of participants' binned signed asynchronies (50 bins per trial) on the x-axis in ms. Positive asynchronies indicate that participant B's tap occurred later in time than participant A's respective tap. Negative asynchronies indicate that participant A's tap occurred after participant B's. Asynchronies around 0 indicate that participants' taps were closely coordinated.

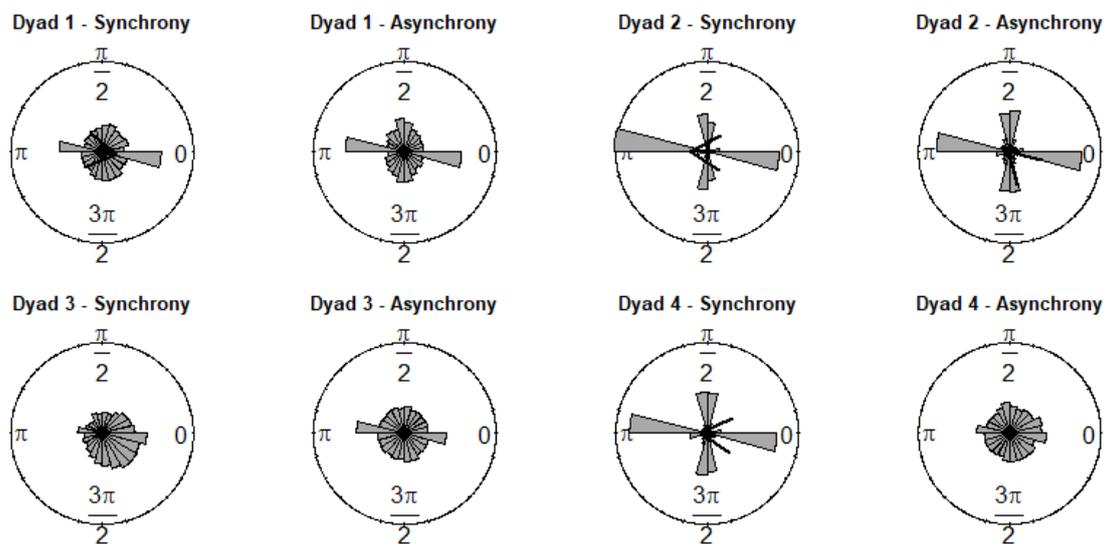
Continuous Coordination

Participants' mean velocity did not substantially differ between synchronous ($4 \pm 16\text{mm/s}$) and asynchronous coordination trials ($6 \pm 34\text{mm/s}$): $BF_{01} = 2.97$, and neither did the difference in velocity between members of the same dyad ($13 \pm 59\text{mm/s}$ in the synchrony

condition; $4 \pm 13\text{mm/s}$ in the asynchrony condition): $BF_{01} = 2.16$.³⁸ There was anecdotal evidence for a difference in mean vector length, however: $BF_{10} = 1.19$, with a greater mean value in the synchrony ($M = 0.19$) compared to the asynchrony condition ($M = 0.05$; see Figure 23). This suggests that participants were in a somewhat more stable phase relationship during synchronous coordination trials compared to asynchronous coordination trials, but that their velocity did not differ between conditions.

Figure 23

Sliding Task: Relative Phase in each Trial



Note. This figure depicts participants' relative phase angles in each trial of the sliding task. Each rose diagram represents one trial. Labels above the diagrams indicate which dyad and which condition the diagram depicts. Each rose diagram depicts relative phase angles in radians, ranging from 0 to 2π (360 degrees). A relative phase angle of 0 indicates in-phase synchrony. A relative phase angle of π (180 degrees) indicates anti-phase synchrony. The arrow

³⁸ For participants' mean velocity in each trial of the sliding task see Supplementary Table 6.

originating from the center of each diagram represents the mean vector (i.e., the mean relative phase angle) and its length. The longer the arrow, the more stable the phase relationship.

Interview Data

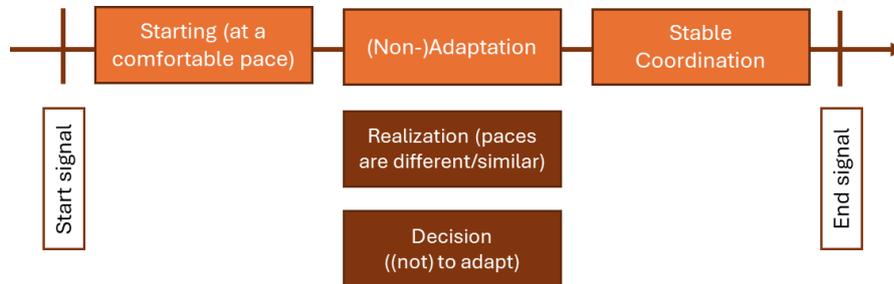
In our analysis of the specific diachronic structures of participants' experience of interpersonal coordination, we were able to extract one key generic structure, encompassing three phases of the experience, as well as some structural variations. We also found differences between individuals, as well as some variation between tasks and conditions.

Generic Diachronic Structure

The generic diachronic analysis of four dyads' experience of interpersonal coordination under different conditions and in different tasks allowed us to identify one key generic structure of the experience of interpersonal coordination. This generic structure describes three phases of the experience of interpersonal coordination: an initial phase of 'starting (at a comfortable pace)' (phase 1), followed by a phase of '(non-)adaptation' (phase 2), and a phase of 'stable coordination' (phase 3). Phase 2 was further characterized by two sub-phases: 'realization (that paces are different/similar)' (phase 2.1), and 'decision ((not) to adapt)' (phase 2.2; see Figure 24).

Figure 24

Generic Diachronic Structure of the Experience of Interpersonal Coordination



Note. This figure depicts the generic diachronic structure of the experience of discrete asynchronous coordination. The different phases of the experience are depicted in order of occurrence from left to right, along the brown arrow. Brown vertical lines mark the occurrence of the start and end signal. Each orange box on the brown arrow represents one phase of the experience. Brown boxes under the arrow represent sub-phases.

Phase 1: Starting

The phase ‘starting (at a comfortable pace)’ is characterized by participants’ experience of starting to move after the start signal occurs. Participants often described not giving their initial tempo much thought, stating, “I just started without thinking much about that pace” (1B4, line 350), or “I just started clicking. I wasn’t really, uh, thinking with about a speed to go.” (2A2, line 63). In some cases, however, participants had already thought about their initial tempo before starting and began moving with a specific tempo in mind – e.g., “my plan was to do the same as yesterday and start as slow as possible” (2B2, line 14).

Participants often described their initial pace as comfortable – e.g., “I figured out I’ll grab some pace, which is good for me, which is comfortable.” (1A4, line 14) – or natural – e.g., “I chose the tempo that would feel natural to me” (3A1, line 26). In some cases, it was also accompanied by an imaginary metronome – e.g., “I generated a metronome sound in my

head. Uh-Uh, so that I wouldn't fall. Um, I wouldn't lose the tempo. So I used that to, um, continue on the tempo” (3A2, line 120) – or counting – e.g., “at some point I was counting like [taps] one [tap] one [tap] one [tap] one” (1A4, line 171).

Phase 2: (Non-)Adaptation

The phase ‘(non-)adaptation’ is characterized by participants’ experience of adjusting their movements to their task partner’s pace – e.g., “I was following her pace” (1A1, line 115) – or by their experience of intentionally not adjusting their movements to their task partner’s pace – e.g., “I didn't get, uh, slower. I just kept the pace” (3A1, line 192). This phase was further characterized by two sub-phases: ‘realization (that paces are different/similar)’ and ‘decision ((not) to adapt)’.

The sub-phase ‘realization (that paces are different/similar)’ is characterized by participants’ experience of noticing that their own pace was different from their task partner’s (when they were instructed to move at the same tempo) – e.g., “I noticed that she was much more slower than I was” (3A1, line 192) – or that it was similar to their task partner’s (when they were instructed to move at different tempos) – e.g., “I realized that we were at a similar pace at some point” (1A4, line 26). It typically occurred via seeing – e.g., “when I saw that [gestures] this one is my literally my blink stops and then her begins and my stops and her begins [stops gesturing] I realized it's not even close” (1A2, line 325) – or hearing – e.g., “at some point I noticed that the strong gliding sound she was making matched my, um, match, my metronome in my head” (3A4, line 280).

The sub-phase ‘decision ((not) to adapt)’ is characterized by participants’ experience of adapting their movements, or trying to adapt their movements, to their task partner’s pace – e.g., “I tried to adjust to her speed and when she would go slower, try to go faster, vice versa” (2B1, line 40) – or choosing not to adapt their movements to their task partner’s pace – e.g., “I

chose the tempo that would feel natural to me, and I tapped it and expected the other one to comply to it” (3A1, line 26). Some participants described this to be a decision – e.g., “I decided to go very fast [because it was as different as possible from what my partner was doing]” (2B1, line 311). Others only described their actions – e.g., “I was tapping and then I was also looking at my partner's, uh, way of tapping. [...] Um, and then accordingly, accordingly, I would, uh, um, arrange myself within this” (4A1, line 143) – or attitudes – e.g., “I didn't really care. I kept-I kept with my pace.” (1A4, line 18).

The two sub-phases typically occurred in this order. However, some participants described their decision to adapt or not to adapt their movements to their task partner's pace at the beginning or even before the start of the task – e.g., “I wasn't planning to adapt” (1A2, line 264). In these cases, the sub-phase ‘realization (that paces are different/similar)’ either occurred after the sub-phase ‘decision ((not) to adapt)’ or did not occur at all.

Phase 3: Stable Coordination

The phase ‘stable coordination’ is characterized by participants’ experience of keeping a stable pace (stable intrapersonal coordination) – e.g., “I tried to, um, keep my tempo stable and my hand movement, um, as uniform as possible” (3A2, line 19) – or of keeping a consistent relationship between their own pace and their task partner's pace (stable interpersonal coordination) – e.g., “we had a pretty constant tempo” (2B4, line 138).

This phase often included an aspect of timekeeping. For some, this took the form of an imagined metronome – e.g., “I didn't generate a metronome sound in the actual world, but in my head I was I tried to hear the metronome in my head” (3A2, line 79). For others, it was bodily – e.g., “I tried to keep a constant tempo by um. Also how to say imitating the rhythm with my foot” (2B2, line 155). And for some, it was by counting – e.g., “I was counting the seconds” (2A2, line 204).

Some participants also described an aspect of sensory exploration during this phase, commenting on the appearance of the visual feedback from the button in the tapping task: “I realized that it's not even blinking when I when I'm clicking that fast” (1A4, line 22) or on additional auditory feedback provided by the sound of the finger touching the screen: “I realized that I not only liked the pace like it was physically comfortable for my hand, but also the, um. The sound like the rhythm” (1B4, line 227). In some cases, this also helped participants keep a stable tempo – e.g., “it was more like, um, enabling me to feel the time.” (1B4, line 265).

Structural Variations

Absence of Phase 1: Starting

For some participants, the initial phase of ‘starting (at a comfortable pace)’ was not present in their (description of the) experience. These participants did not describe starting at their own pace, but rather immediately started moving in accordance with their task partner – e.g., “completely unconsciously, I started following him” (2A4, line 38), or “from the very beginning, I tried to just, you know, um, go with her pace” (3B3, line 187).

Absence of Phase 2: (Non-)Adaptation

The phase of ‘(non-)adaptation’ was not present in all descriptions. Some participants described immediately starting at different paces (in the asynchrony condition) and therefore not needing to adapt to one another – e.g., “I figure out immediately that I would go so slow and that that's going to be just enough not to be at the same pace” (1A3, line 14). In some dyads, this was mutual, “intuitively I started it at a high pace and then I just kept doing that” (1B3, line 141), with neither member describing a phase of (non-)adaptation in their experience.

In others, only one member of the dyad described the experience in these terms – e.g., “Just find my pace. Um, and the other one will go at a different pace anyway.” (3A3, line 13)

or “I knew that he's going to do that. That's why I chose a different strategy.” (4B2, line 39). For one participant, this also occurred in the synchrony condition: “when I first heard the sound, I again, again felt, a natural tempo and I would go with it. [...] I saw that. She was approximately, um, at the pace that I was [...] I kept doing this movement. Uh huh. As smoothly as possible” (3A2, line 19).

Absence of Phase 3: Stable Coordination

Some participants did not describe a phase of stable coordination in their experience. In these cases, the phase of (non-)adaptation was often accompanied by some discomfort – e.g., “it was a little bit uncomfortable, but it's more like an emotional uncomfort” (3B3, line 146) – or unpleasant feelings – e.g., “she made changes, but I didn't tend [stand?] to, Um, adjust myself to, to those changes. Uh, yes. I, I felt not that good” (4A4, line 38). One dyad described it as a kind of missing – e.g., “it was like, for example, the feeling as if you're trying to throw a ball into a basket or like, reach the point, a certain point, but you're not, uh, like you're always missing. You come close, but you're always missing” (1B2, 170) or “I was like, okay, we are missing each other for from the very single second. Like it's blinking like [gestures] so so that's how I realized that we are not touching that, you just literally can I was able to see that my blinks and shots and then her blinks and shots. So like, you know, it's like we are not touching it at the same time.” (1A2, line 55).

Additional Phase of Experimenting

In addition to the phases described above, some participants' descriptions of their experience also included a phase of experimenting. This phase was characterized by participants' experience of trying out different kinds of movements, tempos, or rhythms to see what would happen, or how their partner would respond. This phase often included an aspect of playfulness – e.g., “I just playfully changed it and I realized that he is actually paying more

attention. So I more confidently [laughs] changed it without really paying much attention.” (2A1, line 210). Participants also described an aspect of reciprocity – e.g., “I think there was some music in the tempo that we took. So I tried to adjust it a little bit so that it reminded me of a, uh, a musical rhythm. And then I think my partner picked up on that pretty easily, and we started to have some jamming sessions” (2B4, line 128) or “he kind of did my movement, and I had to change it right away. [...] And then this thing happened again [...] And then I had to borrow what he was doing [laughs]” (4B1, line 50). However, the aspect of reciprocity was not always present – e.g., “I switched the movements and I took the lead and, uh, started to speed. [...] speed it up. And then I noticed that my partner is not following me” (4B3, line 28).

Inter-Individual Variation

Adaptation vs. Nonadaptation

The phase ‘(non-)adaptation’ comprised experiences of mutual adaptation, one-sided adaption, and non-adaptation. These different sub-categories were present to varying degrees in individual participants’ experiential reports.

One dyad primarily showed mutual adaption – e.g., “we were both trying to, um. Trying to adjust to the other person's tempo” (2B4, line 62). This was linked to different feelings about the other person’s pace – e.g., “I was happy that I felt like he's speeding up” (2A3, line 213) or “I was surprised that she took the exact same tempo that I took yesterday” (2B2, line 113) – and about their own pace – e.g., “I regretted why I started so slow” (2A3, line 13). This dyad’s experiential reports most often included a phase of stable coordination (7 out of 8 interviews).

In the remaining three dyads, one-sided adaptation was equally or more common than mutual adaptation, though to varying degrees. Within each of these dyads, there was one member who typically tried to adapt their movements to the other person’s pace (adapters), and

one member who tried not to adapt and instead to remain stable in their own movements (nonadapters).³⁹

Two nonadapters described their task partner's movements as less stable as their own. One participant described his task partner's movements saying, "there were oscillations in her pace" (1A1, line 136), and "I saw her finger like, I don't know how to call that just it's her finger was not moving the same all the time" (1A2, line 59). This participant tried to adapt to his task partner at first, then took on a nonadaptive stance. The other nonadapter chose not to adapt to her task partner's pace from the start, stating, "if I, uh, try to adapt to her tempo, then, um, it wouldn't work, because if I try to adapt hers, maybe she has a weaker sense of tempo. So she will go wayward, and I will go [small laugh] wayward too" (3A1, 79). Instead, this participant tried to keep her own tempo stable and waited for her task partner to adapt.

Both of these nonadapters' task partners tried to adapt their movements to their task partner's pace but described this as difficult. One participant described, "I also couldn't make myself slower, um, because, um, I saw that trying to adjust that pace will make me stop." (1B1, line 224), "it's very hard to regulate" (1B2, line 262), and "it seemed that my, for my partner, it was um, it took less effort to move that circle." (1B1, line 264). The other similarly described, "it was a little bit hard for me to just move along" (3B2, line 14), and "I was more focused on myself because I was, you know, like, uh, experiencing kind of like a problem. Uh, but at the same time, like, I was also, um, seeing her circle, and it was kind of moving, like, smoothly" (3B2, line 47). The second adapter also described some emotional discomfort, describing, "I felt like I had no choice" (3B3, line 104), and "I felt like she's just, uh, like, she doesn't care about the task" (3B3, line 220).

³⁹ All nonadapters indicated that they had experience with making music or playing an instrument, whereas all adapters indicated that they did not, suggesting that this may be related to music experience.

The third nonadapter approached the coordination tasks in a different way, stating, “I perceived it as a strategic game” (4A2, line 20), “like a competition” (4A2, line 38), and “my kind of purpose was to get the floor under my influence” (4A2, line 48). Whereas he adapted to his task partner in most of the discrete coordination task, his stance was different in the continuous coordination task, in which he states, “I felt taking the control over the game” (4A4, line 38). In these sessions, this participant decides not to adapt to his task partner, despite describing, “I felt not that good” (4A4, line 38). His task partner, meanwhile, describes, “I took a lead and I changed it, but he didn't follow my paces [...] I saw that he didn't do anything and then just switched to what he was doing again” (4B4, line 30).

Experimenting vs. Exploring

A phase of experimenting was present in two dyads' experience but was fully absent in both other dyads' experiential reports. The dyads who described a phase of experimenting were also those who showed more mutual adaptation. By contrast, members of the other two dyads who reached a phase of stable (intrapersonal) coordination, described an aspect of exploration, which was not present in experimenting dyads' reports.

For one of the experimenting dyads, members of the dyad took turns experimenting, with the other member following the experimenting member's pace. Each member reported experimenting in one session – “I realized that, okay, he's taking care of this [laughs]. So I can just go with my own experiment to see, uh, which is really is more fun for example” (2A1, line 338) and “once I started to make musical rhythms, I had other musical rhythms popping into my mind. And then I wanted to try those” (2B4, line 138). The other member, respectively, described how they tried to adapt – “I tried to adjust to her speed and when she would go slower, try to go faster, vice versa.” (2B1, line 40) and “He's doing a rhythm like. [...] I had to uh recognize the rhythm and then follow it” (2A4, line 108).

The other experimenting dyad showed a different pattern of experience. One member of this dyad reported a phase of experimenting in her experience of all sessions of the study – e.g., “I was just doing what, what's in my head and like oh what if I just start really slowly and then moving slowly upwards and then backwards” (4B2, line 45). The other member of this dyad only did so for the first session – “I wanted to make a difference, to make some changes. And then I, no matter what kind of I intervened and then kind of kept doing, uh, differently, kind of changed my speed” (4A1, line 172). He then took on a stance of nonadaptation with respect to his task partner’s experimenting for the following sessions.

Dyads who did not report a phase of experimenting in their experience instead described an aspect of sensory exploration during stable (intrapersonal) coordination. This aspect was closely related to the qualities or effects of their own movements. Some participants described it in relation to the visual effects of their movements – e.g., “I was looking at both of our circles' pace, and then I noticed that they leave such a footprint” (1B1, line 280). Others described it with respect to the auditory qualities of the movement – e.g., “I made uh like pronounced sounds with my fingers” (3A3, line 13) or “I tried to hear the metronome in my head, um, and I also barely, um, heard the sound, the really light sound that my fingers were making” (3A2, line 79). One member of these dyads did not describe a phase of stable coordination in any of the interviews and did not report an aspect of sensory exploration.

Differences between Tasks and Conditions

To see whether there are differences in the experiences of interpersonal coordination between the different types of tasks and the different conditions used in this study, we summarized the occurrence and non-occurrence of the different phases for each participant in each session. The generic structure was most prevalent in the context of discrete asynchronous coordination (see Table 2), described by half of participants. Discrete synchronous

coordination was most often experienced without a phase of starting (50% of participants) but included a phase of (non-)adaptation for all participants, as well as an additional phase of (non-)adaptation for two participants in response to their task partner's experimenting (see Table 3).

Table 2

Discrete Asynchronous Coordination

Dyad	Member	Starting	(Non-) Adaptation	Stable Coordination	Experimenting
1	A	✓	✓	✓	
	B	✓	✓	✓	
2	A	✓	✓	✓	
	B	✓	✓	✓	
3	A	✓		✓	
	B		✓		
4	A	✓	✓	✓	✓
	B	✓	✓		✓

Note. This table illustrates whether participants' descriptions of their experience of discrete asynchronous coordination included the following phases: starting (column 3), a (non-) adaptation (column 4), a stable coordination (column 5), and experimenting (column 6).

Table 3*Discrete Synchronous Coordination*

Dyad	Member	Starting	(Non-) Adaptation	Stable Coordination	Experimenting
1	A	✓	✓		
	B		✓		
2	A		✓	✓	(non-) adaptation
	B		✓	✓	✓
3	A	✓	✓	✓	
	B		✓		
4	A	✓	✓	✓	(non-) adaptation
	B	✓	✓	✓	✓

Note. This table illustrates whether participants' descriptions of their experience of discrete synchronous coordination included the following phases: starting (column 3), a (non-) adaptation (column 4), a stable coordination (column 5), and experimenting (column 6).

Continuous asynchronous coordination most often lacked a phase of (non-)adaptation (3 out of 8 participants), and equally often lacked a phase of starting and of stable coordination (see Table 4). Continuous synchronous coordination most often lacked a phase of stable coordination (for 62.5% of participants) and had only one report of experimenting (see

Table 5). Numerically, the generic structure was less frequently present in participants' experience of the continuous coordination task (twice) than in their experience of the discrete coordination task (five times, or eight times if experiences that include a phase of experimenting are included too). Only one description of synchronous coordination did not include a phase of (non-)adaptation, whereas four participants reported on an experience of asynchronous coordination that did not include a phase of (non-)adaptation.

Table 4

Continuous Asynchronous Coordination

Dyad	Member	Starting	(Non-) Adaptation	Stable Coordination	Experimenting
1	A	✓		✓	
	B	✓		✓	
2	A	✓	✓	✓	✓
	B	✓	✓	✓	
3	A	✓	✓		
	B		✓		
4	A		✓		
	B	✓		✓	✓

Note. This table illustrates whether participants' descriptions of their experience of continuous asynchronous coordination included the following phases: starting (column 3), a (non-) adaptation (column 4), a stable coordination (column 5), and experimenting (column 6).

Table 5*Continuous Synchronous Coordination*

Dyad	Member	Starting	(Non-) Adaptation	Stable Coordination	Experimenting
1	A	✓	✓		
	B		✓	✓	
2	A	✓	✓		
	B	✓	✓	✓	
3	A	✓		✓	
	B	✓	✓		
4	A	✓	✓		
	B	✓	✓		✓

Note. This table illustrates whether participants' descriptions of their experience of continuous synchronous coordination included the following phases: starting (column 3), a (non-) adaptation (column 4), a stable coordination (column 5), and experimenting (column 6).

Regarding participants' adaptation, one participant (4A) showed a stronger tendency not to adapt to his task partner in the continuous coordination task. However, another participant (1A) showed the opposite pattern, suggesting that this is not a structural difference related to features of the task. Similarly, the frequency of reports of a phase of experimenting, as well as the presence of an aspect of exploration, was similar across tasks and conditions.

Relationship between Behavioral and Experiential Data

To gain a better understanding of the relationship between participants' experiential reports and their behavior during the tasks, we additionally report our descriptive observations

of the behavioral data in sessions when participants indicated that they experienced a stable coordination phase and in sessions when they did not.

