Acting in the We-Mode:
The Content and Structure of
Joint Task Representations

by

Maximilian Marschner

Submitted to

Central European University

Department of Cognitive Science

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Cognitive Science

Primary supervisor: Günther Knoblich

Secondary supervisors: David Dignath & Marcel Brass

Vienna, Austria

2025

Declaration of Authorship

I, the undersigned, Maximilian Marschner, candidate for the PhD degree in Cognitive Science declare herewith that the present dissertation titled *"Acting in the We-Mode: The Content and Structure of Joint Task Representations"* 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 dissertation infringes on any person's or institution's copyright. I also declare that no part of the dissertation has been submitted in this form to any other institution of higher education for an academic degree.

This dissertation includes work that appears in the following publications:

Marschner, M., Galang, M., & Brass, M. (2025). *Us and them: Anticipated imitation between groups*. PsyArXiv. https://doi.org/10.31234/osf.io/n4drs_v1

Marschner, M., Dignath, D., & Knoblich, G. (2024). Me or we? Action-outcome learning in synchronous joint action. *Cognition*, *247*. https://doi.org/10.1016/j.cognition.2024.105785

Marschner, M., Knoblich, G., & Dignath, D. (2025). *Do Individual and Joint Action Goals Modulate Imitative Response Tendencies?* OSF Preprints. https://doi.org/10.31219/osf.io/3ctsx_v2

Other parts of this dissertation contain research conducted in collaboration and will be submitted for publication with the following co-authors:

- Chapter 4.2 with Günther Knoblich and David Dignath

Vienna, 10. December 2025

MAXIMILIAN MARSCHNER

Copyright Notice

Abstract

The ability to coordinate our actions with others to reach shared goals is a hallmark of human sociality. What are the cognitive foundations for this ability? In the present dissertation, I addressed this question by investigating the content and structure of joint task representations, i.e., of the mental representations that guide action planning and control in joint action by specifying what co-actors need to do to reach their goals.

Distinguishing different theoretical proposals of how joint task representations could be structured, the present dissertation tested central assumptions of one of these proposals, referred to as the *relational structure* of joint task representations. This proposal posits that joint task representations encode group-level relations between co-actors' individual task contributions, thereby capturing how joint actions are performed at the level of the group rather than at the level of the separate interacting individuals.

The present dissertation comprises four empirical research projects testing this proposal by investigating anticipated imitation effects in social interactions between groups (Study I), action-outcome learning in synchronous joint action (Study II), and modulations of imitative response tendencies in joint task settings (Study III and IV). Study I showed that anticipated imitation effects can be shaped by group-level congruency relations but also pointed to limitations in how readily people integrate others' actions into their own task representations. Study II showed that action-outcome learning is sensitive to group-level relations between co-actors' individual task contributions. Yet, it also indicated that the application of joint task representations embedding a relational structure is cognitively costly and not necessarily the default in task contexts featuring other co-actors. Finally, Study III and Study IV revealed that shared goals that specify relations between the outcomes of own and others' individual task contributions are limited in modulating imitative response tendencies in interactive task settings. This indicates that the representations guiding action selection and control in

interactive task settings may often remain anchored at individual rather than group-level performance.

Taken together, the findings of the present dissertation provide novel support for the proposal that joint task representations can encode group-level relations between co-actors' individual task contributions. At the same time, they reveal important limits on how readily these representations get employed in interactive task settings. These findings indicate that joint task representations do not default to a fixed structure, but become flexibly adapted to the coordination demands, temporal dynamics and specific characteristics of the task at hand. These findings advance our understanding of the cognitive architecture underlying joint action, indicating that the content and structure of joint task representations is determined by deliberate construal processes balancing costs and benefits of different representational structures.

Acknowledgements

This dissertation comprises work conducted over the course of five years. In these years, I witnessed the move of a university and a global pandemic, lived in three different cities, and was more than once doubting if I will make it through the end. Managing this time and completing this dissertation would have been impossible without the help and support of others who accompanied me on this journey.

First and foremost, I would like to express my gratitude to Günther Knoblich. Thank you for supervising my work, for trusting and shaping my ideas, for motivating me in times of doubt, for being a constant source of guidance and inspiration, and for reminding me that everything gets easier with the right amount of wit and humour.

Second, I would like to thank David Dignath. Thank you not only for supervising most of the research presented in this dissertation and for helping me with your scientific expertise and methodological rigour. Thank you foremost for having become a mentor and role model over the last years who helped me to grow up in academia and who sparked my interest in joint action research in the first place.

Third, I am very thankful to Marcel Brass and Mike Galang. Marcel, thank you for welcoming me to your research group and for offering your valuable guidance and support for my research. I am very grateful that you made me able to experience a new research environment, to learn new methods, and to broaden the scope of my research interests. Mike, thank you for pushing our collaboration, for sharing your ideas and tuning mine, and for your expertise on the nuts and bolts of rigorous experimental research. I am also very thankful to all other members of the Social Intelligence Lab at HU Berlin for welcoming me as a new colleague and for making me feel right at home.

Fourth, a big Thank You to all members of the Department of Cognitive Science at CEU for creating the most inspiring research environment one could wish for. Special