Synchronous Coordination

In the context of synchronous coordination, there were two sessions in which neither member of the dyad reported experiencing a phase of stable coordination. One of these sessions involved discrete coordination and was reflected in the dyad's asynchronies, which were not centered around 0ms, but rather showed a distribution with the highest peak between 50 and 100ms (see

Figure 22, dyad 1), suggesting that their taps were not well-synchronized. The other session involved continuous coordination and does not match the behavioral data very well, as the distribution of relative phase angles shows that this dyad spent the majority of the trial in a relative phase relationship of either 0 or 180 degrees (see Figure 23, dyad 4), which are considered the most stable phase relationships. By contrast, there were also two sessions of synchronous coordination in which both members of the dyad reported experiencing a phase of stable coordination. Both of these sessions involved discrete coordination and were reflected in the behavioral data to the extent that the distribution of asynchronies was centered around 0ms (see

Figure 22, dyads 2 and 4).

Asynchronous Coordination

In the context of asynchronous coordination, there was only one session, in which neither member of the dyad reported experiencing a phase of stable coordination. This session involved continuous coordination, and the distribution of relative phase angles is indeed spread out across the entire circle, with only small peaks around 0 and 180 degrees (see Figure 23, dyad 3). By contrast, there were four sessions of asynchronous coordination, in which both members of a dyad described experiencing a phase of stable coordination. Two of these sessions involved discrete coordination; two involved continuous coordination; and all sessions were reported by the same two dyads.

In the context of discrete coordination, both members of dyad 1 reported stable intrapersonal coordination, with both participants experiencing themselves as going at a stable pace, different from the other person's. This is supported by the large difference between their mean ITIs (more than 180ms; see Supplementary Table 4). The distribution of their asynchronies, however, shows that their taps were quite closely coordinated in time, with a sharp spike at 0ms (see

Figure 22). Both members of dyad 2 also reported stable intrapersonal coordination in this task, though participant A also reported stable interpersonal coordination in that they were successfully tapping at different tempos, and thereby fulfilling the task. This is reflected in their behavioral data, with the distribution of asynchronies being relatively wide and having no sharp peak (see

Figure 22).

In the context of continuous coordination, both members of dyad 1 also described stable intrapersonal coordination. This is reflected in the numerically greater difference between their mean velocities in the asynchronous condition of the sliding task compared to the synchronous condition of the sliding task (see Supplementary Table 6). Visual inspection of the distribution of relative phase angles during this trial compared to continuous synchronous coordination (see Figure 23), however, suggests that their coordination was similarly stable during both conditions. Both members of dyad 2 reported a phase of stable interpersonal coordination, in which they were moving at different paces, fulfilling the task instruction. The relatively small difference between their mean velocities, similar to the difference in the synchrony condition (see Supplementary Table 6), as well as visual inspection of the distribution of relative phase angles (see Figure 23), however, suggest that their velocities were not substantially different and that their coordination was relatively stable.

Discussion

The present chapter set out to investigate the experience of interpersonal synchrony in continuous and discrete coordination contexts by comparing it to the context of asynchronous movement coordination. To do so, we invited four pairs of participants to the lab to take part in four sessions, in which they completed a short coordination task before being interviewed about their experience during the task. We found that participants' experience of interpersonal coordination could be described with a generic structure made up of three phases: an initial phase of 'starting (at a comfortable pace)', a phase of '(non-)adaptation', and a phase of 'stable coordination'. We found that some participants' experiential reports lacked descriptions of one or more of these phases, and that some participants also described an additional phase of 'experimenting'. We also found some variation between individuals, with participants differing

in their approaches to adaptation or non-adaptation and experimenting as opposed to exploring. Finally, we aimed to compare participants' structure of experience between tasks and conditions, and to compare their experiential reports to the recorded behavioral data, with mixed results.

In contrast to the studies presented in Chapters 2 and 3, the tasks used in the present chapter did not include a metronome at the start of the trial to pace participants and instead gave participants the freedom to choose their own pace. We opted for this approach after pilot testing a version which included a metronome, as pilot participants' experiential reports were largely centered around the strategies they used to try to keep the tempo and contained much less information about participants' attempts to coordinate with their task partner. While this may have been informative with respect to gaining a better understanding of participants' experience of the particular tasks we used in our studies, we reasoned that it would be more informative with respect to the underlying processes of interpersonal coordination to use a task in which participants were forced to coordinate with one another and did not have the option to simply rely on an external reference instead.

Interestingly, despite our efforts to force participants to coordinate with one another, we found that most participants reported on some kind of timekeeping mechanism – counting (participants 1A and 2A), generating an imaginary metronome (participant 3A), tapping with their foot (participant 2B), or using external sounds (participant 1B), consistent with the alignment of internal models proposed to underlie interpersonal synchronization (Repp & Su, 2013). Moreover, we found that all participants reported focusing only on themselves or their own movements at some point during the study. This strong focus on keeping a tempo, albeit arbitrary, was also reflected in participants' limited willingness to adapt to their task partner,

with three participants taking on a relatively consistent stance of nonadaptation, opting instead to move at a stable pace and expecting their partner to adapt.

While this stance may likely be a byproduct of our design, which involved repeated interactions with the same person, it provides support for the previously documented pattern of unidirectional adaptation, or distribution of distinct roles in the context of joint actions (Richardson et al., 2015). Moreover it extends findings on the effectiveness of non-arbitrary leader-follower role distributions in the context of the availability of (perceptual) information (Curioni et al., 2019) by providing an indication that competence may play a similar role. Two of the three participants who chose not to adapt to their task partner appeared to do so due to a perceived lack of competence on their task partner's behalf. Moreover, linking their adaptation behavior to their questionnaire responses with respect to their experience with music, we found that all three non-adapters had experience with music, whereas all three adapters did not.

However, our sample also included one dyad who primarily showed mutual adaptation. This, too, has been shown to lead to successful coordination (Curioni et al., 2019), particularly in experts, who thereby enter a state of co-leadership (Noy et al., 2011). Mutual adaptation has been found to lead to better performance than arbitrarily assigned leader-follower roles (Noy et al., 2011; Curioni et al., 2019; Varlet et al., 2020). Unlike the other dyads, in which only one member indicated that they had experience with music, both members of dyad 2, who showed mutual adaptation, indicated that they had at least some experience with music. While this does not render them experts, it aligns with previous findings that temporal prediction abilities, which may be linked to musical training, are linked to synchronization performance (Pecenka & Keller, 2011).

It is also worth noting that this dyad described more playfulness and responsiveness to the other member's behavior in their reports, and that in subsequent phases of experimentation

they also described moving at different rhythms and having jam sessions, possibly akin to the contexts of joint improvisation described in the above mentioned studies. Interestingly, the dyad who primarily showed mutual adaptation was also one of the two dyads who described an additional phase of experimenting when reporting on their experience. As the other dyad who reported a phase of experimenting also reported on experiences of mutual adaptation, this may suggest that not only attention to the other person but also reciprocity, or mutual adaptation, are necessary prerequisites to engage in ‘experimenting’, or joint improvisation (see Noy et al., 2011; Varlet et al., 2020; Zhai et al., 2016).

Our data also suggests that there may be individual differences with respect to the different phases of the experience. With respect to the phase of experimenting, there was one participant who reported the presence of such a phase in each of her experiential reports. By stark contrast, other participants never reported the presence of such a phase, with one participant also failing to report on a phase of stable coordination in any of her reports. Interestingly, this participant also failed to report on a phase of starting in three of the four sessions, and primarily described her efforts to adapt to her task partner during the interviews. This suggests the presence of different coordination strategies across different individuals, which may be linked to particular individual characteristics or personality traits, such as internal or external locus of control (Fairhurst et al., 2014; Washburn et al., 2021), but may also be related to the particular combination of individuals paired to complete the task together.

Interestingly, participants’ experiential reports primarily focused on descriptions of the actions participants performed to coordinate their movements. The interviews also included descriptions of the relation between their paces, as well as descriptions of the quality of the movement and positive responses to it, particularly during phases of stable coordination or experimenting. However, participants’ experiential reports typically did not feature

descriptions of their relation with the other person or any change in attitude towards them. This is particularly interesting in relation to findings on the effect of interpersonal synchrony on people's affective responses, which have been documented in different 'real-life' contexts (e.g., Fischer et al., 2014), but often could not be replicated in lab settings (e.g., Reddish et al., 2013), suggesting that some effects of interpersonal synchrony may also be determined by the overall context or setting of the interaction and not simply the relation between one's own and others' movements.

It is worth noting with respect to the behavioral data that without the presence of a metronome to guide their tempo, as in Chapter 2, or the additional performance criteria of Chapter 3, participants' movements were not substantially more coordinated in the synchrony condition compared to the asynchrony condition. The analysis of discrete coordination showed anecdotal evidence for no difference between conditions. While the analysis of continuous coordination showed anecdotal evidence for a difference between conditions, it was not substantial. This is noteworthy despite the small sample size given that the experiments described in previous chapters found extreme evidence for a difference between participants' synchronous and asynchronous coordination. This suggests that, by comparison, our manipulation in the present study was not successful in eliciting more closely coordinated movements in the synchrony condition than in the asynchrony condition.

Upon visual inspection, this is particularly salient for dyads 1 and 3 in the tapping task, as both of these dyads showed a large proportion of taps with an asynchrony close to 0ms in the asynchrony condition, visible as a spike around 0 in the histogram (see Figure 2), whereas their asynchronies were more spread out in the synchrony condition (and the highest peak in the histogram was not around 0ms). In fact, only dyad 2 showed visibly more coordinated tapping in the synchrony condition than in the asynchrony condition, with a clear peak around

0ms in the synchrony condition and relatively evenly spread out asynchronies in the asynchrony condition. That being said, dyad 4 also showed more closely coordinated taps in the synchrony condition, based on visual inspection, but the peak around 0ms was visible in the histogram for both conditions.

Interestingly, this was not the case in the sliding task. Instead, visual inspection here suggests that participants' coordination patterns were relatively similar across conditions for three of the four dyads. Only dyad 4 showed visibly more stable coordination in the synchrony condition, spending most of the trial in stable in- or anti-phase coordination (0 or 180 degrees, respectively; see Figure 3), than in the asynchrony condition, with relative phase angles relatively evenly distributed across the full 360 degree circle. This suggests that although participants did not succeed overall in generating more closely coordinated movements in the synchrony condition compared to the asynchrony condition in either task, the continuous coordination task may have been easier for participants to complete.⁴⁰

This may be related to the fact that participants did not receive additional auditory cues beyond those they generated by performing the movements on the screen (i.e., there were no auditory sound effects as in Chapter 3) to support their coordination. Reliance on visual cues alone may have supported successful coordination in the continuous coordination task, with some participants making reference to the ends of the track functioning as reference points. However, as some participants pointed out, this was more challenging in the discrete coordination task, with participants resorting to auditory or even bodily cues. This may be due to the better temporal resolution of auditory compared to visual information (Vroomen & Keetels, 2010), which may have supported more precise coordination, particularly in the discrete task.

⁴⁰ This is consistent with participants' ratings of task difficulty in Chapter 2.

However, it is worth noting that our instructions to participants emphasized tempo, not the stability of their coordination. As participants were not given a tempo to follow and were only instructed to move at the same tempo (in the synchrony condition) or at different tempos (in the asynchrony condition), different tempos could also have been achieved by one person going at twice the speed as the other person. The latter approach would look like a stable coordination pattern in participants' asynchronies or mean relative phase, but it would also fulfill the instructions participants were given for the asynchrony condition. This discrepancy becomes apparent in the case of discrete asynchronous coordination for dyad 1. Whereas descriptively, there was a large difference in their ITIs, suggesting that they were tapping at very different tempos, the distribution of their asynchronies showed a large spike around 0ms, suggesting that many of their taps were perfectly synchronized. This may suggest that these participants settled on a coordination strategy in which one participant taps at double the speed of the other, resulting in half of their taps being perfectly coordinated (with an asynchrony of 0ms).

However, if this was the case one would expect to see a bimodal distribution of asynchronies, with half of the asynchronies centered around 0ms, and the other half centered around the faster participants' tapping speed (e.g., 200ms). As in this particular case, as well as the cases of other dyads who showed a sharp spike around 0ms, the remaining asynchronies were relatively evenly distributed (or did not show a clear second peak), it is not likely that this would have been the case. Instead, these distributions suggest that participants spontaneously synchronized their taps, similar to the behavior described by Richardson et al. (2007), despite being instructed to tap at different paces, and that the sharp peak at 0ms for three out of four dyads was not the result of a stable coordination pattern that fulfilled the instructions but would not be considered interpersonal asynchrony.

The lack of a clear difference between the coordination of participants' movements in the synchrony compared to the asynchrony condition, however, did not exclusively stem from participants' difficulty completing the asynchrony condition. Participants also struggled with the synchrony condition. This is particularly noteworthy in the case of discrete coordination, where participants show a strong tendency to coordinate their movements in the asynchrony condition, but seemed to struggle to do so in the synchrony condition. In particular, only two dyads' asynchronies were centered around a peak at 0ms in the synchrony condition, with both other dyads showing more skewed distributions. It is possible that the intention to synchronize, combined with sub-optimal cues for synchronization in this task, may ultimately have been detrimental to dyads' coordination success.

Interestingly, participants' experience of the discrete coordination task suggests that most participants did in fact experience a phase of stable coordination in the synchrony condition of this task. Contrary to the behavioral data for continuous coordination, this was not the case in the discrete task, with only three of eight participants reporting on a phase of stable coordination in their experience of synchronous coordination. Most surprisingly, neither member of the dyad with the most visible difference in coordination between the synchrony and asynchrony condition reported on a phase of stable coordination experience, despite visual inspection of the data indicating relatively stable coordination overall (see Figure 3, dyad 4).

While this appears to be at odds with the behavioral data, it is possible that what this reflects is an extended process of (successful) adaptation in the case of continuous synchronous coordination, without participants reaching a point at which their coordination felt stable. It is also possible that participants were not successful in reaching their desired level of coordination, and therefore did not report on a phase of stable coordination, but that their coordination remained relatively stable nonetheless. Similarly, it is equally possible that other

aspects of the experience were simply more salient to participants and overshadowed their experience of the degree of coordination between their movements.

By contrast, participants' experiential reports on discrete synchronous coordination matched the behavioral data quite well. Four of the five participants who reported experiencing a phase of stable coordination were members of the two dyads whose asynchronies indeed showed a clear peak around 0ms. The fifth participant, however, reported experiencing a phase of stable interpersonal coordination, despite their behavioral data suggesting that this phase would likely have involved relatively imprecise temporal coordination. The fact that the other member of this dyad did not report on a phase of stable interpersonal coordination, however, suggests that the same interaction may be experienced differently by each respective member of the interaction. Together with participants' reports of their experience of continuous synchronous coordination, this suggests that measured indices of coordination need not always line up with participants' experience of the interaction.

It is also worth noting that the lack of detailed descriptions of participants' relationship with their task partner and the discrepancies between the measured behavioral data and the extracted descriptions of subjective experience may differ due to methodological limitations of the interview procedure. Although the interviewer was trained both with respect to the micro-phenomenological interview technique and the micro-phenomenological analysis procedures, circumstantial factors could have allowed for some amount of bias to be retained and for certain details to be missing in the experiential reports. First, as the interviewer was present during the experience which the interviews were centered on, it may have been more difficult for the interviewer to put aside her own biased view of the situation (and sequence of events), and for the interviewee to share parts of their subjective experience, as the interviewer also featured in the experience. This may also have led to the omission of details that were regarded as common

knowledge due to the interviewer's presence during the experience, despite the interviewer's efforts to ask seemingly redundant questions. Second, the fact that participants were both interviewed by the same person (and were paired for all four sessions of the study) may have prompted interviewees to experience greater inhibitions about sharing details of their experience that pertained to the other person, as they were aware that they would have repeated interactions with the person in question (and that the interviewer would also be part of these interactions). Third, the lack of video recordings for most sessions meant that the evaluation of the reliability of the interviewees' utterances was restricted to verbal and paraverbal clues, and often could not take nonverbal clues, such as gaze shifts or the presence of gestures, into account. Finally, the analyses were conducted by the interviewer, with some initial support from a fellow researcher for the first set of interviews but no additional set of eyes of the final set of analyses, suggesting that some of her own biases will likely have gone unnoticed throughout the entire research process.

Overall, the present chapter demonstrates that despite struggling to synchronize their movements effectively in the synchrony condition and spontaneously falling into the same pace in the asynchrony condition without a metronome to provide a reference tempo, participants tried to coordinate their movements using different coordination strategies. Some participants opted for unidirectional adaptation, others mutually adapted to one another; some participants showed phases of sensory exploration, often with an aspect of timekeeping or trying to sustain interest in the task, while others opted to experiment, trying out different tempos and seeing how their task partner would respond. While their experiential reports did not always perfectly line up with the measured behavioral data, suggesting that objective and subjective measures of coordination may not always go hand in hand, the present findings contribute towards a better understanding of the experience of interpersonal coordination.

Supplementary Information

Results

Coordination Data

Supplementary Table 4

Tapping Task: Mean ITI in each Trial

Dyad	Member	ITI (M ± SD)	
		Synchrony Condition	Asynchrony Condition
1	A	379 ± 93.2ms	295 ± 255ms
	B	404 ± 48.8ms	477 ± 174ms
2	A	342 ± 88.3ms	910 ± 108ms
	B	336 ± 79.8ms	366 ± 92.7ms
3	A	863 ± 67.9ms	611 ± 36.1ms
	B	860 ± 80.4ms	991 ± 148ms
4	A	510 ± 398ms	710 ± 316ms
	B	420 ± 402ms	383 ± 286ms

Note. This table summarizes the ITIs (mean ± standard deviation) in each trial of the tapping task. The data are organized by dyad (column 1), member of the dyad (participant A or B; column 2), and condition (synchrony or asynchrony; columns 3 and 4).

Supplementary Table 5

Tapping Task: Mean Asynchrony in each Trial

Dyad	Type of Asynchrony	Asynchrony (M ± SD)	
		Synchrony Condition	Asynchrony Condition
1	Absolute	77.2 ± 37.1ms	41.1 ± 43.7ms
	Signed	+60.3 ± 60.9ms	-0.96 ± 60ms
2	Absolute	31.8 ± 25.9ms	87 ± 52.4ms
	Signed	-8.58 ± 40.2ms	+6.73 ± 102ms
3	Absolute	70.2 ± 51.3ms	89.7 ± 101ms
	Signed	-46.0 ± 73.9ms	-9.62 ± 135ms
4	Absolute	71.6 ± 58.7ms	127 ± 92.6ms
	Signed	+25.1 ± 89.2ms	-23.2 ± 156ms

Note. This table summarizes the asynchronies (mean ± standard deviation) in each trial of the tapping task. The data are organized by dyad (column 1), type of asynchrony (absolute or signed; column 2), and condition (synchrony or asynchrony; columns 3 and 4).

Supplementary Table 6

Sliding Task: Mean Velocity in each Trial

Dyad	Member	Velocity (M \pm SD)	
		Synchrony Condition	Asynchrony Condition
1	A	-0.012 \pm 3.66 cm/s	0.031 \pm 1.41cm/s
	B	-0.021 \pm 3.71 cm/s	-0.006 \pm 4.34cm/s
2	A	0.007 \pm 1.18 cm/s	0.003 \pm 3.01cm/s
	B	-0.005 \pm 1.2 cm/s	0.001 \pm 1.63cm/s
3	A	0.019 \pm 11.1 cm/s	-0.024 \pm 6.46cm/s
	B	-0.01 \pm 5.9 cm/s	0.016 \pm 4.09cm/s
4	A	-0.015 \pm 1.94 cm/s	-0.002 \pm 14.4cm/s
	B	0.085 \pm 2.16 cm/s	0.011 \pm 3.79cm/s

Note. This table summarizes the velocity (mean \pm standard deviation) in each trial of the sliding task. The data are organized by dyad (column 1), member of the dyad (participant A or B; column 2), and condition (synchrony or asynchrony; columns 3 and 4).

CHAPTER FIVE: DOES INTERPERSONAL SYNCHRONY AFFECT SELF-OTHER ALIGNMENT IN 18-MONTH-OLDS?

Whereas Chapters 2-4 dealt with interpersonal synchrony in the context of dyadic interactions involving two adults, the present chapter turns to interpersonal synchrony in development. As proposed mechanisms of the effects of interpersonal synchrony often rely on the idea that interpersonal synchrony relies on a recategorization of the self in relation to others, the present chapter sets out to investigate the effects of interpersonal synchrony on self-other alignment in 18-month-old toddlers, who are at the cusp of developing objective self-awareness. We therefore ask whether interpersonal synchrony affects toddlers' motivation to socially align and how this relates to their developing self-concept.

Infants demonstrate an ability to detect sensorimotor contingencies from a very young age (Jacquey et al., 2020). Yet, synchrony between their own and others' movements only appear to influence their behavior towards others in their second year of life. For instance, Tunçgenç et al. (2015) showed that 12-month-olds, but not 9-month-olds, exhibit a preference for synchronous over asynchronous social partners, but do not show a preference in a non-social context. Moreover, interpersonal synchrony has been shown to affect toddlers' helping behavior from around 14 months of age: After being bounced in synchrony with an experimenter, they are more likely to help the experimenter (Cirelli, Einarson, et al., 2014) or her affiliates (Cirelli et al., 2016), but not a neutral stranger (Cirelli, Wan, et al., 2014), than after being bounced asynchronously with her.

Drawing on the partner choice model of prosociality (Kuhlmeier et al., 2014), which suggests that individuals are selective in their social interactions and choose to cooperate with and act prosocially towards 'good' social partners, these findings have been taken to suggest

that interpersonal synchrony functions as a social cue toddlers use to direct their behavior, similar to cues for self-similarity such as language, but developing later due to their multimodal complexity (Cirelli, 2018). This line of reasoning has found support from studies in older children, who were found to help one another more and showed better cooperative problem solving (Kirschner & Tomasello, 2010b) and also showed greater bonding with out-group members (Tunçgenç & Cohen, 2016b) after interpersonal synchrony compared to asynchronous movement at age four. Moreover, eight-year old children indicated higher perceived similarity and self-other overlap with their task partner after tapping synchronously with one another compared to tapping asynchronously or not tapping at all (Rabinowitch & Knafo-Noam, 2015) and displayed greater cooperation following synchronous singing compared to other group activities (Good & Russo, 2016).

This suggests that interpersonal synchrony may act as a cue to social group membership early in development, starting in the second year of life. An important development that occurs during the second year of life is that of a self-concept, or objective self-awareness. Whereas infants attain subjective self-awareness, or a bodily self, shortly after birth, they only develop an objective sense of self, as public objects of others' attention and awareness, around 18 months (Rochat, 2018). Whether toddlers have a self-concept or not is typically assessed using the mirror mark test (Amsterdam, 1972), a test of mirror self-recognition. Toddlers in Western societies typically pass this test by 18-24 months of age (Cebioğlu & Broesch, 2021), but they succeed as early as 14 months with training (Chinn et al., 2024).

Whereas mirror self-recognition may appear trivial, it presents a challenge not only to human infants but also for most non-human animals (Gallup & Anderson, 2020). In human toddlers, it has been linked to capacities such as self-other comparison (Kampis et al., 2022), detection of perspective conflict (Yeung et al., 2022), as well as the emergence of a self-

reference effect (Grosse Wiesmann et al., 2024), and it has been associated with increased functional connectivity in brain areas associated with self-related processing (Bulgarelli et al., 2019a). Mirror self-recognition has also been shown to be related to increased compassion and readiness to help (Bischof-Köhler, 2012), as well as increased use of personal pronouns (Lewis & Ramsay, 2004a) and displays of self-conscious emotions (Lewis et al., 1989). However, it has not been found to be linked to earlier contingency preference and detection (Klein-Radukic & Zmyj, 2020). Mirror self-recognition thus appears to mark the development of a set of capabilities linked to objective self-awareness, distinct from a simple contingency detection mechanism.

During their second year of life, toddlers show a selective preference to learn from individuals who share features with themselves or their caregivers. They preferentially imitate individuals who share their native language at 14 and 18 months of age (Buttelmann et al., 2013; Altnok et al., 2022a) and they show a preference for individuals who treat similar others well and dissimilar others poorly around 14 months (Hamlin et al., 2013). While this selectivity begins to emerge during the first year of life, possibly reflecting a preference for familiar input (Southgate, 2024), during their second year of life toddlers also start to become more sensitive to other features of the people around them, which are not as easily explained by a preference for familiarity. From around 16 months toddlers start to point at novel objects more when interacting with competent individuals (Begus & Southgate, 2012a), and at 14 months they show more imitative behavior for competent over incompetent models (Buttelmann & Zmyj, 2020). This may suggest that a developmental trajectory from the detection of familiar (self-similar) features to genuine self-other comparison occurs during the second year of life.

The present study aimed to test whether interpersonal synchrony with a stranger would influence 18-month-olds' motivation to socially align with them and to investigate how this

relates to their developing self-concept. To do so, we used an adapted version of the paradigm developed by Cirelli and colleagues (Cirelli et al., 2014a) for our manipulation of interpersonal synchrony, in which toddlers were bounced either synchronously or asynchronously with an experimenter. This task was followed by an adapted version of the sticker task (Kampis et al., 2022) as a measure of toddlers' motivation to align their appearance with the experimenter's, an imitation task (Meltzoff, 1988) as a measure of their behavioral alignment, or imitative tendencies, and a pointing task (Begus & Southgate, 2012a; Lucca & Wilbourn, 2019) as measures of alignment of perspectives or mental states. These tasks were preceded by a mirror self-recognition task (Amsterdam, 1972) to test for toddlers' objective self-awareness.

We hypothesized that moving in synchrony with a stranger would increase toddlers' motivation to socially align. More specifically, we expected toddlers in the synchrony condition to show a stronger tendency to pick the same sticker as the experimenter was wearing and to show a stronger tendency to place the sticker on themselves or their caregiver in the sticker task than toddlers in the asynchrony condition. We also expected toddlers in the synchrony condition to more faithfully imitate the demonstrated action in the imitation paradigm and to point more to draw the experimenter's awareness to objects she could not see in the pointing task compared to toddlers in the asynchrony condition.

However, we also reasoned that bouncing asynchronously with the experimenter could make toddlers more sensitive to the differences between themselves and the experimenter and that this could lead to the opposite effect, with toddlers showing more self-other alignment in the asynchrony than in the synchrony condition. Thus, if toddlers show more self-other alignment in the synchrony condition, this would suggest that moving in synchrony with another person affects toddlers' motivation to socially align. However, if toddlers show more self-other alignment in the asynchrony condition, this would indicate that these behaviors are

modulated rather by cognitive mechanisms supporting the understanding of self-other distinction and individuals as unique, thereby allowing toddlers to better appreciate that if they aim to be more like the other, they need to bring about this state.

We also expected to see a relationship between our primary dependent measures and mirror self-recognition.⁴¹ Regarding toddlers' sticker choice, we broadly saw two alternatives. If mirror mark test passers are more sensitive to asynchronous bouncing, highlighting self-other distinction, then we reasoned that they should be more likely to pick the non-matching sticker in the asynchrony condition compared to non-passers. In line with this, if mirror mark test non-passers are more sensitive to the synchronous bouncing, emphasizing self-other overlap, then they should be more likely to pick the matching sticker in the synchrony condition compared to passers. However, if mirror mark test passers are more sensitive to the synchronous bouncing, as they have a higher a priori awareness of the difference between self and other, then they should be more likely to pick the matching sticker in the synchrony condition compared to non-passers. In line with this, if mirror mark test non-passers are more sensitive to asynchronous bouncing than passers, as they experience more a priori self-other overlap, then they should be more likely to pick the non-matching sticker in the asynchrony condition compared to passers.

Methods

Participants

We recruited 97 Danish-speaking 17- to 18-month-old toddlers ($M_{\text{age}} = 530.9$ days, range = 511-552 days; 52 girls, 44 boys, 1 diverse) to participate in this study.⁴² Participants

⁴¹ These hypotheses were preregistered as exploratory.

⁴² The protocol for the pointing task was only finalized after the third participant had already been tested, so the first three participants of the study were excluded from the analysis of the pointing task data.

were recruited via the participant database at the University of Copenhagen. Toddlers received a diploma and a small toy in return for their participation. The study took place in the laboratory of the Centre for Early Childhood Cognition at the University of Copenhagen in Denmark. The procedures were approved by the ethics board of the Department of Psychology at the University of Copenhagen.

We used a sequential design (Schönbrodt & Wagenmakers, 2018b) with a minimum sample size of 40, following the recommendations of Schönbrodt et al. (2017) to arrive at a minimum of 20 participants in each group, and a maximum sample size of 80. Once our minimum sample size was reached, we computed a Bayes factor (BF) comparing participants' choices in the sticker task (our main dependent measure) under the assumptions of H_0 , H_+ , and H_- . We continued data collection until we had moderate evidence (i.e., a BF greater than 3 or smaller than 1/3) either in support of the best alternative model (H_+ or H_-) or in support of H_0 , re-evaluating the evidence after every four participants to allow for even counterbalancing between conditions, or until our maximum sample size was reached.

Once our minimum sample size was reached, we also coded participants' behavior in the mirror mark test. Until we reached our minimum sample size, participants were randomly assigned to the synchrony or asynchrony condition prior to arriving at the lab, and both experimenters (E1 and E2) were aware of the experimental condition the participant was assigned to. Once our minimum sample was reached, participants were assigned to the synchrony or asynchrony group after completing the mirror mark test to arrive at a more balanced sample of mirror mark test passers and non-passers in each condition. Once the sample was balanced, participants were randomly assigned to one of the two conditions prior to arriving at the lab. The non-random assignment was limited to sets of four participants to ensure for even counterbalancing of factors for each re-evaluation of our stopping criterion.

Materials and Procedure

Each session began with a 10-minute warm-up phase, during which children were familiarized with the space and both experimenters. After the warm-up phase, E1 explained the study procedures to the caregiver and obtained their informed consent. Children then completed five tasks: a mirror mark test, a rhythmical movement task, a sticker task, an imitation task, and a pointing task (see Figure 25). Caregivers were asked to fill out a short demographics questionnaire during the session as well as a longer take-home questionnaire prior to coming to the lab.

Figure 25

Procedure



Note. Participants completed five tasks, depicted here from left to right: a mirror mark test, a rhythmical movement task, a sticker task, an imitation task, and a pointing task.

Mirror Mark Test

The mirror mark test followed the protocol of Kamps et al. (2022). Children entered the testing room and were seated in a chair facing the wall where the mirror was hidden. Children who did not want to sit on the chair either stood facing the mirror or were seated on their caregiver's leg facing the mirror. E2 revealed the mirror together with the child and drew the child's attention towards the mirror by tapping on the mirror or flipping her fingers in front of it (phase 1). Once children had either looked at their mirror image for 10 seconds or fixated towards it three times, E2 hid the mirror again and handed the child a toy to play with.

While the mirror was hidden, E2 discreetly placed a lipstick smear on the child's nose while pretending to blow their nose (phase 2). She waited for 10 seconds to ensure that the child had not noticed the mark application without seeing it in the mirror, then proceeded to unveil the mirror again together with the child (phase 3). Once the mirror was visible again, E2 drew the child's attention towards it by tapping on it or flipping her fingers in front of it. Once children reached up to touch their nose or had looked at their mirror image for 10 seconds or fixated towards it three times, E2 drew the child's attention towards the mirror image again and asked, 'Who is that?' (in Danish; phase 4). If the child did not respond, she repeated the question after 5 seconds and again after another 5 seconds had passed.

During the entire task, caregivers were asked to sit quietly in the corner on a chair. In some cases, the caregiver was invited to sit on the floor next to the mirror to allow the child to sit on their legs during the task. In these cases, caregivers were instructed to make sure their face was not visible to the child through the mirror and to avoid making eye contact with the child through the mirror.

Rhythmical Movement Task

After completing the mirror mark test, children completed a rhythmical movement task adapted from Cirelli et al. (2014). For this task, children were placed in a carrier on their caregiver's body, facing outwards. The caregiver then stood facing E1, about 120 cm away, and was given headphones to wear for the duration of the task. Once they were ready, E2 started the music (a midi version of Twist and Shout by The Beatles), which played from a speaker right behind the caregiver (slightly to their left) for 145s.

Together with the music, the bounce instruction soundtrack started in the respective headphones for the caregiver and E1. Caregivers were instructed to bounce to the soundtrack as best they could. Both the music and the soundtrack that was played to caregivers was taken

from Cirelli et al., (2014). If children were assigned to the synchrony condition, E1 listened to the same bounce instruction soundtrack as the caregiver and bounced according to the instructions. If children were assigned to the asynchrony condition, E1 listened to a soundtrack that was 33% faster than the soundtrack which was played to parents. E1 directed her attention towards the child during this task, maintaining eye contact when possible and generally maintaining a friendly expression throughout the task (in both conditions).

Parents were informed about the condition they were in prior to completing the rhythmical movement task. They received a short information video about the task together with the study invite to familiarize themselves with what they would be asked to do before coming to the study session. Caregivers were told that they did not have to look at E1 during this task but could instead close their eyes or look around the room if that made the task easier to complete. They were instructed to keep their body facing E1 for the full duration of the task.

During the rhythmical movement task, E1 wore an accelerometer on her chest and held a mouse in her right hand which she used to track eye contact between herself and the baby during the movement phase by clicking when she and the child had direct eye contact and releasing when they did not. A second accelerometer was placed on front of the carrier in a small pouch or the child's back when parents opted to bring their own carrier or to perform the task without a carrier but by holding their child instead.

Sticker Task

After the rhythmical movement task, E1 took the caregiver and child to the waiting area and helped caregivers take the child back out of the carrier. She then went back to the testing room, leaving the caregiver and child with E2, to prepare for the sticker task (adapted from Kampis et al. (2022)). While out of the child's sight, E1 placed a sticker on her face (either on

her forehead or her cheek). When the caregiver and child returned to the testing room, they saw her with the sticker on her face and sat down facing her.

E2 then kneeled in front of them and offered the child two stickers to choose from – one matching the one E1 was wearing, one not. E2 encouraged children to choose one of the two stickers. If children did not choose a sticker, they were assigned the matching sticker. Once children had chosen a sticker (or been assigned a sticker), the spontaneous phase began. For 60 seconds, the child’s spontaneous behavior with the sticker was observed.

If the 60 seconds passed without the child doing anything with the sticker, E2 prompted them, asking ‘What should we do with the sticker?’ (in Danish). Then, after 20 seconds, ‘Where should we put it?’ (in Danish), and after another 20 seconds, ‘Should we put it on you?’ (in Danish), and finally, after another 20 seconds, ‘Should we put it on [caregiver]?’ (in Danish). If the child placed the sticker on their own face or on their caregiver’s face, E2 immediately asked them about the other of the two (i.e., whether they would like to put the sticker on their caregiver if they had placed it on themselves, or whether they would like to put the sticker on themselves if they had placed it on their caregiver).

Caregivers were instructed not to direct their child’s behavior during the task (they were asked not to suggest anything the children could do with the sticker), but they were told that they could make broad encouraging statements if the child did not seem particularly interested in the sticker (e.g., ‘Oh, what a nice sticker!’).

Imitation Task

Once the sticker task was complete, participants moved on to the imitation task (adapted from Altınok et al., 2022). Caregivers were asked to sit on a chair with their child on their lap, while E1 prepared a table, sat across them at it, and brought out a lamp (see Figure 26). Once

everyone was seated, E1 drew the child's attention to the lamp and uncovered it. She then made eye contact with the child and began the demonstration.

At the beginning of each demonstration, E1 placed her hands on the table, next to the lamp. Then she lowered her head towards it and turned it on using her forehead, then sat back up after about two seconds. E1 demonstrated this action three times, smiling at the child before starting each demonstration. Then she covered the lamp again and took the child and the caregiver back to the waiting area, where they then spent five minutes playing with E2.

Figure 26

Imitation Task: Apparatus



Note. The lamp used in the imitation task was mounted on a black box. The lamp could be operated by pressing on the white surface and remained lit until the surface was no longer being pressed. At the start of the demonstration and response phase the lamp was covered by a grey cloth. During the response phase, the black box was stuck to the table using blu tack to keep the apparatus in place while toddlers interacted with it.

Once the five minutes were up, E1 invited the caregiver and child to return to the testing room, asking them to sit in front of the (covered) lamp. She then revealed the lamp, told then child, 'Your turn!' (in Danish), and left the room. E2 timed 60 seconds from the child's initial

engagement with the apparatus, then asked parents to move to the other side of the table and signaled to E1 to return to the testing room for the pointing task.

Pointing Task

The pointing task was adapted from Begus & Southgate (2012) and Lucca & Wilbourn (2019). E1 re-entered the room, removed the lamp from the table and sat down opposite the child and their caregiver with a small box of toys on her lap (see Figure 27), hidden from the child's sight. She took the first toy out of the box and began playing with the child. Meanwhile, E2, from behind a large curtain, began to display an unknown object (see Figure 28) that the child and their caregiver could see but E1 could not (see Figure 29).

Figure 27

Pointing Task: Toys

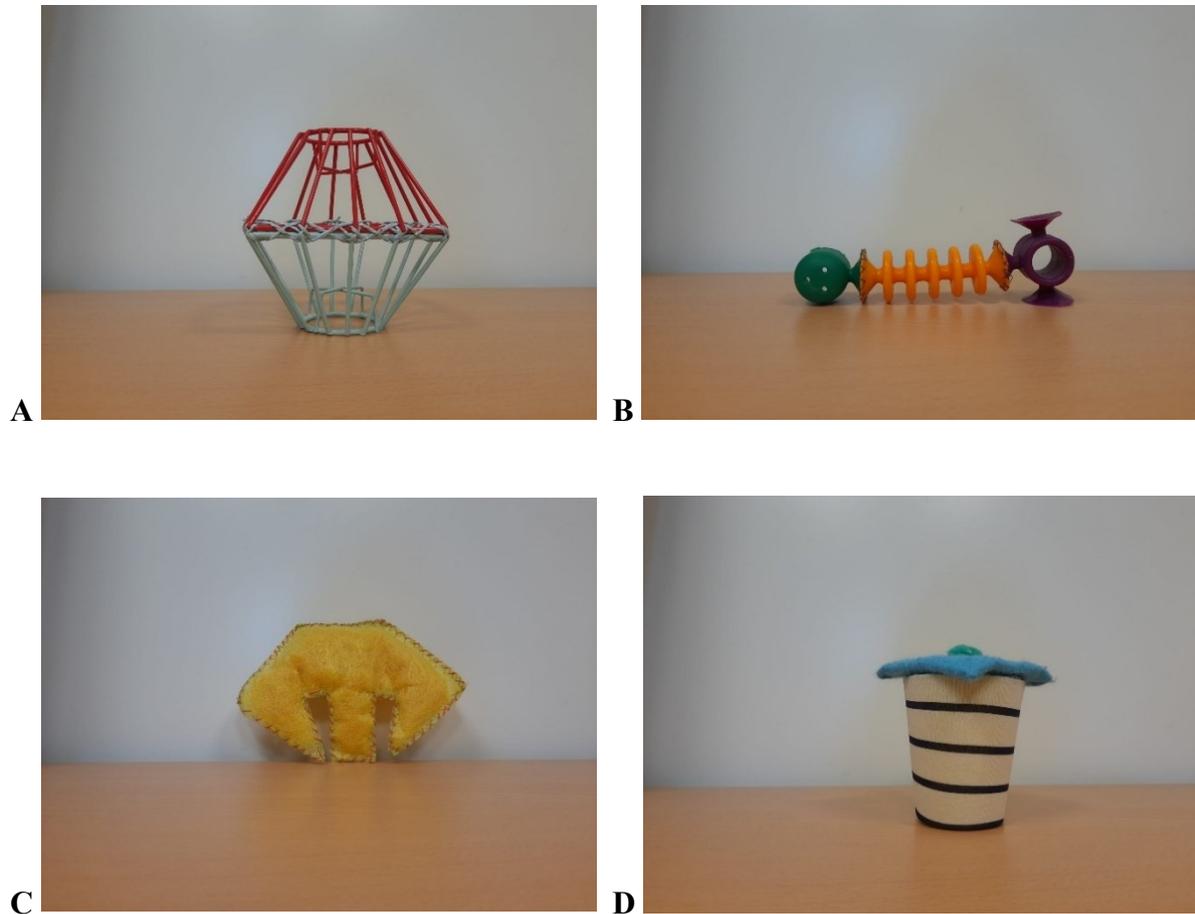


Note. During the pointing task, E1 had a box containing the six toys depicted in this figure on her lap and played with toddlers using one of the depicted toys per object that was presented. Objects were selected based on piloting to ensure toddlers were interested enough to proceed with the task but not too enthralled by the toys to lack interest in the objects being presented.

E2 displayed the unknown object for up to 60 seconds, waiting for a signal from E1 indicating that the child had looked at the location at which the object was being displayed. If the child fixated either to the left or the right above her head, E1 scratched the back of her neck to indicate to E2 that the child had looked at the unknown object, without turning to face it. If the child did not look at the object, E2 removed it after 60 seconds.

Figure 28

Pointing Task: Objects



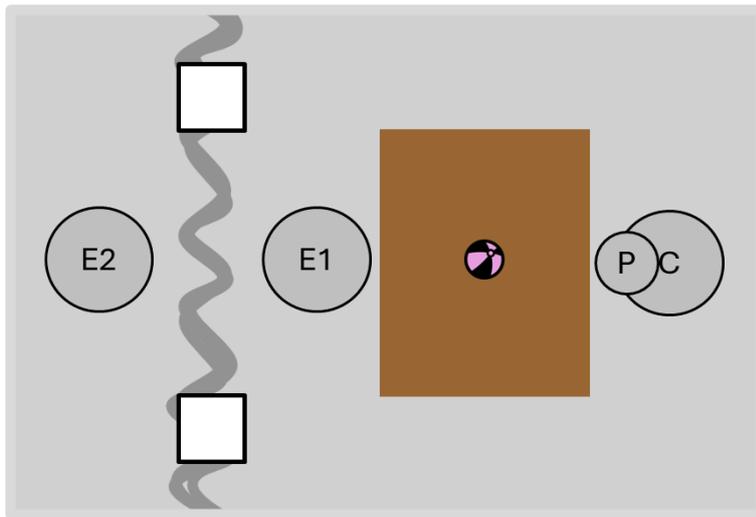
Note. During the pointing task, toddlers were shown the four unknown objects depicted in this figure. The objects were designed for another study (Grosse Wiesmann et al., 2024) and selected here, as they were both unfamiliar to toddlers and sufficiently different from one another for toddlers to be able to discriminate them.

If E1 signaled that the child had looked at the object, E2 continued to display the object for another 15 seconds or until the child pointed at the object. If the child pointed at the object, E1 turned towards it, then turned back to the child and said, “Wow, a [label]!” (in Danish). She repeated this once more, then E2 removed the object. The labels used in this task were Danish pseudowords borrowed from Kamps (personal communication): *trøft*, *girk*, *sæm*, and *stryk*.

Once the object was removed, E1 switched out the toy they had been playing with for a new one, and the procedure was repeated three more times, with a novel object, toy, and label in each iteration.

Figure 29

Pointing Task: Room Layout



Note. During the pointing task participants (P) sat on their caregiver's (C) lap at a table, facing E1. E1 played with the participants using the set of toys depicted in Figure 27, while E2 discretely displayed the objects depicted in Figure 28 from behind a curtain one by one. Each object was shown in one of the two locations marked with white squares, behind E1 but within the participant's line of sight.

If the child pointed to at least two of the objects, and thus received two labels, there was an additional test phase after all four objects had been displayed. For this, E2 discretely handed E1 the first two objects the child had pointed at from behind the curtain.⁴³ E1 then showed the

⁴³ The first two labels provided in this task were always [trøft] and [girk]. Their order was counterbalanced across participants, and they were the only two labels used at test.

objects to the child⁴⁴ and asked for one of them, “Look! Where is the [label]?” (in Danish). She then placed the objects on either side of the table, moved them closer to the child, then placed her right palm face-up in the middle of the table and asked, “Can you give me the [label]?” This question was repeated up to three times.

Questionnaires

Caregivers were sent a link to fill out a survey including the mirror mark test questionnaire used by Kampis et al. (2022) and six sub-scales of the short version of the Early Childhood Behavior Questionnaire (ECBQ; Putnam et al., 2006) prior to their study participation and filled out a shorter demographics questionnaire during the test session. Caregivers who completed the test session but did not fill out the take-home questionnaire prior to participating were asked to complete the questionnaire as soon after their participation as possible. The selected ECBQ sub-scales were attentional shifting, cuddliness, inhibitory control, low intensity pleasure, shyness, sociability.

Data Treatment

Mirror Mark Test

Children’s behavior in each phase of the mirror mark test was coded to make sure they met the inclusion criteria (looking at themselves for 10 seconds or three times in phase 1, not interacting with the mark in phase 2, looking at themselves for 10 seconds or three times or interacting with the mark in phase 3). Children were coded as mirror mark test passers if they showed mark directed behavior in phase 3 or 4. They were coded as non-passers if they met the inclusion criteria but did not display any mark directed behavior. The data were coded by

⁴⁴ Initially, E1 simply held up both objects for the child to see and then placed them on either side of the table. This procedure was modified after approximately one third of the sessions were completed such that the experimenter first held the objects in front of her face to make sure the child had seen both objects before asking for one of them and placing the objects on either side of the table.

two independent coders. Inter-coder agreement was near perfect ($\kappa = 0.93$). Coding disagreements were resolved by discussion or by involving a third party.

Rhythmical Movement Task

The total duration of mouse clicks by E1 during the rhythmical movement task was summed up for each session to obtain a measure of how much eye contact there was between the child and the experimenter during the rhythmical movement task. The overall duration was entered into a Bayesian independent samples t test to compare the duration of direct eye contact between conditions.

The x, y, and z coordinates from the accelerometers were used to compute acceleration for both E1 and the child, using their sum of squares. The acceleration data was mean centered and interpolated using the `approx()` function in R to reach a consistent sampling rate of 50Hz (one sample every 20ms). We computed individual phase angles for both E1 and the child using the Hilbert transform and then computing the arctan of the transformed signal. We obtained their relative phase by subtracting one time series (the child's) from the other (E1's) and used this to calculate their synchronization index (Mardia & Jupp, 2009) as a measure of the stability of each dyad's coordination during the rhythmical movement task. We compared dyads' synchronization indices between conditions to test whether our manipulation was successful.

In addition to this, we used the video recordings of the rhythmical movement task to generate three sets of 30-second clips, which were manually coded by a researcher who was blind to the purposes of this study. One set of clips included only the recording of E1 (the child and parent were not visible) 60 seconds into the rhythmical movement task. As in Cirelli et al. (2017), the coder was asked to code (a) how happy, smiley, attentive, interactive, and connected to her bounce partner E1 seemed on a 10-point Likert scale, based on these clips, and (b) to

judge whether E1 was bouncing synchronously or asynchronously with her out-of-view movement partner in a two-alternative forced-choice question. We generated a second set of clips from the same time window, this time removing E1 from view, and asked the coder to rate the child's and the caregiver's enjoyment of the interaction, respectively. Finally, we generated a third set of clips, in which both E1 and the parent-child dyad were in view, and we asked the coder to rate how coordinated E1's and the parent-child-dyad's movements were and to guess whether they were trying to bounce synchronously or asynchronously. We removed the audio track from all clips to ensure that coders based their ratings on movement and facial expressions.

Sticker Task

We coded whether children chose a sticker, and if they did, whether they chose the same sticker E1 was wearing (the matching sticker) or not. The data were coded by two independent coders. Inter-coder agreement on whether children chose a sticker ($\kappa = 0.63$) and on whether the chosen sticker matched the one E1 was wearing ($\kappa = 0.79$) was substantial but not perfect due to participant movements obscuring coders' view of the sticker, the impossibility of adjusting camera angles during the task, and low-resolution camera footage. Coding disagreements were resolved by discussion or by involving a third party.

For our main analysis, we computed a Bayesian A/B test comparing three models: H_0 (no difference between the synchrony and asynchrony group; prior probability: 0.5), H_+ (more children in the synchrony group choose the matching sticker than in the asynchrony group; prior probability: 0.25), and H_- (fewer children in the synchrony group choose the matching sticker compared to the asynchrony group; prior probability: 0.25). Children who did not choose a sticker were excluded from this analysis.

We also coded any attempts made to place the sticker on their own face or their caregiver's face and whether they occurred during the spontaneous phase of the task or after one of the prompts. Inter-coder agreement for these items was substantial (all $\kappa \geq 0.68$). Coding disagreements were resolved by discussion or by involving a third party. We used the coding of individual behaviors to create one composite score, as preregistered, indicating whether participants placed the sticker on their own face or their caregiver's face in either the spontaneous or the prompted phase of the task.

We analyzed composite sticker placement scores following the same logic as for the sticker choice, using a Bayesian A/B test to compare three models: H_0 , H_+ , and H_- for placement on their own face and on their caregiver's face combined. Children who did not complete the task were excluded from the analysis. Children who did not choose a sticker but completed the remainder of the task with the assigned sticker were included in this analysis.

Imitation Task

We coded whether children used their head in the first 20 seconds of the response phase ($\kappa = 0.95$), their frequency of head ($\kappa = 0.98$) and hand touches ($\kappa = 0.66$), respectively, over the course of the entire response phase, and the latency of their first attempt to turn on the lamp ($\kappa = 0.4$). The frequency data was used to compute proportions of head touches during the full 60-second response phase for each child. These data were entered in a Bayesian t test to compare children's tendency to use their head between conditions. The latency of their first touch was also compared between conditions using a Bayesian t test. The analysis of whether they spontaneously used their head (in the first 20 seconds of the response phase) followed the same logic as the analysis of sticker choice and placement: We used a Bayesian A/B test to compare toddlers' behavior between groups.

Pointing Task

For each of the displayed objects, we coded whether children looked at it ($\kappa = 1$) and whether they pointed at it ($\kappa = 1$). For children who completed the test, we additionally coded whether, at test, their first look was at the correct object ($\kappa = 0.58$), whether their first reach was for the correct object ($\kappa = 1$), and whether they handed E1 the correct object ($\kappa = 1$). We analyzed toddlers' pointing behavior using a Bayesian generalized linear mixed model (GLMM) with Subject ID and Object Number as random effects grouping factors. We compared participants' behavior at test between conditions using a Bayesian A/B test.

Questionnaire Data

Eight caregivers of toddlers who completed the rhythmical movement task did not complete the take-home questionnaire. For the remaining 88 toddlers, we computed ECBQ scores for the attention shifting, cuddliness, inhibitory control, low intensity pleasure, shyness, sociability⁴⁵ sub-scales following the scoring procedure made available at <https://research.bowdoin.edu/rothbart-temperament-questionnaires/instrument-descriptions/the-early-childhood-behavior-questionnaire/> (last accessed 2025-06-01). We then tested for a relationship between the ECBQ sub-scales and participants' matching sticker choices, sticker placement, and imitation behavior (head touches in the first 20 seconds of the response phase) in three Bayesian logistic regressions (one for each measure).

⁴⁵ One caregiver selected 'does not apply' for all items on the sociability sub-scale, thus no sociability score could be computed for this child.

Results

Mirror Mark Test

All but one child, who did not let E2 put the mark on their nose, completed the mirror mark test. Of the remaining 96 participants, seven were excluded because they touched their nose after application of the mark but before the mirror was revealed. Two more participants were excluded due to ambiguous behavior (one), and because they did not complete the rhythmical movement task (one). Among the remaining 87 participants, there were 48 children who passed the mirror mark test, and 39 who did not. Non-passers were relatively evenly split across the synchrony condition (20) and the asynchrony condition (19). There were slightly more passers in the asynchrony condition (26) than in the synchrony condition (22).

Rhythmical Movement Task

All but one child completed the rhythmical movement task. 48 of the remaining 96 children completed the synchrony condition; 48 children completed the asynchrony condition.

Eye Contact

Due to experimenter error, eye contact was not recorded in nine sessions. The data of the remaining 87 sessions (42 in the synchrony condition; 45 in the asynchrony condition) was entered into a Bayesian independent samples t test to compare the duration of eye contact between child and experimenter in the synchrony and asynchrony condition. The analysis showed moderate evidence for H_0 : $BF_{01} = 4.23$, indicating that the duration of children's eye contact with the experimenter during the rhythmical movement task did not differ between the synchrony ($53.6 \pm 29s$) and asynchrony condition ($55.7 \pm 28.5s$).

We followed up on this result to test for a difference between mirror mark test passers and non-passers, as well as a possible interaction with condition, in a 2x2 Bayesian ANOVA.

The analysis showed that the model that included only the factor Mirror Mark Test Passing was the best model of the data ($P_{M|data} = 0.53$), 2.27 times more likely to be true than the null model, with mirror mark test passers showing a longer overall duration of eye contact during the rhythmical movement task ($60.4 \pm 28s$) than non-passers ($45.8 \pm 28s$).

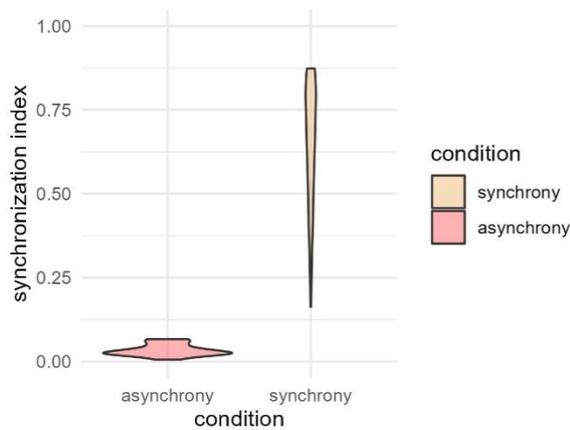
Synchronization

As the accelerometers were not included in the protocol for the first three experimental sessions, these three participants' data were excluded from the analysis. An additional nine participants' data was excluded from the analysis of synchronization between the child and E1 due to experimenter error (in one case the data file was mistakenly deleted; in the other eight the accelerometer was not correctly placed on E1). The remaining 84 participants' data was not normally distributed and therefore entered into a Bayesian Mann-Whitney U test to test for a difference in synchronization between conditions. The analysis showed extreme evidence that dyads' synchronization indices were greater in the synchrony condition (0.65 ± 0.18) compared to the asynchrony condition (0.03 ± 0.02): $BF_{10} = 133023.66$ (see

Figure 30). This suggests that coordination during the rhythmical movement task was much more stable in the synchrony than in the asynchrony condition, demonstrating that our manipulation was successful.

Figure 30

Synchronization Index by Condition



Note. This violin plot depicts the distribution of synchronization index values (y axis) in the synchrony (yellow) and asynchrony (red) condition. Larger values indicate more stable coordination.

Video Ratings

The coder was also able to predict which condition they were watching based on recordings of only E1 bouncing at 95.7% accuracy in the synchrony condition, and at 80.9% accuracy in the asynchrony condition.⁴⁶ Moreover, the Bayesian independent samples t test comparing video ratings of E1’s facial expressions between the synchrony and the asynchrony condition also showed extreme evidence that the ratings of E1’s facial expression were higher in the asynchrony condition (8.49 ± 1.33) compared to the synchrony condition (7.29 ± 0.94): $BF_{10} = 133.18$, suggesting that E1’s attitude towards the child during the rhythmical movement task was not consistent across conditions and that she was scored higher on the rating of how

⁴⁶ However, due to technical error, the coder was able to see the caregiver-child dyad’s movement in one of the videos and used this information to predict the condition in subsequent videos. Excluding this ratings and those performed thereafter (18.75% of the videos), the coder’s accuracy dropped to 92.5% for the synchrony condition and 73.68% for the asynchrony condition, suggesting that the coder’s accuracy was high even before seeing the faulty recording.

“happy, smiley, interactive, attentive, and connected” to her bounce partner she appeared, when bouncing asynchronously with them.

However, the Bayesian independent-samples *t* tests comparing the video ratings of caregiver and child enjoyment between the synchrony and asynchrony condition showed anecdotal evidence for no difference in the caregiver’s enjoyment ($BF_{01} = 2.78$) and moderate evidence for no difference in the child’s enjoyment of the rhythmical movement task between conditions ($BF_{01} = 3.35$). The Bayesian independent samples *t* test comparing video ratings of coordination in the synchrony and asynchrony condition also revealed extreme evidence for better movement coordination in the synchrony condition (8.85 ± 1.1) compared to the asynchrony condition (2.4 ± 1.04): $BF_{10} = 33661.99$, and the coder’s prediction of which condition they were watching based on videos showing both E1 and the parent-child dyad were near perfect: 100% accuracy in the synchrony condition; 97.9% accuracy in the asynchrony condition. This suggests that neither participants nor their caregivers had a clear preference for one of the two conditions, and that the manipulation was successful in bringing about more closely coordinated movement in the synchrony condition and less closely coordinated movement in the asynchrony condition.

Sticker Task

Sticker Choice

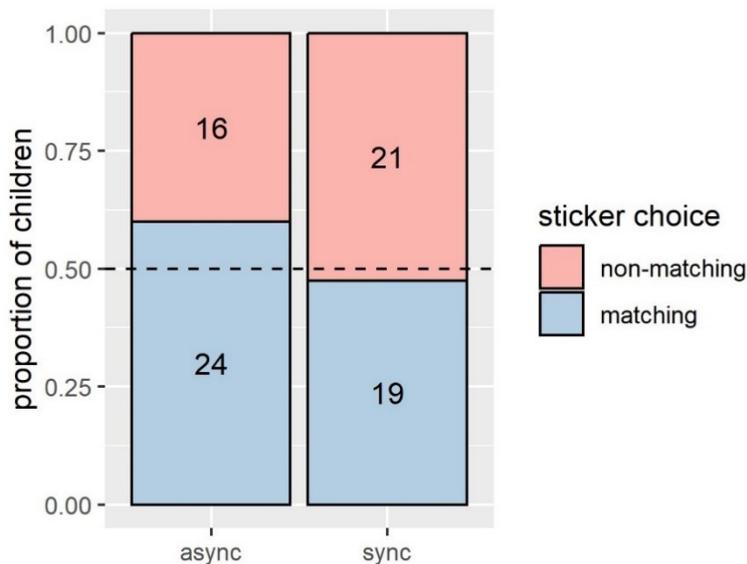
16 children (eight in the synchrony group, eight in the asynchrony group) did not choose a sticker and were therefore excluded from the analysis of sticker choice. The remaining 80 children’s data were entered into a Bayesian A/B test, our preregistered main analysis. The results showed anecdotal evidence for H_1 compared to H_0 ($BF_{-0} = 1.19$), suggesting that children in the asynchrony group are slightly more likely to pick the matching sticker (60%) than children in the synchrony group (47.5%; see

Figure 31). However, the analysis also revealed that H_0 remained the best model, given the data ($P_{M|data} = 0.59$), and that H_0 is 4.7 times more likely to be true than H_+ , indicating that overall, children in the synchrony group were not more likely to pick the matching sticker than children in the asynchrony group.

We followed up on these results by comparing sticker choice in each condition to chance (set to 0.5) using a non-directional Bayesian binomial test. The analyses showed moderate evidence that children in the synchrony group did not have a preference between the two stickers ($BF_{01} = 4.9$) and anecdotal evidence that children in the asynchrony group did not have a preference between the two stickers either ($BF_{01} = 2.34$). This suggests that overall, children did not have a clear preference for the matching or the non-matching sticker in either condition.

Figure 31

Sticker Choice by Condition

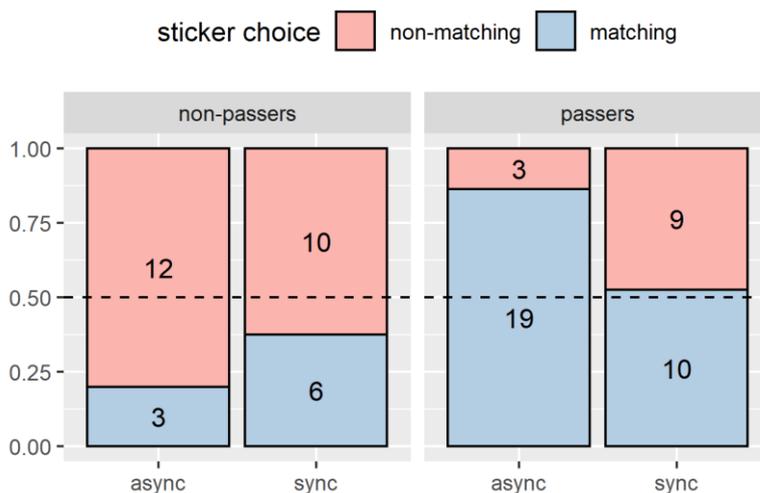


Note. The figure shows the proportion of children who picked the matching sticker (blue) compared to the non-matching sticker (red) in the asynchrony (left) and synchrony condition (right). The dashed line represents chance level (0.5).

Next, we conducted separate Bayesian A/B tests for mirror mark test passers and non-passers as preregistered comparing their matching sticker choices in each condition.⁴⁷ The Bayesian A/B test comparing mirror mark test passers' sticker choices in the synchrony and asynchrony condition showed moderate evidence for H_+ compared to H_0 ($BF_{10} = 7.89$), suggesting that mirror mark test passers were more likely to choose the matching sticker in the asynchrony condition (86.4%) compared to the synchrony condition (52.6%; see Figure 32, right panel). The Bayesian A/B test comparing mirror mark test non-passers' sticker choices in the synchrony and asynchrony condition showed anecdotal evidence for H_+ compared to H_0 ($BF_{10} = 1.48$), suggesting that mirror mark test non-passers may have had a preference for the matching sticker in the synchrony (37.5%) compared to the asynchrony condition (20%; see Figure 32, left panel). However, H_0 that was the best model, given the data ($P_{M|data} = 0.52$), 2.81 times more likely than H_+ .

Figure 32

Sticker Choice by Condition and Mirror Mark Test Passing



Note. This figure depicts the proportion of children who picked the matching sticker (blue) compared to the non-matching sticker (red), split by mirror self-recognition: non-passers'

⁴⁷ Among the participants who chose a sticker, there were 41 children who passed the mirror mark test (22 in the asynchrony condition) and 31 children who did not pass the mirror mark test (15 in the asynchrony condition).

choices are depicted in the left panel, and passers' choices are depicted in the right panel. Each panel depicts their choices in the asynchrony (left bar in each panel) and synchrony condition (right bar in each panel). The dashed line represents chance level (0.5).

We followed up on these results by comparing sticker choice in each sub-group to chance (set to 0.5) using a non-directional Bayesian binomial test. The analyses showed extreme evidence that mirror mark test passers in the asynchrony group had a for the matching sticker ($BF_{10} = 118.42$) and moderate evidence that mirror mark test non-passers in the asynchrony had a preference for the non-matching sticker ($BF_{10} = 3.52$). The analyses also revealed moderate evidence that mirror mark test passers in the synchrony group did not have a preference between the two stickers ($BF_{01} = 3.52$) and anecdotal evidence that mirror mark test non-passers in the synchrony group did not have a preference between the two stickers either ($BF_{01} = 2.08$). This suggests that whereas neither passers nor non-passers showed a clear preference for either sticker in the synchrony condition, children in the asynchrony group had a clear preference for one of the two stickers. This preference varied between mirror mark test passers and non-passers, with mirror mark test passers preferring the matching sticker and mirror mark test non-passers preferring the non-matching sticker.

To gain a better understanding of the relationship between the effect of synchrony on motivation to align and the development of the self-concept, we additionally computed a Bayesian logistic regression with the factors Condition and Mirror Mark Test Passing, using a uniform prior, on children's sticker choices. The analysis showed that the full model, including both factors as well as their interaction, was the best model, given the data: $P_{M|data} = 0.65$, and that it was 148.94 times more likely to be true than the null model. The outcome of this analysis suggests that there was an interaction between the factors Condition and Mirror Mark Test Passing that affected children's sticker choices ($BF_{inclusion} = 7.52$), such that mirror mark test passers were less likely to pick the matching sticker in the synchrony condition ($\beta = -1.3$, 95%

confidence interval [-3.67, 0.02]), but that the factor Mirror Mark Test Passing alone may also have acted as a predictor of children's choices ($BF_{inclusion} = 100.16$), with mirror mark test passers being more likely to pick the matching sticker overall ($\beta = 2.12$, 95% confidence interval [0.5, 3.9]).

We repeated this analysis including the following additional predictors: the synchronization index (measured during the rhythmical movement task)⁴⁸, the duration of eye contact between the child and E1 during the rhythmical movement task, ratings of E1's facial expressions during the rhythmical movement task⁴⁹, ratings of the child's enjoyment of the rhythmical movement task, ratings of the caregiver's enjoyment of the rhythmical movement task, and the child's age.⁵⁰ The analysis showed that the best model of the data was one that included both the factors Mirror Mark Test Passing and Condition, as well as their interaction and the additional predictors Synchronization Index, and Rating of Caregiver Enjoyment: $P_{M|data} = 0.03$, 382.87 times more likely than the null model, and 2.64 times more likely to be true than the model that included only the factors Mirror Mark Test Passing and Condition, and their interaction.⁵¹ Synchronization Index acted as a positive predictor ($\beta = 0.85$, 95% confidence interval [-1.85, 5.2], $BF_{inclusion} = 1.07$), as did Rating of Caregiver Enjoyment ($\beta = 0.16$, 95% confidence interval [-0.14, 0.5], $BF_{inclusion} = 2.46$). This suggests that the degree of coordination between E1 and the parent-child dyad during the rhythmical movement task as

⁴⁸ As there was extreme evidence for a correlation between the synchronization and the video ratings of coordination during the rhythmical movement task ($\tau = 0.61$, $BF_{10} = 12,860.000,000,000$), we included only the synchronization index as a predictor in the analysis, as it provides a more fine-grained quantification of the degree of coordination between the experimenter and the caregiver-child dyad.

⁴⁹ There was also extreme evidence for a correlation between the ratings of E1's facial expressions during the rhythmical movement task and the synchronization index ($\tau = -0.34$, $BF_{10} = 1,889.57$), as well as between the ratings of E1's facial expressions during the rhythmical movement task and the ratings of coordination during the rhythmical movement task ($\tau = -0.38$, $BF_{10} = 6,9975.53$). However, as the correlation coefficients were relatively small (see Dormann et al., 2013), and the predictors capture meaningfully different constructs, we kept both the synchronization index and the ratings of E1's facial expressions during the rhythmical movement task as predictors in the analysis.

⁵⁰ For a heatmap visualization of all correlation coefficients, see Supplementary Figure 9.

⁵¹ For a table of β coefficients, see Supplementary Table 7.

well as the caregiver's enjoyment of the task additionally affected whether children chose the matching sticker or not.⁵²

Sticker Placement

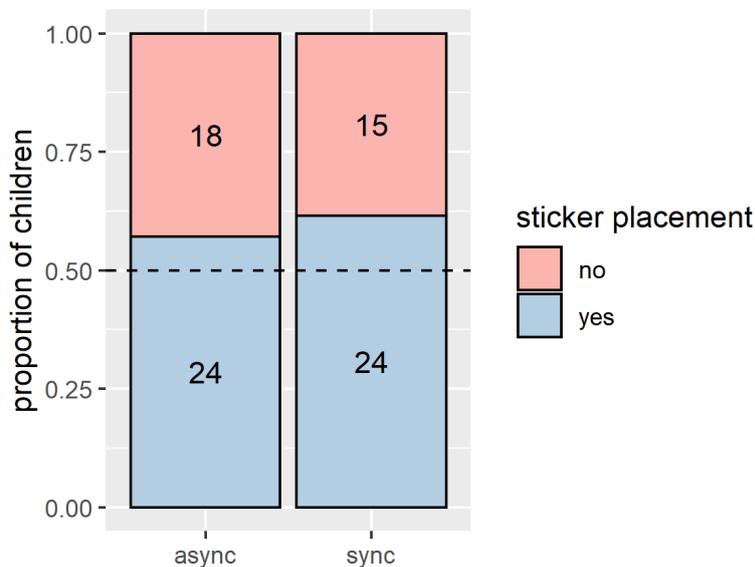
Fifteen children were excluded from the analysis of sticker placement because they did not complete the task. Four of these children chose a sticker but did not complete the rest of the task; 11 children did not complete either part of the task. Five children did not choose a sticker but completed the remainder of the task with the assigned (matching) sticker. The remaining 81 children were included in the analysis of sticker placement, as preregistered.

As preregistered, we computed a Bayesian A/B test to compare children's placement of the sticker onto their own or onto their companion's face between conditions. The analysis showed that H_0 was the best model of the data: $P_{M|data} = 0.69$ and that H_0 is 1.72 times more likely to be true than H_+ and 3.09 times more likely to be true than H_- . This suggests that children did not show a substantially stronger tendency to place the sticker on their own or their caregiver's face after moving in synchrony with the experimenter (61.5%) compared to moving asynchronously with her (57.1%; see Figure 33).

⁵² As gender has been found to affect overimitation in children (Frick et al., 2017) and may therefore be a relevant predictor of self-other alignment in development more broadly, we recomputed this analysis with Gender as an additional predictor. However, in order to compute this analysis, one child had to be excluded (gender reported as 'diverse'). The outcome of this analysis showed that the best model of the data was one that included both the factors Mirror Mark Test Passing and Condition, as well as their interaction and the additional predictors Synchronization Index, Rating of Caregiver Enjoyment, and Gender: $P_{M|data} = 0.01$, 392.39 times more likely than the null model.

Figure 33

Sticker Placement on Own or Companion's Face by Condition



Note. This figure depicts the proportion of children who placed the sticker on their own or their companion's face (blue) in the asynchrony (left) and synchrony condition (right). The dashed line marks a proportion of 0.5.

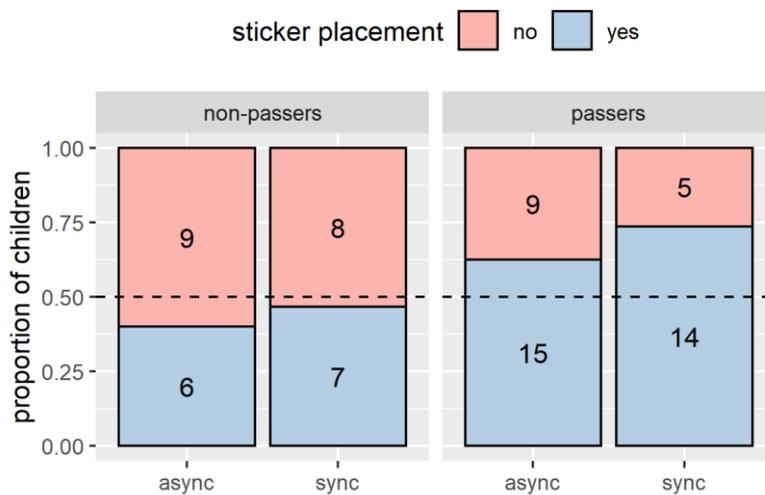
Next, we computed a Bayesian A/B test comparing mirror mark test passer' and non-passers' sticker placement behavior.⁵³ The analysis showed moderate evidence for H_+ over H_0 ($BF_{10} = 5.01$), suggesting that mirror mark test passers were more likely to place the sticker on their own or their caregiver's face (67.4%) than non-passers (43.3%), consistent with the previous findings of Kamps et al. (2022). As for the sticker choice, we also conducted separate Bayesian A/B tests for mirror mark test passers and non-passers comparing their sticker placement in each condition. The Bayesian A/B test comparing mirror mark test passers' sticker placement in the synchrony and asynchrony condition showed that H_0 was the best model of the data ($P_{M|data} = 0.6$), 2.84 times more likely to be true than H_+ and equally likely to be

⁵³ Among the participants who completed the sticker task, there were 43 children who passed the mirror mark test (24 in the asynchrony condition) and 30 children who did not pass the mirror mark test (15 in each condition).

true as H_+ , suggesting that mirror mark test passers were not substantially more likely to place the sticker on their own or their caregiver's face in the synchrony condition (73.7%) compared to the asynchrony condition (62.5%; see Figure 34, right panel). The Bayesian A/B test comparing mirror mark test non-passers' sticker placement in the synchrony and asynchrony condition also showed that H_0 was the best model of the data ($P_{M|data} = 0.61$), 2 times more likely to be true than H_- and 1.28 times more likely to be true than H_+ , suggesting that mirror mark test non-passers also were not substantially more likely to place the sticker on their own or their caregiver's face in the synchrony condition (46.7%) compared to the asynchrony condition (40%; see Figure 34, left panel).

Figure 34

Sticker Placement on Own or Companion's Face by Condition and Mirror Mark Test Passing



Note. This figure depicts the proportion of children who spontaneously placed the sticker on their own face (blue), split by mirror self-recognition. Mirror mark test non-passers' data is shown in the left panel; mirror mark test passers' data is shown in the right panel. In each panel, the left bar shows children's behavior in the asynchrony condition, and the right bar shows their behavior in the synchrony condition. The dashed line marks a proportion of 0.5.

To gain a better understanding of the relationship between the effect of synchrony on children's tendency to place the sticker on their own or their companion's face and the development of the self-concept, we additionally followed up on these results in a Bayesian logistic regression, as we did for the sticker choice. The analysis showed that the best model, given the data, was one that included only the factor Mirror Mark Test Passing: $P_{M|data} = 0.36$ and that it was 1.42 times more likely to be true than the null model. This suggests that passing the mirror mark test affected children's tendency to place the sticker on their own or their companion's face, to the effect that mirror mark test passers were 1.53 times more likely to place the sticker than mirror mark test non-passers ($\beta = 0.42$, 95% confidence interval [-0.34, 1.33]). Together, these analyses suggest that mirror mark test passers were more likely than non-passers to place the sticker on their own or their caregiver's face, but that sticker placement was not substantially affected by condition in either group.⁵⁴

We repeated this analysis including the same additional predictors as for the sticker choice (see above), as well as whether children chose the matching sticker or not. The analysis showed that the best model of the data was one that included both the factors Mirror Mark Test Passing and Condition, as well as the additional predictors Matching Sticker Choice, Synchronization Index, and Rating of Child Enjoyment: $P_{M|data} = 0.006$, 1.15 times more likely than the null model, and 1.96 times more likely to be true than the model including only the factor Mirror Mark Test Passing.⁵⁵ Matching Sticker Choice acted as a negative predictor of sticker placement, suggesting that children who had chosen the matching sticker were less likely to place it on their own or their caregiver's face compared to children who had chosen the non-matching sticker ($\beta = -0.23$, 95% confidence interval [-1.28, 0.58], $BF_{inclusion} = 1.34$).

⁵⁴ We followed up on these results by looking at sticker placement on their own face and sticker placement on their companion's face separately; the results showed that the model that included only the factor Mirror Mark Test Passing was the best model, given the data, both for sticker placement on their own face ($P_{M|data} = 0.32$, 1.24 times better than the null model) and for sticker placement on their companion's face ($P_{M|data} = 0.35$, 1.34 times better than the null model).

⁵⁵ For a table of β coefficients, see Supplementary Table 8.

Synchronization Index acted as a positive predictor ($\beta = 0.52$, 95% confidence interval [-1.26, 3.91], $BF_{inclusion} = 1.23$), suggesting that the more stable coordination the coordination was during the rhythmical movement task, the more likely children were to place the sticker on their own or their caregiver's face. Rating of Child Enjoyment emerged as a negative predictor, suggesting that less enjoyment of the rhythmical movement task predicted a greater tendency to place the sticker on their own or their caregiver's face ($\beta = -0.13$, 95% confidence interval [-0.66, 0.13], $BF_{inclusion} = 1.62$). This suggests that, at least anecdotally, in addition to the factors Mirror Mark Test Passing and Condition, whether or not children chose the matching sticker, the degree of coordination during the rhythmical movement task, and the child's enjoyment of it affected whether or not children placed the sticker on their own or their caregiver's face.⁵⁶

Imitation Task

Two children's data were excluded from the analysis, as they did not engage with the apparatus at all. The remaining 94 children's data were entered into the analyses.

Latency of First Touch

As the inter-rater agreement of the two raters' coding of the latency of children's first touch of the lamp was only fair at Cohen's $\kappa = 0.4$, we report the respective analysis in the supplementary information.

Head Touch (20s)

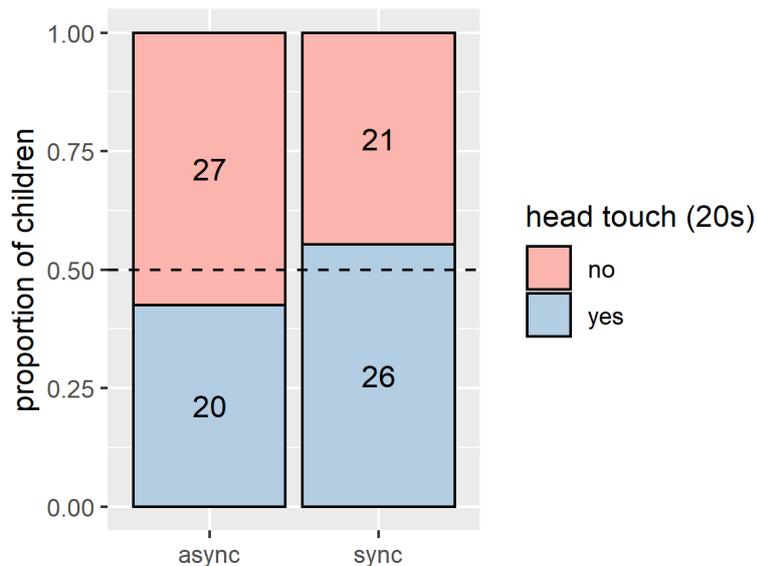
The Bayesian A/B test comparing the proportion of children who either used their head or attempted to use their head to turn on the lamp in the first 20 seconds of the response phase between the synchrony and the asynchrony condition showed moderate evidence in support of H_0 compared to H_+ : $BF_{0+} = 5.39$ and anecdotal evidence for H_+ compared to H_0 : $BF_{+0} = 1.3$,

⁵⁶ The results did not change when gender was included in the analysis as well.

suggesting that slightly more children in the synchrony condition (55.3%) compared to the asynchrony condition (42.6%) used their head in the first 20 seconds of the response phase (see Figure 35), but that the difference was not substantial.

Figure 35

Head Touch (20s) by Condition



Note. This figure depicts the proportion of children who (attempted to) turn on the lamp using their head in the first 20 seconds of the response phase (blue) in the asynchrony (left) and synchrony condition (right). The dashed line marks a proportion of 0.5.

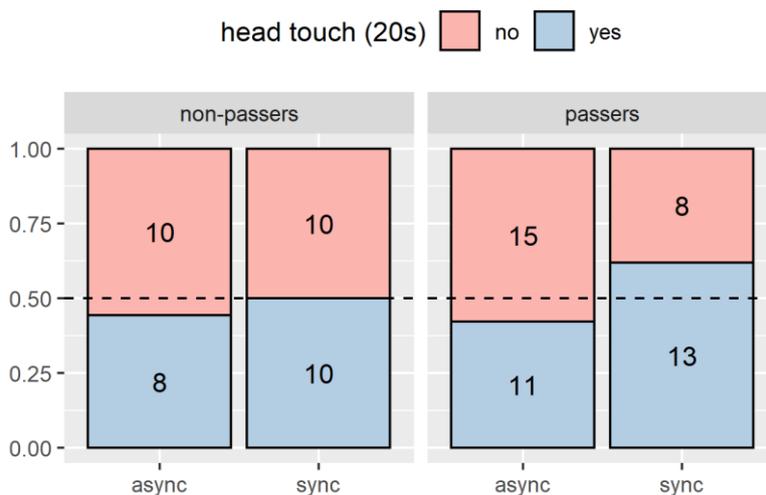
Next, we conducted separate Bayesian A/B tests for mirror mark test passers and non-passers as preregistered comparing the proportion of children who either used their head or attempted to use their head to turn on the lamp in the first 20 seconds of the response phase between conditions.⁵⁷ The Bayesian A/B test comparing mirror mark test passers' head touch behavior in the first 20 seconds of the response phase between the synchrony and the asynchrony condition showed anecdotal evidence for H_+ compared to H_0 ($BF_{10} = 1.76$),

⁵⁷ Of the children included in this analysis, 47 children passed the mirror mark test (26 in the asynchrony condition) and 38 children did not pass the mirror mark test (18 in the asynchrony condition).

suggesting that mirror mark test passers were more likely to use their head to (attempt to) turn on the lamp in the first 20 seconds of the response phase in the synchrony condition (61.9%) compared to the asynchrony condition (42.3%; see Figure 36). However, H_0 that was the best model, given the data ($P_{M|data} = 0.5$), 4.02 times more likely than H_+ . The Bayesian A/B test comparing mirror mark test non-passers' head touch behavior between the synchrony and the asynchrony condition showed that H_0 was the best model of the data ($P_{M|data} = 0.63$), 1.39 times better than H_+ , and 2.27 times better than H_- , suggesting that mirror mark test non-passers' head touch behavior did not substantially differ between conditions.

Figure 36

Head Touch (20s) by Condition and Mirror Mark Test Passing



Note. This figure depicts the proportion of children who (attempted to) turn on the lamp using their head in the first 20 seconds of the response phase (blue), split by mirror self-recognition. Mirror mark test non-passers' data is shown in the left panel; mirror mark test passers' data is shown in the right panel. In each panel, the left bar shows children's behavior in the asynchrony condition, and the right bar shows their behavior in the synchrony condition. The dashed line marks a proportion of 0.5.

To gain a better understanding of the relationship between the effect of synchrony on children's tendency to imitate E1's actions and the development of the self-concept, we additionally followed up on these results in a Bayesian logistic regression, as we did for the sticker choice and sticker placement. The analysis showed that the null model was the best model given the data: $P_{M|data} = 0.42$, 1.94 times better than the next best model, including only the factor Condition.

We followed up on this result by repeating the analysis with the same additional predictors as for the sticker placement, as well as whether children placed the sticker on their own or their caregiver's face. This analysis showed that the best model was one that included the factor Condition, as well as the Synchronization Index, Ratings of Child Enjoyment, and Sticker Placement: $P_{M|data} = 0.003$, 1.46 times more likely than the null model.⁵⁸ Condition acted as a positive predictor of imitative tendencies, suggesting that children in the synchrony condition were more likely to imitate the experimenter's actions compared to children in the asynchrony condition ($\beta = 0.67$, 95% confidence interval [-1.04, 4.06], $BF_{inclusion} = 1.36$). Synchronization Index acted as a negative predictor ($\beta = -0.87$, 95% confidence interval [-5.05, 1.94], $BF_{inclusion} = 1.54$), suggesting that the more stable coordination the coordination was during the rhythmical movement task, the less likely children were to imitate the experimenter's actions. Rating of Child Enjoyment emerged as a positive predictor, suggesting that more enjoyment of the rhythmical movement task predicted a greater tendency to imitate the experimenter's actions ($\beta = 0.11$, 95% confidence interval [-0.2, 0.66], $BF_{inclusion} = 1.41$). Sticker Placement also emerged as a positive predictor, suggesting that children who had placed the sticker on their own or their caregiver's face were more likely to imitate the experimenter's actions ($\beta = 0.11$, 95% confidence interval [-0.49, 0.81], $BF_{inclusion} = 0.95$). This suggests that, anecdotally, the factor Condition affected children's imitation of the

⁵⁸ For a table of β coefficients, see Supplementary Table 9

experimenter's actions, when considered in combination with the degree of coordination, the child's enjoyment of the task, and whether they placed the sticker on their own or their caregiver's face.⁵⁹

Proportion of Head Touches

19 participants' data were excluded for the analysis of the proportion of head touches relative to head and hand touches, as the recording ended before the conclusion of the 60-second response phase (17) or due to caregiver interference (in one case, the task was cut short by the caregiver; in the other, the caregiver placed the apparatus on the floor such that it was no longer visible in the recording). The remaining 74 participants' data were entered into a Bayesian Mann-Whitney U test comparing the proportion of children's head touches, relative to head and hand touches, over the full 60-second response phase between the synchrony and the asynchrony condition. The analysis showed moderate evidence for H_0 : $BF_{01} = 3.71$, suggesting that children did not display a greater proportion of head to hand touches in the synchrony condition (0.24 ± 0.23) compared to the asynchrony condition (0.27 ± 0.3).

Pointing Task

Three participants were excluded from the analysis of the pointing task, as the procedure had not yet been finalized. One additional participant was excluded because they did not complete the task. The remaining 92 participants' data was entered into the analyses.⁶⁰

⁵⁹ Including the factor Gender in the analysis too revealed that the best model of the data was one that included the factor Condition, as well as the additional predictors Synchronization Index, Rating of Child Enjoyment, Sticker Placement, and Gender: $P_{M|data} = 0.002$, 1.48 times more likely than the null model.

⁶⁰ 47 of these children completed the synchrony condition (22 mirror mark test passers, 20 non-passers); 45 complete the asynchrony condition (24 mirror mark test passers, 19 non-passers).

Looking Behavior

We computed a Bayesian generalized linear mixed model (GLMM) with the factors Condition and Mirror Mark Test Passing as fixed effect variables, and with Subject ID as well as Object Number as random effects grouping factors to ensure that participants' looking behavior did not differ between groups. The 95% credible interval for the main effects of Condition ($M = 0.66$, 95% CI $[-0.18, 1.54]$) and Mirror Mark Test Passing ($M = -0.24$, 95% CI $[-1.12, 0.62]$), as well as the interaction between Condition and Mirror Mark Test Passing ($M = -0.28$, 95% CI $[-1.47, 0.87]$) include 0, suggesting that participants' looking behavior did not differ between groups.

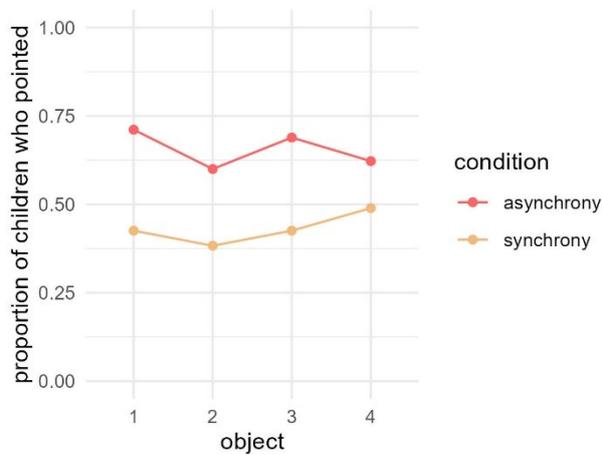
Pointing Behavior

To test for an effect of condition and mirror mark test passing on participants' pointing behavior, we computed a Bayesian generalized linear mixed model (GLMM) on participants' pointing with the factors Condition and Mirror Mark Test Passing as fixed effect variables, and with Subject ID as well as Object Number as random effects grouping factors. The analysis showed credible evidence for a main effect of Condition ($M = 1.18$, 95% CI $[0.1, 2.38]$), with participants in the asynchrony group pointing to more objects (2.62 ± 1.5) than children in the synchrony group (1.73 ± 1.6 ; see

Figure 37). The 95% credible interval for the main effect of Mirror Mark Test Passing ($M = 0.56$, 95% CI $[-0.55, 1.67]$) and for the interaction between Condition and Mirror Mark Test Passing ($M = -0.67$, 95% CI $[-2.26, 0.8]$) both include zero, suggesting a lack of strong evidence for these effects.

Figure 37

Pointing by Condition



Note. This figure depicts the proportion of children (y axis) who pointed at each of the four objects (x axis) in the synchrony condition (yellow) and in the asynchrony condition (red).

Test

57 of the 92 children who completed the pointing task (61.96%) completed the test (73.33% of children in the asynchrony condition; 51.06% of children in the synchrony condition).⁶¹ One of these children had to be excluded due to procedural error; the remaining 56 children's data were entered into the analyses.

The Bayesian A/B tests on children's first look (correct/incorrect), their first reach (correct/incorrect), and on whether they handed E1 the correct object, showed that H_0 was the best model of the data in all cases ($P_{M|data} = 0.66, 0.58, \text{ and } 0.6$, respectively). This suggests that children in the synchrony condition did not look at (39.1%), reach for (52.2%), or hand E1 the correct object (31.3%) substantially more often than children in the asynchrony condition (43.8%, 38.7%, and 42.9%, respectively).

⁶¹ 39 of the 57 children who completed the test did so following the modified procedure, the outcomes of the analyses using only the data from these participants is reported in the supplementary information.

Questionnaire Data

We found extreme evidence for a positive correlation between low intensity pleasure and attentional shifting: $\tau = 0.29$, $BF_{10} = 458.38$, and between low intensity pleasure and cuddliness: $\tau = 0.28$ $BF_{10} = 228.34$, as well as strong evidence for a positive relationship between low intensity pleasure and shyness: $\tau = 0.22$ $BF_{10} = 12$.⁶² As the correlation coefficients were relatively low, we computed the analysis as planned and dropped the predictor low intensity pleasure if the $BF_{inclusion}$ was smaller than 1.

The Bayesian logistic regressions with the factors Condition, Mirror Mark Test Passing, their interaction, and the ECBQ variables Attention Shifting, Cuddliness, Inhibitory Control, Low Intensity Pleasure, Shyness, and Sociability on toddlers' matching sticker choice showed extreme evidence that the model that included the factors Condition, Mirror Mark Test Passing, their interaction, and the ECBQ variables Attentional Shifting, Cuddliness, and Low Intensity Pleasure was the best model of the data: $P_{M|data} = 0.04$, 1,177 times more likely to be true than the null model, and 2.75 times more likely to be true than the model that included only the factors Condition, Mirror Mark Test Passing, and their interaction.⁶³

The Bayesian logistic regressions with the factors Condition, Mirror Mark Test Passing, their interaction, and the ECBQ variables Attentional Shifting, Cuddliness, Inhibitory Control, Low Intensity Pleasure, and Shyness on toddlers' sticker placement showed anecdotal evidence against the inclusion of Low Intensity Pleasure in the model ($BF_{inclusion} = 0.84$), so we recomputed the analysis without this sub-scale. The results showed that the model that included only the factors Mirror Mark Test Passing and Sociability was the best model of the data: $P_{M|data}$

⁶² For a heatmap visualization of all correlation coefficients, see Supplementary Figure 10.

⁶³ As the inclusion of the factor Low Intensity Pleasure only anecdotally improved model fit ($BF_{inclusion} = 1.15$), we recomputed this analysis without this sub-scale. The outcome of this analysis showed that the model included the factors Condition, Mirror Mark Test Passing, their interaction, and the ECBQ variable Attentional Shifting was the best model of the data: $P_{M|data} = 0.05$, 1,166 times more likely to be true than the null model, and 1.8 times more likely to be true than the model that included only the factors Condition, Mirror Mark Test Passing, and their interaction.

= 0.02, anecdotally better than the null model: $BF_{10} = 1.14$, and the model that only included the factor Mirror Mark Test Passing: $BF_{10} = 1.4$.

The Bayesian logistic regressions with the factors Condition, Mirror Mark Test Passing, their interaction, and the ECBQ variables Attentional Shifting, Cuddliness, Inhibitory Control, Low Intensity Pleasure, Shyness, and Sociability on toddlers' head touch behavior in the first 20 seconds of the response phase showed anecdotal evidence against the inclusion of Low Intensity Pleasure in the model ($BF_{inclusion} = 0.72$), so we recomputed the analysis without this sub-scale. The results showed that the null model was the best model of the data: $P_{M|data} = 0.02$, 1.27 times better than the next best model, which included only the factor Cuddliness.

Summary

The outcomes of the analyses including only the factors Condition, Mirror Mark Test Passing and their interaction for the key dependent measures Matching Sticker Choice, Sticker Placement, Head Touch (20s), and Pointing are summarized below (Table 6).

Table 6

Result Summary

Measure	Condition	Mirror Mark Test Passing	Condition x Mirror Mark Test Passing
Matching Sticker Choice	✓	✓	✓
Sticker Placement		✓	
Head Touch (20s)			
Pointing	✓		

Note. This table visualizes the factors (columns) included in the best model of the for the Measures (rows) Matching Sticker Choice, Sticker Placement, and Head Touch (20s), and the factor that showed substantial evidence for Pointing.

Discussion

The present study aimed to investigate whether interpersonal synchrony affects 18-month-old toddlers' motivation to socially align and how this relates to their developing self-concept. Our results do not provide a clear-cut answer to this question but rather suggest that interpersonal (a-)synchrony has a complex relationship with self-other alignment in toddlers who are at the cusp of developing cognitive self-awareness. In particular, we found that after moving asynchronously with another person, toddlers were overall more motivated draw her attention to objects they could not see, and those who showed evidence for cognitive self-awareness were also more likely to demonstrate a shared preference with her.

In the sticker task, toddlers who showed evidence for cognitive self-awareness were more likely to choose the matching sticker following asynchronous movement with the experimenter, suggesting that a conceptual understanding of themselves allowed them to appreciate their ability actively intervene to bring about alignment by choosing the matching sticker. Toddlers who passed the mirror mark test were also more likely to place the sticker on their own or their caregiver's face, consistent with Kamps et al. (2022), further suggesting that there may be a link between the development of a self concept and a greater sense of agency, or a more pronounced awareness of one's ability to bring about changes in one's external environment, including one's own physical appearance. Toddlers who did not show evidence for cognitive self-awareness, by contrast, were less likely to choose the matching sticker overall, and particularly following interpersonal asynchrony, suggesting that they too were sensitive to basic cues to self-other similarity, but that they went along with it rather than actively intervening on self-other alignment.

The imitation task, however, showed the opposite pattern: Toddlers, particularly those who showed evidence of cognitive self-awareness, were anecdotally more likely to align their

own behavior with the experimenter's after bouncing in synchrony with her. This suggests that here, experiencing interpersonal synchrony may have resulted in self-other alignment in the form of social learning from the other. As we did not find evidence for an interaction between passing the mirror mark test and our synchrony manipulation, this suggests that a basic sensitivity to self-other similarity may be enough to boost learning from a similar other. However, considering that the evidence for an effect of interpersonal synchrony was only anecdotal, and stronger in mirror mark test passers, this may suggest that our analysis was underpowered, and that selective imitation based on self-other similarity, as conveyed via interpersonal synchrony, may indeed rely on a conceptual understanding of self.

Toddlers' pointing behavior, however, was more pronounced following interpersonal asynchrony, independent of their self concept development. This suggests that moving asynchronously with another person may have made toddlers more sensitive to differences between self and other overall, including their perspective, similar to the effects of imitation inhibition training (Kampis et al., 2023). Toddlers in the asynchrony condition were not more likely to learn the new labels, suggesting that greater tendency to point was not driven by epistemic motives, but that it was rather the result of a greater sensitivity to the difference between their own and the other's perspective or the manifestation of an attempt to affiliate with the experimenter.

We also found evidence for the inclusion of the synchronization index, our measure of the stability of the coordination between the experimenter and the caregiver-child dyad during the rhythmical movement task, as well as either the caregiver's or the child's enjoyment of the interaction as additional predictors in our model across different tasks. This suggests that beyond the categorical distinction between conditions, toddlers' motivation to align with the experimenter was additionally affected by the quality of the interaction – in particular, by how

stable and enjoyable the interaction was. Furthermore, toddlers' alignment in a given task was influenced by their behavior in the preceding task, suggesting that toddlers' flexibly adapt their alignment behavior based on previous interaction.

In this study, we based our classification of toddlers regarding their developing self concept on their behavior in the mirror mark test (Amsterdam, 1972). While this test has been linked to other markers of the development of cognitive self-awareness (Bulgarelli et al., 2019; Lewis & Ramsay, 2004), it is also possible that factors other than the presence of a self concept affected toddlers' behavior in this task. In particular, whereas actively reaching for the mark on their own face provides an indication that toddlers were successfully able to map the external representation of their face in the mirror onto their own body, the absence of such a behavior does not necessarily mean that toddlers do not yet possess a self concept. For this reason, conclusions about mirror mark test non-passers' behavior should be drawn with caution, and future studies should aim to disentangle the underlying processes that determine whether toddlers pass the mirror mark test.

Our accelerometer data, as well as the video ratings, indicate that, despite the inevitable variation in movement dynamics between caregivers in different sessions of the study, our manipulation was successful, and the caregiver-child dyad's bouncing was more coordinated with the experimenter's in the synchrony condition than in the asynchrony condition. The considerable variability in the stability of coordination between dyads, particularly in the synchrony condition, is, however, also visible in the data (see Figure 5). Given that our measure of the stability of coordination during this task consistently emerged as a relevant predictor of toddlers' subsequent behavior, tapping into different aspects that affect the quality of the interaction, such as the degree or stability of coordination, may be a fruitful avenue for future research.

The fact that the duration of eye contact did not differ between groups suggests that moving in synchrony with another person did not lead to increased attention to or salience of the experimenter, as has previously been proposed (Macrae et al., 2008; Miles et al., 2010). However, our finding that toddlers who show evidence for cognitive self-awareness also showed more direct eye contact during the rhythmical movement task may provide an indication that passing the mirror mark test is linked to greater sensitivity to social cues, in line with its documented link to increased prosocial behavior, imitation, declarative pointing, and pretend play (Goncharova & Ross, 2024). The duration of eye contact during the rhythmical movement task did not emerge as a relevant predictor of toddler's alignment behavior, however, suggesting that our effects were not driven by increased attention to the experimenter during interpersonal (a-)synchrony, nor by ostensive cues.

However, one limitation of the current study is that we did not counterbalance the tempo at which the experimenter and the caregiver-child dyad bounced between participants. This means that children in the synchrony condition not only experienced an experimenter who was bouncing at the same pace as them, the experimenter was also always bouncing at a relatively slow pace. Conversely, children in the asynchrony condition only experienced the experimenter bouncing at a fast pace. We opted for this implantation based on Cirelli et al. (2014, 2016), who did not find differences between different tempos, and because the slower pace was also easier for caregivers to follow. This systematic covariation of bouncing speed and condition, however, made it possible for our coder to accurately guess the experimental condition based only on the experimenter's behavior.

Moreover, we found the experimenter's facial expressions to systematically vary between conditions, with higher combined ratings of how happy, smiley, attentive, interactive, and connected to her bounce partner she seemed in the asynchrony condition compared to the

synchrony condition. Although this could reasonably be assumed to have affected toddlers' motivation to align with the experimenter, this factor did not emerge as a relevant predictor in our analyses of participants' subsequent alignment behavior, suggesting that it did not substantially affect behavioral outcomes. Furthermore, neither the toddlers' nor the caregivers' enjoyment substantially differed between conditions, and the correlation coefficients between the experimenter's facial expressions and these ratings were small (see Supplementary Figure 9), suggesting that the systematic difference in the experimenter's facial expressions between conditions did not substantially affect the quality of the interaction.

Regarding toddlers' sticker choice, we expected to see more matching sticker choices after being bounced in synchrony with the experimenter, despite there being no systematic variation in the original paper (Kampis et al., 2022), due to an effect of interpersonal synchrony on toddlers' motivation to align their sticker preference with their movement partner's. While this primary hypothesis was supported by our pilot data, we also preregistered an alternative hypothesis: If moving asynchronously with another person makes toddlers more sensitive to differences between themselves and the other person, toddlers should be more likely to pick the matching sticker in the asynchrony condition. Although we did not see an effect of interpersonal synchrony overall, we did find evidence for the latter hypothesis in mirror mark test passers.

More specifically, we find evidence for an interaction between self concept development and interpersonal synchrony: Following asynchronous movement with the experimenter, toddlers who pass the mirror mark test show a preference for the matching sticker, whereas toddlers who do not pass the mirror mark test show a preference for the non-matching sticker. This suggests that moving asynchronously with the experimenter highlighted the difference between self and other for both mirror mark test passers and non-passers, but

that they responded differentially. Whereas toddlers who did not show evidence for a self concept appear to simply have followed the cue to self-other difference, showing a moderate preference for the non-matching sticker, toddlers who show evidence for a self concept seem to have actively tried to intervene, showing a pronounced preference for the matching sticker. This could indicate an attempt to repair the relationship with the experimenter (Morton, 2016), or be the result of stronger prosocial tendencies overall (Goncharova & Ross, 2024). However, it might more plausibly be linked to a greater awareness of their own ability to intervene and bring about changes in the environment, or a stronger sense of agency.

As for toddlers' sticker placement, we expected to see a stronger tendency to place the sticker on their own or their caregiver's face following interpersonal synchrony due to greater motivation to socially align. Whereas descriptively, a greater proportion of toddlers placed the sticker on their own or their caregivers face following synchronous compared to asynchronous bouncing, our analysis showed anecdotal evidence for no difference between groups, suggesting that interpersonal synchrony did not affect toddlers' motivation to align their own or their caregiver's physical appearance with the experimenter's. Despite not finding evidence for our main hypothesis in this task, we did find a stronger tendency to place the sticker on their own or their caregiver's face in toddlers who passed the mirror mark test compared to those who did not, replicating the findings of Kamps et al. (2022), and further suggesting that there may be a link between the underlying capacities in these two tasks.

However, the fact that our results with respect to sticker choice differ from those for sticker placement suggests that these two parts of the task likely tap into different processes. In the original paper by Kamps et al. (2022), the authors found a link between mirror self-recognition and sticker placement, suggesting that these may both rely on a common motivational component related to the self in relation to other and that it could be this social

component that drives the relationship between these behaviors. This interpretation fits nicely with our findings regarding toddlers' sticker placement, with mirror self recognition predicting self-other alignment. However, the authors do not report finding a similar pattern in toddlers' sticker choice, suggesting that our findings may be at odds with their data. This may be explained by the fact that in the original paper, toddlers saw their caregiver, with whom they have a very close relationship already, wearing the sticker on their face, whereas in the present study, it was the experimenter, with whom they had only had a short interaction at the beginning of the study. It is possible that in an interaction with a stranger, choosing the matching or non-matching sticker may have acted as more of a communicative signal, conveying a shared preference (or not), than in the context of an interaction with their caregiver.

Beyond this, matching sticker choice acted as a negative predictor of sticker placement, as did the child's enjoyment of the bouncing task, suggesting toddlers who seemed to enjoy the interaction more and toddlers who chose the matching sticker, were less likely to place the sticker on their own or their caregiver's face than those who did not enjoy the interaction and/or chose the non-matching sticker. Kamps et al. (2022) did not find toddlers who chose the matching sticker to show a stronger tendency to place the sticker on their own face than not place the sticker at all, but the authors do not report whether toddlers' sticker choice was predictive of their sticker placement overall. Though neither of these was a strong independent predictor (see Supplementary Table 8), this further suggests that the two parts of the task measure likely different constructs, and that different processes may be involved in toddlers' motivation and decisions to signal a shared preference or engage in self-other alignment.

It is also possible that the second phase of the sticker task, once toddlers had made their sticker choice, allowed for more self-determined action on the part of the child, as they were only prompted to do something with the sticker after exploring it on their own for 60 seconds.

It is therefore possible that toddlers who found greater enjoyment in self-directed exploration may have been more likely to place the sticker on themselves or their caregiver, and may have been less likely to enjoy the rhythmical movement task, which they primarily took part in as a patient. The relevance of the child's enjoyment of the rhythmical movement task and their sticker choice is further highlighted by the fact that, in conjunction with these two negative predictors, the synchronization index emerged as a positive predictor, while condition was found to be a negative predictor of sticker placement, as in the case of sticker choice. Although none of these additional factors emerged as strong additional predictors, this finding nonetheless seems to suggest that there is a relevant link between the (lack of) enjoyment of the rhythmical movement interaction and toddlers' motivation to socially align.

With respect to toddlers' imitation behavior, we expected toddlers to show a stronger tendency to faithfully imitate the experimenter's actions following interpersonal synchrony due to increased affiliation with her. We found anecdotal evidence for this hypothesis, with mirror mark test passers showing a somewhat more pronounced, but still anecdotal, effect than non-passers. This appears to sit well with the literature on the effects of interpersonal synchrony on affiliation in adults (Cross et al., 2019a), which documents effects in most but not all studies. This suggests that this effect may be smaller than previously assumed and require relatively large samples to detect. It is thus possible that there may indeed be a link between toddlers' developing self-awareness and their affiliative or social learning tendencies following interpersonal synchrony. However, further research is needed for more conclusive results.

As in the case of sticker placement, participants' imitation behavior was predicted by the child's enjoyment of the task, the synchronization index, and condition, as well as the child's behavior in the previous task. However, unlike the sticker task, participants' imitative tendencies were negatively predicted by the synchronization index and positively predicted by

condition (and the child's enjoyment of the rhythmical movement task). This suggests that less stable coordination, being in the synchrony condition, and enjoying the rhythmical movement interaction, together with having placed the sticker on themselves or their caregiver (despite anecdotal evidence against the inclusion of this predictor, see Supplementary Table 9), predicted toddlers' tendency to faithfully imitate the experimenter's actions in the imitation task.

Regarding toddlers' pointing behavior, we expected to see a stronger tendency to point at objects the experimenter could not see following interpersonal synchrony. We reasoned that there could be two mechanisms behind this. If toddlers are more likely to represent the experimenter as a reliable source of information after moving in synchrony with her, then we reasoned that they should be more likely to point and they should better remember the labels she provides in the synchrony condition compared to the asynchrony condition. However, if toddlers are simply motivated to draw the experimenter's attention to an object she can't see due to increased affiliation with her following interpersonal synchrony, they should also be more likely to point following the synchrony compared to the asynchrony condition, but we should not have reason to predict systematic differences in toddlers' retention of the object labels.

However, our data suggest that our initial hypothesis was incorrect. Toddlers' pointing behavior systematically differed between condition but was more prevalent in the asynchrony condition compared to the synchrony condition. As toddlers' retention of the object labels was low, this suggests that toddlers who were bounced asynchronously with the experimenter were more motivated to draw her attention to the objects. The fact that toddlers' looking behavior did not differ between conditions suggests that this effect was not driven by increased attention to, and by extension, interest in the novel objects. Instead, it is likely that moving

asynchronously highlighted differences between self and other, acting as a kind of training in self-other distinction (Santesteban et al., 2012), and thereby making participants more sensitive to the difference between their own and the experimenter's perspective (Kampis et al., 2023).

It is worth noting that our study is not the first to find different effects of interpersonal synchrony on different outcome variables. For instance, Rabinowitch and Meltzoff found that in four-year-old children, moving either synchronously or asynchronously with a peer elicits more generous allocation of resources than not moving together at all (Rabinowitch & Meltzoff, 2017a) but that moving in synchrony with a peer gives rise to more successful cooperation and greater intentional communication between peers than asynchronous or no movement at all (Rabinowitch & Meltzoff, 2017b). Similarly, although studies with 14-month olds found more helping behavior after being bounced in synchrony with an experimenter than after asynchronous bouncing (Cirelli et al., 2014a), a study investigating the effect of interpersonal synchrony on novel word learning in two and a half year old children did not find evidence for an effect (Bazhydai et al., 2022). Moreover, like the present study, the aforementioned studies used passive movement paradigms, in which children were moved in synchrony or asynchronously with another person. It is possible that being a patient rather than an actively contributing member of the synchronous or asynchronous interaction weakened the effect of the manipulation and made other aspects of the interaction more salient.

Together, our findings suggest that interpersonal synchrony affects toddlers' self-other alignment in multifaceted ways, and that its effects are related to toddlers' emerging cognitive self-awareness. We found interpersonal (a-)synchrony to highlight self-other similarity or lack thereof to toddlers across tasks. Whereas this cue was sufficient to guide toddlers' pointing behavior, it was found to interact with self concept development in other tasks. Toddlers'

responses to this cue in the sticker task appear to also have been driven by an additional factor linked to cognitive self-awareness, allowing them to intervene and signal a shared preference to their movement partner, or not, and maintain the status quo. While this suggests that the processes underlying the effects of interpersonal synchrony may not be unitary, the present study presents a first step into their systematic investigation. We conclude that interpersonal (a-)synchrony can highlight self-other similarity or lack thereof to toddlers, which is an important social cue that affects how they align with others, communicate and learn from them, especially once they are able to conceive of their own self.

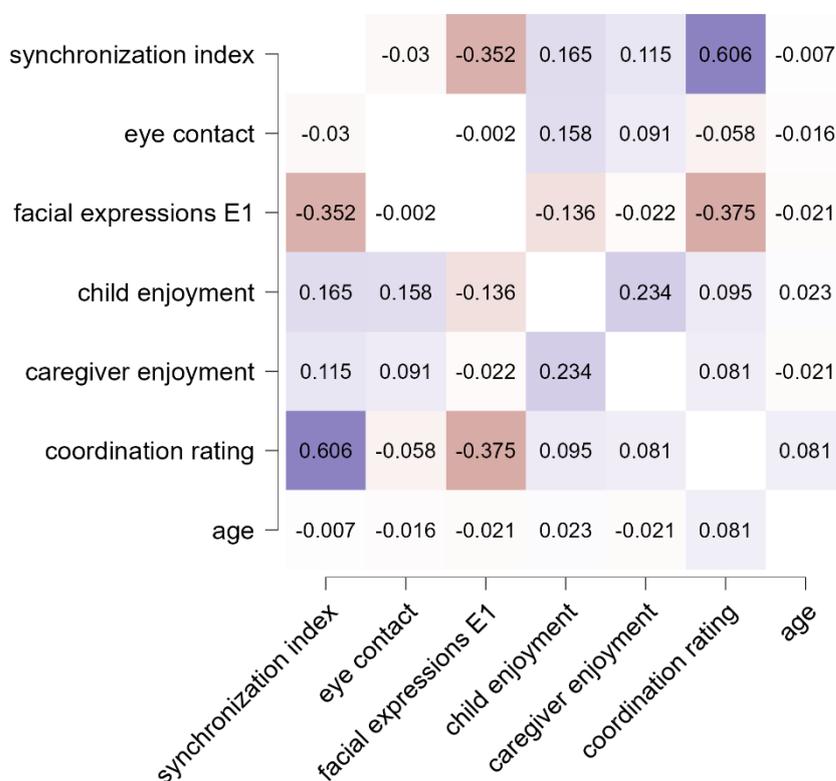
Supplementary Information

Results

Rhythmical Movement Task

Supplementary Figure 9

Heatmap Visualization of Correlation of Rhythmic Coordination Task Variables and Age



Note. This figure depicts the correlation (Kendall's τ) between variables measured during the rhythmic coordination task: the synchronization index, the duration of eye contact, the ratings of E1's facial expressions, the ratings of the child's enjoyment of the task, the ratings of the caregiver's enjoyment of the task, the ratings of coordination, and the child's age. Different shades of purple visualize positive correlations; different shades of brown visualize negative correlations. The darker the shade, the stronger the correlation.

Sticker Task

Sticker Choice

Supplementary Table 7

Logistic Regression: β Coefficients of Sticker Choice Predictors

Coefficient	BF_{inclusion}	Mean \pm SD	95% Credible Interval
<i>Intercept</i>	1	-7.19 \pm 15.96	[-56.61, 17.64]
<i>Condition (synchrony)</i>	2.11	-0.69 \pm 1.39	[-3.98, 1.76]
<i>Mirror Mark Test Passing (1)</i>	341.57	2.02 \pm 0.86	[0.37, 3.86]
<i>Synchronization Index</i>	1.07	0.85 \pm 1.77	[-1.85, 5.2]
<i>Eye Contact</i>	0.58	0.0000003 \pm 0.000007	[-0.00002, 0.00002]
<i>E1 Facial Expressions</i>	0.67	0.04 \pm 0.19	[-0.44, 0.51]
<i>Child Enjoyment</i>	0.65	-0.04 \pm 0.19	[-0.52, 0.41]
<i>Caregiver Enjoyment</i>	2.46	0.16 \pm 0.18	[-0.14, 0.5]
<i>Age</i>	0.61	0.01 \pm 0.02	[-0.03, 0.07]
<i>Condition (synchrony): Mirror Mark Test Passing (1)</i>	3.19	-0.5 \pm 0.98	[-2.97, 0.78]

Note. This table depicts the BF_{inclusion} (column 2), the mean β coefficient \pm its standard deviation (column 3) and the 95% credible interval (column 4) for each coefficient included as a predictor (listed in column 1) in the logistic regression on toddlers' matching sticker choice, as well as the intercept.

Sticker Placement

Supplementary Table 8

Logistic Regression: β Coefficients of Sticker Placement Predictors

Coefficient	BF_{inclusion}	Mean \pm SD	95% Credible Interval
<i>Intercept</i>	1	1.01 \pm 13.81	[-31.1, 32.28]
<i>Condition (synchrony)</i>	1	-0.24 \pm 0.83	[-2.08, 1.51]
<i>Mirror Mark Test Passing (1)</i>	1.77	0.36 \pm 0.53	[-0.58, 1.53]
<i>Synchronization Index</i>	1.23	0.52 \pm 1.2	[-1.26, 3.91]
<i>Eye Contact</i>	0.81	-0.0000009 \pm 0.000005	[-0.00001, 0.000009]
<i>E1 Facial Expressions</i>	0.62	-0.01 \pm 0.11	[-0.26, 0.28]
<i>Child Enjoyment</i>	1.62	-0.13 \pm 0.2	[-0.66, 0.13]
<i>Caregiver Enjoyment</i>	0.77	0.004 \pm 0.07	[-0.14, 0.18]
<i>Age</i>	0.76	0.0005 \pm 0.02	[-0.04, 0.04]
<i>Matching Sticker Choice (1)</i>	1.34	-0.23 \pm 0.46	[-1.28, 0.58]
<i>Condition (synchrony): Mirror Mark Test Passing (1)</i>	1.12	0.44 \pm 0.36	[-0.74, 1.09]

Note. This table depicts the BF_{inclusion} (column 2), the mean β coefficient \pm its standard deviation (column 3) and the 95% credible interval (column 4) for each coefficient included as a predictor (listed in column 1) in the logistic regression on toddlers' sticker placement, as well as the intercept.

Imitation Task

Latency of First Touch

The Bayesian Mann-Whitney U test comparing the latency of children's first touch of the lamp, relative to their initial engagement with the apparatus, between the synchrony and the asynchrony condition showed moderate evidence for H_0 : $BF_{01} = 3.65$, suggesting that children in the synchrony condition were not more eager to engage with the lamp ($1.94 \pm 4.1s$) than children in the asynchrony condition ($3.54 \pm 8.11s$).

Head Touch (20s)

Supplementary Table 9

Logistic Regression: β Coefficients of Head Touch (20s) Predictors

Coefficient	BF_{inclusion}	Mean \pm SD	95% Credible Interval
<i>Intercept</i>	1	-5.63 \pm 15.13	[-41.33, 25.47]
<i>Condition (synchrony)</i>	1.36	0.67 \pm 1.23	[-1.04, 4.06]
<i>Mirror Mark Test Passing (1)</i>	0.76	0.03 \pm 0.34	[-0.71, 0.84]
<i>Synchronization Index</i>	1.54	-0.87 \pm 1.72	[-5.05, 1.94]
<i>Eye Contact</i>	0.88	0.000001 \pm 0.000005	[-0.000009, 0.00001]
<i>E1 Facial Expressions</i>	0.98	-0.05 \pm 0.14	[-0.36, 0.2]
<i>Child Enjoyment</i>	1.41	0.11 \pm 0.21	[-0.2, 0.66]
<i>Caregiver Enjoyment</i>	1.18	0.04 \pm 0.09	[-0.1, 0.26]
<i>Age</i>	0.83	0.003 \pm 0.02	[-0.03, 0.04]
<i>Matching Sticker Choice (1)</i>	0.77	-0.23 \pm 0.29	[-0.68, 0.69]
<i>Sticker Placement (1)</i>	0.95	0.11 \pm 0.3	[-0.49, 0.81]
<i>Condition (synchrony): Mirror Mark Test Passing (1)</i>	0.85	-0.02 \pm 0.32	[-0.94, 0.74]

Note. This table depicts the BF_{inclusion} (column 2), the mean β coefficient \pm its standard deviation (column 3) and the 95% credible interval (column 4) for each coefficient included as a predictor (listed in column 1) in the logistic regression on toddlers' head touches, as well as the intercept.

Pointing Task

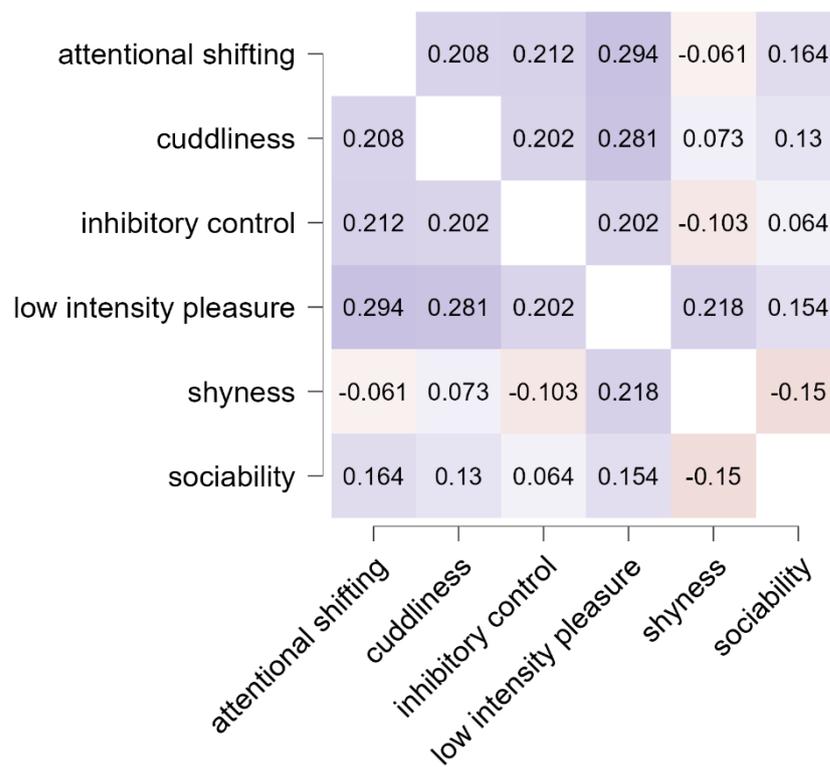
Test

39 of the 57 children who completed the test did so following the modified procedure (22 in the asynchrony condition, 17 in the synchrony condition), according to which E1 displayed both objects in the middle, in the child's line of sight, before asking for one of them. The Bayesian A/B tests on these children's first look (correct/incorrect), their first reach (correct/incorrect), and on whether they handed E1 the correct object, showed that H_0 was the best model of the data in all cases ($P_{M|data} = 0.49, 0.58, \text{ and } 0.5$, respectively). This suggests that children in the synchrony condition did not look at (35.3%), reach for (58.8%), or hand E1 the correct object (31%) substantially more often than children in the asynchrony condition (57.1%, 45%, and 53.3%, respectively).

Questionnaire Data

Supplementary Figure 10

Heatmap Visualization of Correlation of ECBQ Variables



Note. This figure depicts the correlation (Kendall's τ) between the ECBQ variables attentional shifting, cuddliness, inhibitory control, low intensity pleasure, shyness, and sociability. Different shades of purple visualize positive correlations; different shades of brown visualize negative correlations. The darker the shade, the stronger the correlation.

CHAPTER SIX: GENERAL DISCUSSION

The present thesis set out to investigate self-other relations in interpersonal synchrony. To do so, Chapters 2 and 3 aimed to test the hypothesis that moving in synchrony with others leads to an increase in motor-level self-other overlap but did not find evidence for it. Instead, the findings reported in these chapters demonstrate that interpersonal synchrony affects only perceived self-other overlap, and to varying degrees. Chapter 4 then moved on to examine the experience of interpersonal synchrony, finding that its temporal unfolding can be characterized as following a generic diachronic structure made up of three phases, with some variations of this structure across participants and conditions. Finally, Chapter 5 investigated interpersonal synchrony in development, investigating its effects on self-other alignment and its relation with the development of a self-concept, finding evidence for a complex relationship between these factors. On the one hand, 18-month-olds who passed the mirror mark test, and thus showed evidence for cognitive self-awareness, showed a slightly stronger tendency to align with their movement partner after interpersonal synchrony with respect to their imitation behavior and physical appearance. On the other hand, these toddlers clearly showed more aligned preferences with their movement partner following asynchronous movement, suggesting that interpersonal (a-)synchrony highlighted the (dis-)similarity between self and other across tasks but that toddlers' responses differed both between tasks and with their self-concept development.

More specifically, Chapters 2 and 3 aimed to test for an effect of interpersonal synchrony on motor-level self-other overlap using an automatic imitation paradigm. Despite complications due to spatial compatibility in Chapter 2 and failing to elicit an imitation effect in initial versions in both Chapter 2 and Chapter 3, both studies ultimately were able to elicit an effect of imitative compatibility and show moderate evidence against a modulation of

imitative tendencies following interpersonal synchrony. This is at odds with previous findings (Cross, Atherton, et al., 2021a), suggesting that particular task characteristics that differed between our studies and that of Cross et al. (2021), such as the difference in apparatus or automatic imitation task stimuli may have affected our results. Other possible reasons for the discrepant results could be that our analyses may have been underpowered, or that interpersonal synchrony does not have a consistent effect on imitative tendencies.

Chapter 2 additionally included a measure of helping behavior, in order to better situate our findings in the broader context of research on interpersonal synchrony. We administered a task similar to tasks successfully used in previous studies (Kokal et al., 2011; Stupacher et al., 2016). However, we did not find an increase in prosocial behavior following interpersonal synchrony. This goes against a large body of literature documenting prosocial effects of interpersonal synchrony (Mogan et al., 2017a) and may indicate that our implementation of the task was not well suited to test for prosocial behavior directed towards participants' task partner. Alternatively, the findings could indicate that results demonstrating increased prosociality following interpersonal synchrony are not as robust as they have been made out to be.

Both Chapter 2 and 3 also included an additional measure of perceived self-other overlap: a modified version of the Inclusion of Other in Self (IOS) scale. While the different experiments presented in these chapters did not find consistent evidence that interpersonal synchrony elicits substantially greater perceived self-other overlap than asynchronous movement, they did consistently find an increase in perceived self-other overlap following interpersonal synchrony. It is worth noting, however, that across studies participants' ratings of perceived self-other overlap were below 50% on average, ranging from 19 to 45%. These ratings differ from those that have previously been documented in the literature (e.g., Tarr et

al., 2018; Sadaphal et al., 2025), which typically fall onto the upper half of the scale. This may have been an effect of our anchoring of the scale at 50% – participants were shown the two circles with 50% overlap at the start of each rating and may thus have been biased to move them further apart. However, since others have also found lower ratings overall, for instance in the context of VR (Biswas & Brass, 2024), this may also indicate that our manipulation did not elicit a pronounced sense of perceived self-other overlap but only a small increase compared to baseline.

Participants' ratings of perceived self-other overlap following asynchronous movement overall were descriptively lower than those following interpersonal synchrony, ranging from 6% to 37%, but the difference was only significant in Experiment 1 in Chapter 3. Overall, participants' ratings of perceived self-other overlap following asynchronous movement were not consistent in their direction compared to baseline, with Experiment 1 in Chapter 2 finding a decrease in perceived self-other overlap, while Experiments 1 and 2 in Chapter 3 found a (numerical) increase. However, the modification of the scale in Chapter 3, which allowed participants to move the circles apart as well as closer together, showed that participants' baseline ratings of perceived self-other overlap were negative on average, whereas their ratings following interpersonal asynchrony stayed consistently positive on average. It is thus possible that the negative change compared to baseline is the result of the inadequacy of the scale used in Chapter 2 for capturing participants' baseline level of perceived self-other overlap.

Chapter 3 additionally introduced a continuous measure of connectedness during the coordination task to assess participants' experience of connectedness while coordinating. Participants' ratings of connectedness in Experiment 2 of Chapter 3, performed using an expression pedal similar to (Goupil et al., 2021), showed that participants did indeed experience greater connectedness during interpersonal synchrony compared to asynchronous movement.

This suggests that despite feeling more connected during the coordination task when moving in synchrony with their task partner, this did not translate to our measures conducted after the interaction, with participants failing to show substantially greater motor-level or perceived self-other overlap following synchronous compared to asynchronous movement. This finding does, however, confirm that our manipulation did in fact elicit different feelings of connectedness, and that interpersonal synchrony does affect people's experience of interpersonal closeness, while also showing that this experience may be short-lived or too fleeting to persist in the context of a different task.

The investigation of people's experience during interpersonal synchrony was more explicitly thematized in Chapter 4, in which participants completed a short coordination task modeled off the ones used in previous chapters and were subsequently interviewed about their experience during this task. Whereas we found a consistent diachronic structure that described participants' experience of interpersonal coordination, as well as some structural and individual variations, we did not find a pronounced difference between participants' experience of synchronous compared to asynchronous coordination. This may in part be due to participants' suboptimal behavioral performance, given that their movements were not found to be substantially more coordinated in the synchrony compared to the asynchrony condition. At the same time, it is also consistent with our earlier findings that interpersonal synchrony did not substantially affect our behavioral outcome variables.

Chapter 5 then took a different direction, turning to interpersonal synchrony in development. In particular, this chapter examined whether interpersonal synchrony affects 18-month-olds' motivation to socially align and how this relates to their developing self-concept. We found that after moving synchronously with a stranger, toddlers were descriptively more likely to align their appearance with their synchronous movement partner, and they were also

descriptively more likely to imitative her actions. These effects were more pronounced in toddlers who showed evidence for cognitive self-awareness, but they were not substantial. This is consistent with the findings reported above in that the effects of interpersonal synchrony are likely less substantial than has previously been assumed.

With respect to asynchronous movement, we found more substantial effects in two of our measures: Toddlers were overall more likely to point and draw the experimenter's attention to unknown objects that she couldn't see after moving asynchronously with her compared to synchronously, and toddlers who showed evidence for cognitive self-awareness showed clear evidence of a shared preference with the experimenter following asynchronous movement with her. This aligns with our finding from Chapter 3 that asynchronous movement may have a positive effect on the relation between self and other, which has also been found in older children (Rabinowitch & Meltzoff, 2017a). It also sits well with the experiential reports in Chapter 4, which did not substantially differ between the synchrony and asynchrony condition, suggesting that participants' relationship with their task partner was affected in similar ways by both forms of interpersonal coordination. However, the finding is still surprising in light of previous results in younger children and suggests that more research is needed to untangle the effects of synchronous and asynchronous movement on toddlers' self-other alignment, particularly in relation to their self-concept development.

Tying these findings back to our initial definition of interpersonal synchrony, it is worth noting that the paradigm used for our synchrony manipulation in Chapter 5 does not entirely align with it. In particular, while Chapters 2-4 used paradigms, in which both co-actors actively contributed to the interaction, aiming to perform the same movements at the same time as the other person, toddlers in Chapter 5 passively participated in the interaction without actively contributing to the interpersonal (a-)synchrony. While this may have affected our results, it is

not clear what a paradigm involving more active participation on the toddler's part would have looked like, as both paradigms with slightly younger children (e.g., Cirelli et al., 2014; Tunçgenç et al., 2015) and paradigms with slightly older children (e.g., Bazhydai et al., 2022; Rabinowitch & Meltzoff, 2017) have typically relied on synchrony manipulations involving passive movement.

Moreover, it is also worth addressing the involvement of the experimenter in this study. It has previously been argued that the effects of interpersonal synchrony may be reducible to expectancy effects, resulting from top-down expectations (Atwood et al., 2022). However, as Tunçgenç et al. (2023) have argued, this need not necessarily threaten inferential validity, particularly given that expectations about interpersonal synchrony are present from a very young age (Fawcett & Tunçgenç, 2017). In the context of our study in Chapter 5, we found a difference between the experimenter's facial expressions in the two conditions but did not find this variable to affect any of our outcome measures. Thus, while this does not rule out that the presence and involvement of the experimenter may have biased toddlers' responses, it does suggest that their behavior cannot be reduced to the experimenter's facial cues during the experimental manipulation and more broadly could be taken to support the idea that the effects of interpersonal synchrony should not be reduced to simple expectancy effects.

With respect to the mechanisms outlined in the introduction, the studies presented in this thesis do not support the idea that interpersonal synchrony increases affiliation via an increase in motor-level self-other overlap. In particular, we did not find evidence for a modulation of motor-level self-other overlap in Chapters 2 and 3, and we also did not find evidence for an increase in motor-level self-other overlap following interpersonal synchrony in Chapter 3. Moreover, unlike Hove & Risen (2009) we also did not find substantial evidence for a difference in participants' liking of their task partner between conditions. Together, these

findings suggest that in the context of the studies presented in this thesis, interpersonal synchrony did not affect affiliation via an increase in motor-level self-other overlap.

We did not manipulate the presence of a shared goal in any of our studies. However, one crucial difference between the coordination tasks used in Chapter 4 and those used in previous chapters is the more explicit nature of the presence of a shared goal in the asynchrony condition. Whereas participants in Chapters 2 and 3 were instructed to ‘stick to their own pace’ in the asynchrony condition, participants in Chapter 4 were told to tap at different tempos. This difference was not only reflected in higher coordination demands, as participants were not given a predefined tempo to move at in Chapter 4, but it also imposed a shared coordination goal beyond simply sticking to their own pace. The presence of a shared goal in both the synchrony and asynchrony condition of Chapter 4 may thus have framed the coordination task in both conditions as a joint action involving similar coordination demands and may therefore have given rise to qualitatively similar experiences in both conditions of the task.

Failing to find evidence for an account of interpersonal synchrony that centers around close perception-action links, we instead proposed an interpretation of our results in Chapters 2 and 3 in terms of group membership. In particular, we interpreted the finding that moving in synchrony with another person affected perceived but not motor-level self-other overlap as an indication that interpersonal synchrony affects people’s representation of themselves in relation to their movement partner at a more conceptual level, and not simply in terms of motor representations. Following Cross et al. (2019), we took this interpretation to be consistent with the idea that interpersonal synchrony leads to a social recategorization of the self in relation to the other in terms of group membership.

While similar accounts have been proposed for interpersonal synchrony in development (Cirelli, 2018b), our data in this respect is less clear. We found anecdotal evidence for a

modulation of toddlers' imitative tendencies following interpersonal synchrony in a task that has previously been shown to be affected by group membership (Altınok et al., 2022b), suggesting a similar mechanism may have been at work in our study. However, group membership has also been shown to affect toddlers' preferences (Choi & Luo, 2023), and toddlers have been shown to be discriminate in their pointing behavior (Begus & Southgate, 2012b), which does not line up neatly with our findings. Together, this suggests toddlers' motivation to socially align following interpersonal synchrony may in part be driven by synchrony acting as a cue to social group membership, but that there are also other processes at work. Future investigations may focus on untangling the complex set of developments that occur during the second year of life and on determining the role of interpersonal synchrony within it.

While Chapter 4 does not speak to interpersonal synchrony acting as a cue to group membership directly, it does raise other important topics and points to interesting avenues for future research. One interesting finding concerned participants' coordination dynamics with respect to mutual or unidirectional adaptation. In contrast to research suggesting that expert musicians may flexibly adapt to novices in order to achieve better coordination (Wolf et al., 2018), we found that in dyads with a greater asymmetry in terms of their experience with music, the member with music experience chose not to adapt to their novice partner. As the participants in our study indicated varying levels of experience with music, this may suggest that there are intermediate steps between experienced musicians and experts that affect the coordination patterns they use in interactions with novices in order to achieve better coordination. In particular, this may suggest that a greater skillset is required in order to adapt to an unstable novice (i.e., expertise), whereas experience with music may highlight differences in competence or stability with respect to timing of movements but may not suffice to enable

people to adjust to an unstable novice. More broadly, it may also suggest that people's stances with respect to adaptation are affected by their level of expertise in a particular type of domain.

Another interesting finding concerns mutual adaptation and its relationship with joint improvisation. In particular, the dyads who reported a phase of experimenting in their experience were those who also reported phases of mutual adaptation prior to this. This suggests that mutual adaptation may not only be a key component of joint improvisation but also a necessary condition for it to occur. In a similar vein, the fact that only half of the participants in the study reported on a phase of experimenting may also suggest that there are relevant individual differences between people who spontaneously engage in joint improvisation and those who do not. While there have been some attempts to investigate these kinds of questions (Fairhurst et al., 2014; Feniger-Schaal et al., 2016; Washburn et al., 2021), this field of research is still in its beginnings and may yield important insight for our understanding of individual differences in interpersonal coordination.

Finally, Chapter 4 confirmed previous observations about the literature on interpersonal synchrony in lab settings compared to more naturalistic environments (Mogan et al., 2017a) in that it produced a very procedural, action-based description of the temporal unfolding of participants' experiences of interpersonal coordination, with little mention of affective states. This suggests that research on interpersonal synchrony should aim to understand the differences between the phenomena investigated in the laboratory and the forms of interpersonal synchrony people experience in their everyday lives. While researchers have begun to bridge this gap by conducting studies in more naturalistic settings (Sun et al., 2023), there is still much work to be done.

In line with this, it is also worth noting that the majority of research on interpersonal synchrony has focused on neurotypical adults. While there is a growing body of research on

interpersonal synchrony in development, there are still many unanswered questions about the role of interpersonal synchrony in early human relationships, particularly in relation to their self concept development. Similarly, researchers have begun to focus on neurodivergent populations, especially autism (McNaughton & Redcay, 2020; Bowsher-Murray et al., 2022; Carnevali et al., 2024), as well as its potential as a behavioral intervention (Pardo-Olmos et al., 2025), but the consideration of diverse populations and contexts still has much room to grow.

Overall, this thesis has aimed to investigate interpersonal synchrony in self-other relations by conducting lab experiments in adults and toddlers to test how it affects our behavior, as well as phenomenological interviews to gain a better understanding of the associated experiences. We found that interpersonal synchrony does not modulate self-other overlap with one's interaction partner at the motor level but that it does affect perceived self-other overlap. We interpret this finding to suggest that interpersonal synchrony acts as a cue to group membership by prompting a social recategorization of the self in relation to one's movement partner(s). Moreover, we have highlighted that the context in which interpersonal synchrony occurs may affect the way people experience the interaction and proposed that future research should focus on bridging the gap between investigating interpersonal synchrony in lab settings and more naturalistic environments. Finally, we have demonstrated the complex interaction between interpersonal synchrony and the development of the self-concept and have emphasized the need for more research in this domain.

BIBLIOGRAPHY

- Agostinelli, C., & Lund, U. (2023). *R package "circular: Circular Statistics (version 0.5-0)*.
<https://CRAN.R-project.org/package=circular>
- Altınok, N., Király, I., & Gergely, G. (2022a). The Propensity to Learn Shared Cultural Knowledge from Social Group Members: Selective Imitation in 18-month-olds. *Journal of Cognition and Development, 23*(2), 273–288.
<https://doi.org/10.1080/15248372.2021.1966013>
- Altınok, N., Király, I., & Gergely, G. (2022b). The Propensity to Learn Shared Cultural Knowledge from Social Group Members: Selective Imitation in 18-month-olds. *Journal of Cognition and Development, 23*(2), 273–288.
<https://doi.org/10.1080/15248372.2021.1966013>
- Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology, 5*(4), 297–305. <https://doi.org/10.1002/dev.420050403>
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*, 596–612. <https://doi.org/10.1037/0022-3514.63.4.596>
- Atwood, S., Schachner, A., & Mehr, S. A. (2022). Expectancy Effects Threaten the Inferential Validity of Synchrony-Prosociality Research. *Open Mind, 6*, 280–290.
https://doi.org/10.1162/opmi_a_00067
- Bass-Krueger, J. L., Wiedemann, E. G., & Demšar, E. (2024). Comparing Two Methods for Exploring Consciousness: Descriptive Experience Sampling and Micro-Phenomenological Interviews. *Interdisciplinary Description of Complex Systems, 22*(1), 84–106. <https://doi.org/10.7906/indecs.22.1.5>

- Bazhydai, M., Ke, H., Thomas, H., Wong, M. K. Y., & Westermann, G. (2022). Investigating the effect of synchronized movement on toddlers' word learning. *Frontiers in Psychology, 13*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.1008404>
- Begus, K., & Southgate, V. (2012a). Infant pointing serves an interrogative function. *Developmental Science, 15*(5), 611–617. <https://doi.org/10.1111/j.1467-7687.2012.01160.x>
- Begus, K., & Southgate, V. (2012b). Infant pointing serves an interrogative function. *Developmental Science, 15*(5), 611–617. <https://doi.org/10.1111/j.1467-7687.2012.01160.x>
- Bischof-Köhler, D. (2012). Empathy and Self-Recognition in Phylogenetic and Ontogenetic Perspective. *Emotion Review, 4*(1), 40–48. <https://doi.org/10.1177/1754073911421377>
- Biswas, M., & Brass, M. (2024). Syncing online: A methodological investigation into movement synchrony, proxemics, and self-other blurring in virtual spaces. *PLOS ONE, 19*(10), e0308843. <https://doi.org/10.1371/journal.pone.0308843>
- Bowsher-Murray, C., Gerson, S., von dem Hagen, E., & Jones, C. R. G. (2022). The Components of Interpersonal Synchrony in the Typical Population and in Autism: A Conceptual Analysis. *Frontiers in Psychology, 13*. <https://doi.org/10.3389/fpsyg.2022.897015>
- Bratman, M. E. (1993). Shared Intention. *Ethics, 104*(1), 97–113. <https://doi.org/10.1086/293577>
- Bulgarelli, C., Blasi, A., de Klerk, C. C. J. M., Richards, J. E., Hamilton, A., & Southgate, V. (2019a). Fronto-temporoparietal connectivity and self-awareness in 18-month-olds: A resting state fNIRS study. *Developmental Cognitive Neuroscience, 38*, 100676. <https://doi.org/10.1016/j.dcn.2019.100676>

- Bulgarelli, C., Blasi, A., de Klerk, C. C. J. M., Richards, J. E., Hamilton, A., & Southgate, V. (2019b). Fronto-temporoparietal connectivity and self-awareness in 18-month-olds: A resting state fNIRS study. *Developmental Cognitive Neuroscience*, *38*, 100676. <https://doi.org/10.1016/j.dcn.2019.100676>
- Buttelmann, D., & Zmyj, N. (2020). Fourteen-month-olds' imitation is influenced more strongly by a model's competence than by a model's certainty. *Infant Behavior and Development*, *60*, 101458. <https://doi.org/10.1016/j.infbeh.2020.101458>
- Buttelmann, D., Zmyj, N., Daum, M., & Carpenter, M. (2013). Selective Imitation of In-Group Over Out-Group Members in 14-Month-Old Infants. *Child Development*, *84*(2), 422–428. <https://doi.org/10.1111/j.1467-8624.2012.01860.x>
- Carnevali, L., Valori, I., Mason, G., Altoè, G., & Farroni, T. (2024). Interpersonal motor synchrony in autism: A systematic review and meta-analysis. *Frontiers in Psychiatry*, *15*. <https://doi.org/10.3389/fpsy.2024.1355068>
- Catmur, C., & Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 409–421. <https://doi.org/10.1037/a0019325>
- Cebioğlu, S., & Broesch, T. (2021). Explaining cross-cultural variation in mirror self-recognition: New insights into the ontogeny of objective self-awareness. *Developmental Psychology*, *57*(5), 625–638. <https://doi.org/10.1037/dev0001171>
- Chinn, L. K., Noonan, C. F., Patton, K. S., & Lockman, J. J. (2024). Tactile localization promotes infant self-recognition in the mirror-mark test. *Current Biology*, *34*(6), 1370–1375.e2. <https://doi.org/10.1016/j.cub.2024.02.028>
- Choi, Y., & Luo, Y. (2023). Understanding preferences in infancy. *WIREs Cognitive Science*, *14*(4), e1643. <https://doi.org/10.1002/wcs.1643>

- Cirelli, L. K. (2018a). How interpersonal synchrony facilitates early prosocial behavior. *Current Opinion in Psychology*, 20, 35–39. <https://doi.org/10.1016/j.copsyc.2017.08.009>
- Cirelli, L. K. (2018b). How interpersonal synchrony facilitates early prosocial behavior. *Current Opinion in Psychology*, 20, 35–39. <https://doi.org/10.1016/j.copsyc.2017.08.009>
- Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014a). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, 17(6), 1003–1011. <https://doi.org/10.1111/desc.12193>
- Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014b). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, 17(6), 1003–1011. <https://doi.org/10.1111/desc.12193>
- Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014c). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, 17(6), 1003–1011. <https://doi.org/10.1111/desc.12193>
- Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2016). Social Effects of Movement Synchrony: Increased Infant Helpfulness only Transfers to Affiliates of Synchronously Moving Partners. *Infancy*, 21(6), 807–821. <https://doi.org/10.1111/infa.12140>
- Cohen, E. E. A., Ejsmond-Frey, R., Knight, N., & Dunbar, R. I. M. (2009). Rowers' high: Behavioural synchrony is correlated with elevated pain thresholds. *Biology Letters*, 6(1), 106–108. <https://doi.org/10.1098/rsbl.2009.0670>
- Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., Radkova, I., Deschrijver, E., & Brass, M. (2018). Automatic imitation: A meta-analysis. *Psychological Bulletin*, 144(5), 453–500. <https://doi.org/10.1037/bul0000143>

- Cracco, E., & Brass, M. (2019). Reaction time indices of automatic imitation measure imitative response tendencies. *Consciousness and Cognition*, *68*, 115–118. <https://doi.org/10.1016/j.concog.2019.01.001>
- Cracco, E., Genschow, O., & Brass, M. (2024). Covert and overt automatic imitation are correlated. *Psychonomic Bulletin & Review*, *31*(3), 1348–1352. <https://doi.org/10.3758/s13423-023-02420-0>
- Cross, L., Atherton, G., & Sebanz, N. (2021a). Intentional synchronisation affects automatic imitation and source memory. *Scientific Reports*, *11*(1), 573. <https://doi.org/10.1038/s41598-020-79796-9>
- Cross, L., Atherton, G., & Sebanz, N. (2021b). Intentional synchronisation affects automatic imitation and source memory. *Scientific Reports*, *11*(1), 573. <https://doi.org/10.1038/s41598-020-79796-9>
- Cross, L., Michael, J., Wilsdon, L., Henson, A., & Atherton, G. (2020). Still want to help? Interpersonal coordination's effects on helping behaviour after a 24 hour delay. *Acta Psychologica*, *206*, 103062. <https://doi.org/10.1016/j.actpsy.2020.103062>
- Cross, L., Turgeon, M., & Atherton, G. (2019a). How Moving Together Binds Us Together: The Social Consequences of Interpersonal Entrainment and Group Processes. *Open Psychology*, *1*(1), 273–302. <https://doi.org/10.1515/psych-2018-0018>
- Cross, L., Turgeon, M., & Atherton, G. (2019b). How Moving Together Binds Us Together: The Social Consequences of Interpersonal Entrainment and Group Processes. *Open Psychology*, *1*(1), 273–302. <https://doi.org/10.1515/psych-2018-0018>
- Cross, L., Turgeon, M., & Atherton, G. (2021). Moving with the in-crowd: Cooperation and interpersonal entrainment in in- vs. out- groups. *Current Psychology*, *40*(7), 3393–3400. <https://doi.org/10.1007/s12144-019-00283-0>

- Cross, L., Wilson, Andrew D., & Golonka, S. (2020). I'll just watch: Do the pro-social effects of coordination really generalize to non-actors? *The Journal of Social Psychology, 160*(2), 248–262. <https://doi.org/10.1080/00224545.2019.1623161>
- Curioni, A., Vesper, C., Knoblich, G., & Sebanz, N. (2019). Reciprocal information flow and role distribution support joint action coordination. *Cognition, 187*, 21–31. <https://doi.org/10.1016/j.cognition.2019.02.006>
- De Souter, L., Braem, S., Genschow, O., Brass, M., & Cracco, E. (2021). Social group membership does not modulate automatic imitation in a contrastive multi-agent paradigm. *Quarterly Journal of Experimental Psychology, 74*(4), 746–759. <https://doi.org/10.1177/1747021820986528>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Durkheim, É. (2008). *The Elementary Forms of the Religious Life* (J. W. Swain, Trans.). Courier Corporation.
- Ehrenreich, B. (2006). *Dancing in the streets: A history of collective joy* (p. 320). Metropolitan Books/Henry Holt and Company.
- Fairhurst, M. T., Janata, P., & Keller, P. E. (2014). Leading the follower: An fMRI investigation of dynamic cooperativity and leader–follower strategies in synchronization with an adaptive virtual partner. *NeuroImage, 84*, 688–697. <https://doi.org/10.1016/j.neuroimage.2013.09.027>

- Fawcett, C., & Tunçgenç, B. (2017). Infants' use of movement synchrony to infer social affiliation in others. *Journal of Experimental Child Psychology*, *160*, 127–136. <https://doi.org/10.1016/j.jecp.2017.03.014>
- Feniger-Schaal, R., Noy, Lior, Hart, Yuval, Koren-Karie, Nina, Mayo, Avraham E., & Alon, U. (2016). Would you like to play together? Adults' attachment and the mirror game. *Attachment & Human Development*, *18*(1), 33–45. <https://doi.org/10.1080/14616734.2015.1109677>
- Fischer, R., Callander, R., Reddish, P., & Bulbulia, J. (2013). How Do Rituals Affect Cooperation? *Human Nature*, *24*(2), 115–125. <https://doi.org/10.1007/s12110-013-9167-y>
- Fischer, R., Xygalatas, D., Mitkidis, P., Reddish, P., Tok, P., Konvalinka, I., & Bulbulia, J. (2014). The Fire-Walker's High: Affect and Physiological Responses in an Extreme Collective Ritual. *PLOS ONE*, *9*(2), e88355. <https://doi.org/10.1371/journal.pone.0088355>
- Freeman, W. J. (1998). *A neurobiological role of music in social bonding*. <https://escholarship.org/uc/item/9025x8rt>
- Frick, A., Clément, F., & Gruber, T. (2017). Evidence for a sex effect during overimitation: Boys copy irrelevant modelled actions more than girls across cultures. *Royal Society Open Science*, *4*(12), 170367. <https://doi.org/10.1098/rsos.170367>
- Gallup, G. G., & Anderson, J. R. (2020). Self-recognition in animals: Where do we stand 50 years later? Lessons from cleaner wrasse and other species. *Psychology of Consciousness: Theory, Research, and Practice*, *7*(1), 46–58. <https://doi.org/10.1037/cns0000206>
- Genschow, O., Cracco, E., Verbeke, P., Westfal, M., & Crusius, J. (2021). A direct test of the similarity assumption—Focusing on differences as compared with similarities

decreases automatic imitation. *Cognition*, 215, 104824.
<https://doi.org/10.1016/j.cognition.2021.104824>

Gleibs, I. H., Wilson, N., Reddy, G., & Catmur, C. (2016). Group Dynamics in Automatic Imitation. *PLOS ONE*, 11(9), e0162880. <https://doi.org/10.1371/journal.pone.0162880>

Goncharova, Y., & Ross, J. (2024). Who Am I? A Longitudinal Investigation of the Multidimensional Self. *Infant and Child Development*, 33(6), e2556. <https://doi.org/10.1002/icd.2556>

Good, A., Choma ,Becky, & and Russo, F. A. (2017). Movement Synchrony Influences Intergroup Relations in a Minimal Groups Paradigm. *Basic and Applied Social Psychology*, 39(4), 231–238. <https://doi.org/10.1080/01973533.2017.1337015>

Good, A., & Russo, F. A. (2016). Singing Promotes Cooperation in a Diverse Group of Children. *Social Psychology*, 47(6), 340–344. <https://doi.org/10.1027/1864-9335/a000282>

Goupil, L., Wolf, T., Saint-Germier, P., Aucouturier, J.-J., & Canonne, C. (2021). Emergent Shared Intentions Support Coordination During Collective Musical Improvisations. *Cognitive Science*, 45(1), e12932. <https://doi.org/10.1111/cogs.12932>

Grosse Wiesmann, C., Rothmaler, K., Habdank, K., Hasan, E., Yang, C., Yeung, E., & Southgate, V. (2024). *From other-reference to a self-reference effect – Reversal of the classic memory bias in human infant*. <https://doi.org/10.31234/osf.io/my4jg>

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, 14(1), 21–51. <https://doi.org/10.1007/s12110-003-1015-z>

Haidt, J., Patrick Seder, J., & Kesebir, S. (2008). Hive Psychology, Happiness, and Public Policy. *The Journal of Legal Studies*, 37(S2), S133–S156. <https://doi.org/10.1086/529447>

- Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not Like Me = Bad: Infants Prefer Those Who Harm Dissimilar Others. *Psychological Science*, 24(4), 589–594. <https://doi.org/10.1177/0956797612457785>
- Hove, M. J. (2008). Shared circuits, shared time, and interpersonal synchrony. *Behavioral and Brain Sciences*, 31(1), 29–30. <https://doi.org/10.1017/S0140525X07003202>
- Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, 27(6), 949–960. <https://doi.org/10.1521/soco.2009.27.6.949>
- Hurley, S. (2008). The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading. *Behavioral and Brain Sciences*, 31(1), 1–22. <https://doi.org/10.1017/S0140525X07003123>
- Huys, R., Studenka, B. E., Rheaume, N. L., Zelaznik, H. N., & Jirsa, V. K. (2008a). Distinct Timing Mechanisms Produce Discrete and Continuous Movements. *PLOS Computational Biology*, 4(4), e1000061. <https://doi.org/10.1371/journal.pcbi.1000061>
- Huys, R., Studenka, B. E., Rheaume, N. L., Zelaznik, H. N., & Jirsa, V. K. (2008b). Distinct Timing Mechanisms Produce Discrete and Continuous Movements. *PLOS Computational Biology*, 4(4), e1000061. <https://doi.org/10.1371/journal.pcbi.1000061>
- Ivry, R. B., Spencer, R. M., Zelaznik, H. N., & Diedrichsen, J. (2002). The Cerebellum and Event Timing. *Annals of the New York Academy of Sciences*, 978(1), 302–317. <https://doi.org/10.1111/j.1749-6632.2002.tb07576.x>
- Jacquey, L., Fagard, J., Esseily, R., & O'Regan, J. K. (2020). Detection of sensorimotor contingencies in infants before the age of 1 year: A comprehensive review. *Developmental Psychology*, 56(7), 1233–1251. <https://doi.org/10.1037/dev0000916>

- Johnstone, L. T., & Carey, D. P. (2016). Do left hand reaction time advantages depend on localising unpredictable targets? *Experimental Brain Research*, *234*(12), 3625–3632. <https://doi.org/10.1007/s00221-016-4758-6>
- Kampis, D., Grosse Wiesmann, C., Koop, S., & Southgate, V. (2022). Understanding the self in relation to others: Infants spontaneously map another’s face to their own at 16–26 months. *Developmental Science*, *25*(3), e13197. <https://doi.org/10.1111/desc.13197>
- Kampis, D., Lukowski Duplessy, H., Askitis, D., & Southgate, V. (2023). Training self-other distinction facilitates perspective taking in young children. *Child Development*, *94*(4), 956–969. <https://doi.org/10.1111/cdev.13912>
- Kirschner, S., & Tomasello, M. (2010a). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. <https://doi.org/10.1016/j.evolhumbehav.2010.04.004>
- Kirschner, S., & Tomasello, M. (2010b). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364. <https://doi.org/10.1016/j.evolhumbehav.2010.04.004>
- Klein-Radukic, S., & Zmyj, N. (2020). No evidence for infants’ contingency detection as a precursor of toddlers’ self-recognition. *Infant Behavior and Development*, *60*, 101459. <https://doi.org/10.1016/j.infbeh.2020.101459>
- Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized Drumming Enhances Activity in the Caudate and Facilitates Prosocial Commitment—If the Rhythm Comes Easily. *PLOS ONE*, *6*(11), e27272. <https://doi.org/10.1371/journal.pone.0027272>
- Lang, M., Bahna, V., Shaver, J. H., Reddish, P., & Xygalatas, D. (2017). Sync to link: Endorphin-mediated synchrony effects on cooperation. *Biological Psychology*, *127*, 191–197. <https://doi.org/10.1016/j.biopsycho.2017.06.001>

- Launay, J., Dean, R. T., & Bailes, F. (2013). Synchronization Can Influence Trust Following Virtual Interaction. *Experimental Psychology*, *60*(1), 53–63. <https://doi.org/10.1027/1618-3169/a000173>
- Launay, J., Tarr, B., & Dunbar, R. I. M. (2016). Synchrony as an Adaptive Mechanism for Large-Scale Human Social Bonding. *Ethology*, *122*(10), 779–789. <https://doi.org/10.1111/eth.12528>
- Lewis, M., & Ramsay, D. (2004a). Development of Self-Recognition, Personal Pronoun Use, and Pretend Play During the 2nd Year. *Child Development*, *75*(6), 1821–1831. <https://doi.org/10.1111/j.1467-8624.2004.00819.x>
- Lewis, M., & Ramsay, D. (2004b). Development of Self-Recognition, Personal Pronoun Use, and Pretend Play During the 2nd Year. *Child Development*, *75*(6), 1821–1831. <https://doi.org/10.1111/j.1467-8624.2004.00819.x>
- Lewis, M., Sullivan, M. W., Stanger, C., & Weiss, M. (1989). Self Development and Self-Conscious Emotions. *Child Development*, *60*(1), 146–156. <https://doi.org/10.2307/1131080>
- Lorås, H., Sigmundsson, H., Talcott, J. B., Öhberg, F., & Stensdotter, A. K. (2012). Timing continuous or discontinuous movements across effectors specified by different pacing modalities and intervals. *Experimental Brain Research*, *220*(3), 335–347. <https://doi.org/10.1007/s00221-012-3142-4>
- Lucca, K., & Wilbourn, M. P. (2019). The what and the how: Information-seeking pointing gestures facilitate learning labels and functions. *Journal of Experimental Child Psychology*, *178*, 417–436. <https://doi.org/10.1016/j.jecp.2018.08.003>
- Macrae, C. N., Duffy, O. K., Miles, L. K., & Lawrence, J. (2008). A case of hand waving: Action synchrony and person perception. *Cognition*, *109*(1), 152–156. <https://doi.org/10.1016/j.cognition.2008.07.007>

- McNaughton, K. A., & Redcay, E. (2020). Interpersonal Synchrony in Autism. *Current Psychiatry Reports*, 22(3), 12. <https://doi.org/10.1007/s11920-020-1135-8>
- McNeill, W. H. (1997). *Keeping Together in Time: Dance and Drill in Human History*. Harvard University Press.
- Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, 24(4), 470–476. <https://doi.org/10.1037/0012-1649.24.4.470>
- Miles, L. K., Nind, L. K., Henderson, Z., & Macrae, C. N. (2010). Moving memories: Behavioral synchrony and memory for self and others. *Journal of Experimental Social Psychology*, 46(2), 457–460. <https://doi.org/10.1016/j.jesp.2009.12.006>
- Mogan, R., Fischer, R., & Bulbulia, J. A. (2017a). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, 72, 13–20. <https://doi.org/10.1016/j.jesp.2017.03.009>
- Mogan, R., Fischer, R., & Bulbulia, J. A. (2017b). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, 72, 13–20. <https://doi.org/10.1016/j.jesp.2017.03.009>
- Morton, M. (2016). We Can Work it Out: The Importance of Rupture and Repair Processes in Infancy and Adult Life for Flourishing. *Health Care Analysis*, 24(2), 119–132. <https://doi.org/10.1007/s10728-016-0319-1>
- Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, 108(52), 20947–20952. <https://doi.org/10.1073/pnas.1108155108>
- Pardo-Olmos, M., Martí-Vilar, M., Hidalgo-Fuentes, S., & Cabedo-Peris, J. (2025). Interventions Through Music and Interpersonal Synchrony That Enhance Prosocial

Behavior: A Systematic Review. *European Journal of Investigation in Health, Psychology and Education*, 15(3), Article 3. <https://doi.org/10.3390/ejihpe15030029>

Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*, 211(3), 505–515. <https://doi.org/10.1007/s00221-011-2616-0>

Petitmengin, C. (2006). Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5(3–4), 229–269. <https://doi.org/10.1007/s11097-006-9022-2>

Petitmengin, C., Remillieux, A., & Valenzuela-Moguillansky, C. (2019). Discovering the structures of lived experience: Towards a micro-phenomenological analysis method. *Phenomenology and the Cognitive Sciences*, 18(4), 691–730. <https://doi.org/10.1007/s11097-018-9597-4>

Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior and Development*, 29(3), 386–401. <https://doi.org/10.1016/j.infbeh.2006.01.004>

Qian, M., Yu, C., Fu, G., & Cirelli, L. K. (2020). Shaping children's racial bias through interpersonal movement. *Journal of Experimental Child Psychology*, 198, 104884. <https://doi.org/10.1016/j.jecp.2020.104884>

Rabinowitch, T.-C. (2023). Musical engagement as a duet of tight synchrony and loose interpretability. *Physics of Life Reviews*, 44, 122–136. <https://doi.org/10.1016/j.plrev.2022.12.019>

Rabinowitch, T.-C., & Knafo-Noam, A. (2015). Synchronous Rhythmic Interaction Enhances Children's Perceived Similarity and Closeness towards Each Other. *PLOS ONE*, 10(4), e0120878. <https://doi.org/10.1371/journal.pone.0120878>

- Rabinowitch, T.-C., & Meltzoff, A. N. (2017a). Joint Rhythmic Movement Increases 4-Year-Old Children's Prosocial Sharing and Fairness Toward Peers. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01050>
- Rabinowitch, T.-C., & Meltzoff, A. N. (2017b). Synchronized movement experience enhances peer cooperation in preschool children. *Journal of Experimental Child Psychology*, 160, 21–32. <https://doi.org/10.1016/j.jecp.2017.03.001>
- Reddish, P., Bulbulia, Joseph, & Fischer, R. (2014). Does synchrony promote generalized prosociality? *Religion, Brain & Behavior*, 4(1), 3–19. <https://doi.org/10.1080/2153599X.2013.764545>
- Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's Dance Together: Synchrony, Shared Intentionality and Cooperation. *PLOS ONE*, 8(8), e71182. <https://doi.org/10.1371/journal.pone.0071182>
- Reddish, P., Tong, E. M. W., Jong, J., Lanman, J. A., & Whitehouse, H. (2016). Collective synchrony increases prosociality towards non-performers and outgroup members. *British Journal of Social Psychology*, 55(4), 722–738. <https://doi.org/10.1111/bjso.12165>
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 20(3), 403–452. <https://doi.org/10.3758/s13423-012-0371-2>
- Reproducibility Project: Psychology*. (2012). <https://doi.org/10.17605/OSF.IO/EZCUJ>
- Richardson, M. J., Harrison, S. J., Kallen, R. W., Walton, A., Eiler, B. A., Saltzman, E., & Schmidt, R. C. (2015). Self-organized complementary joint action: Behavioral dynamics of an interpersonal collision-avoidance task. *Journal of Experimental Psychology: Human Perception and Performance*, 41(3), 665–679. <https://doi.org/10.1037/xhp0000041>

- Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 26(6), 867–891. <https://doi.org/10.1016/j.humov.2007.07.002>
- Rochat, P. (2018). The Ontogeny of Human Self-Consciousness. *Current Directions in Psychological Science*, 27(5), 345–350. <https://doi.org/10.1177/0963721418760236>
- Sadaphal, D. P., Blum, C. R., Keller, P. E., & Fitch, W. T. (2025). Beyond perfect synchrony: Shared interpersonal rhythmic timing enhances self-other merging judgements. *Royal Society Open Science*, 12(3), 241501. <https://doi.org/10.1098/rsos.241501>
- Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, 122(2), 228–235. <https://doi.org/10.1016/j.cognition.2011.11.004>
- Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 44, e59. <https://doi.org/10.1017/S0140525X20000333>
- Schönbrodt, F. D., & Wagenmakers, E.-J. (2018a). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 25(1), 128–142. <https://doi.org/10.3758/s13423-017-1230-y>
- Schönbrodt, F. D., & Wagenmakers, E.-J. (2018b). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, 25(1), 128–142. <https://doi.org/10.3758/s13423-017-1230-y>
- Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017a). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22(2), 322–339. <https://doi.org/10.1037/met0000061>

- Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017b). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22(2), 322–339. <https://doi.org/10.1037/met0000061>
- Shrout, P. E., & Rodgers, J. L. (2018). Psychology, Science, and Knowledge Construction: Broadening Perspectives from the Replication Crisis. *Annual Review of Psychology*, 69(1), 487–510. <https://doi.org/10.1146/annurev-psych-122216-011845>
- Southgate, V. (2024). *The Origins and Emergence of Self-Representation*. <https://doi.org/10.1146/annurev-devpsych-120621-025747>
- Sowden, S., & Catmur, C. (2015). The Role of the Right Temporoparietal Junction in the Control of Imitation. *Cerebral Cortex*, 25(4), 1107–1113. <https://doi.org/10.1093/cercor/bht306>
- Spencer, R. M. C., Zelaznik, H. N., Diedrichsen, J., & Ivry, R. B. (2003). Disrupted Timing of Discontinuous But Not Continuous Movements by Cerebellar Lesions. *Science*, 300(5624), 1437–1439. <https://doi.org/10.1126/science.1083661>
- Spengler, S., Brass, M., Kühn, S., & Schütz-Bosbach, S. (2010). Minimizing motor mimicry by myself: Self-focus enhances online action-control mechanisms during motor contagion. *Consciousness and Cognition*, 19(1), 98–106. <https://doi.org/10.1016/j.concog.2009.12.014>
- Stupacher, J., Maes, P.-J., Witte, M., & Wood, G. (2017). Music strengthens prosocial effects of interpersonal synchronization – If you move in time with the beat. *Journal of Experimental Social Psychology*, 72, 39–44. <https://doi.org/10.1016/j.jesp.2017.04.007>
- Stupacher, J., Witte, M., & Wood, G. (2016). *Social effects of interpersonal synchronization during listening to music compared to a metronome: What can we learn from implicit measures?*

- Stupacher, J., Wood, G., & Witte, M. (2017). Synchrony and sympathy: Social entrainment with music compared to a metronome. *Psychomusicology: Music, Mind, and Brain*, 27(3), 158–166. <https://doi.org/10.1037/pmu0000181>
- Sullivan, P., & Rickers, K. (2013). The effect of behavioral synchrony in groups of teammates and strangers. *International Journal of Sport and Exercise Psychology*, 11(3), 286–291. <https://doi.org/10.1080/1612197X.2013.750139>
- Sullivan, P., Gagnon, M., Gammage, K., & Peters, S. (2015). Is the Effect of Behavioral Synchrony on Cooperative Behavior Mediated by Pain Threshold? *The Journal of Social Psychology*, 155(6), 650–660. <https://doi.org/10.1080/00224545.2015.1071766>
- Sullivan, P. J., Rickers, K., & Gammage, K. L. (2014). The effect of different phases of synchrony on pain threshold. *Group Dynamics: Theory, Research, and Practice*, 18(2), 122–128. <https://doi.org/10.1037/gdn0000001>
- Sun, Y., Greaves, D. A., Orgs, G., de C. Hamilton, A. F., Day, S., & Ward, J. A. (2023). Using Wearable Sensors to Measure Interpersonal Synchrony in Actors and Audience Members During a Live Theatre Performance. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 7(1), 27:1-27:29. <https://doi.org/10.1145/3580781>
- Tajfel, H. (2010). *Social Identity and Intergroup Relations*. Cambridge University Press.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178. <https://doi.org/10.1002/ejsp.2420010202>
- Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, 11(10), 20150767. <https://doi.org/10.1098/rsbl.2015.0767>

- Tarr, B., Launay, J., & Dunbar, R. I. M. (2014). Music and social bonding: “Self-other” merging and neurohormonal mechanisms. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01096>
- Tarr, B., Launay, J., & Dunbar, R. I. M. (2016). Silent disco: Dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior*, 37(5), 343–349. <https://doi.org/10.1016/j.evolhumbehav.2016.02.004>
- Tarr, B., Slater, M., & Cohen, E. (2018). Synchrony and social connection in immersive Virtual Reality. *Scientific Reports*, 8(1), 3693. <https://doi.org/10.1038/s41598-018-21765-4>
- The Early Childhood Behavior Questionnaire (ECBQ) | Mary Rothbart's Temperament Questionnaires.* (n.d.). Retrieved June 1, 2025, from <https://research.bowdoin.edu/rothbart-temperament-questionnaires/instrument-descriptions/the-early-childhood-behavior-questionnaire/>
- Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, 5. <https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01323>
- Tunçgenç, B., Bamford, J. S., Fawcett, C., & Cohen, E. (2023). The Synchrony-Prosociality Link Cannot Be Explained Away as Expectancy Effect: Response to Atwood et al. (2022). *Open Mind*, 7, 711–714. https://doi.org/10.1162/opmi_a_00103
- Tunçgenç, B., & Cohen, E. (2016a). Movement Synchrony Forges Social Bonds across Group Divides. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00782>
- Tunçgenç, B., & Cohen, E. (2016b). Movement Synchrony Forges Social Bonds across Group Divides. *Frontiers in Psychology*, 7. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.00782>

- Tunçgenç, B., Cohen, E., & Fawcett, C. (2015). Rock With Me: The Role of Movement Synchrony in Infants' Social and Nonsocial Choices. *Child Development, 86*(3), 976–984. <https://doi.org/10.1111/cdev.12354>
- Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion, 11*, 262–266. <https://doi.org/10.1037/a0021302>
- Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46*(4), 693–695. <https://doi.org/10.1016/j.jesp.2010.03.004>
- Valenzuela-Moguillansky, C., & Vásquez-Rosati, A. (2019a). An Analysis Procedure for the Micro-Phenomenological Interview. *Constructivist Foundations, 14*(2), Article 2.
- Valenzuela-Moguillansky, C., & Vásquez-Rosati, A. (2019b). An Analysis Procedure for the Micro-Phenomenological Interview. *Constructivist Foundations, 14*(2), Article 2.
- Varlet, M., Nozaradan, S., Nijhuis, P., & Keller, P. E. (2020). Neural tracking and integration of 'self' and 'other' in improvised interpersonal coordination. *NeuroImage, 206*, 116303. <https://doi.org/10.1016/j.neuroimage.2019.116303>
- Varlet, M., & Richardson, M. J. (2011). Computation of continuous relative phase and modulation of frequency of human movement. *Journal of Biomechanics, 44*(6), 1200–1204. <https://doi.org/10.1016/j.jbiomech.2011.02.001>
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics, 72*(4), 871–884. <https://doi.org/10.3758/APP.72.4.871>
- Washburn, A., Wright, M. J., Chafe, C., & Fujioka, T. (2021). Temporal Coordination in Piano Duet Networked Music Performance (NMP): Interactions Between Acoustic Transmission Latency and Musical Role Asymmetries. *Frontiers in Psychology, 12*. <https://doi.org/10.3389/fpsyg.2021.707090>

- Weinstein, D., Launay, J., Pearce, E., Dunbar, R. I. M., & Stewart, L. (2016). Singing and social bonding: Changes in connectivity and pain threshold as a function of group size. *Evolution and Human Behavior*, 37(2), 152–158. <https://doi.org/10.1016/j.evolhumbehav.2015.10.002>
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and Cooperation. *Psychological Science*, 20(1), 1–5. <https://doi.org/10.1111/j.1467-9280.2008.02253.x>
- Wolf, T., Sebanz, N., & Knoblich, G. (2018). Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing? *Cognition*, 178, 103–108. <https://doi.org/10.1016/j.cognition.2018.05.012>
- Yeung, E., Askitis, D., Manea, V., & Southgate, V. (2022). Emerging Self-Representation Presents a Challenge When Perspectives Conflict. *Open Mind*, 6, 232–249. https://doi.org/10.1162/opmi_a_00065
- Zhai, C., Alderisio, F., Słowiński, P., Tsaneva-Atanasova, K., & Bernardo, M. di. (2016). Design of a Virtual Player for Joint Improvisation with Humans in the Mirror Game. *PLOS ONE*, 11(4), e0154361. <https://doi.org/10.1371/journal.pone.0154361>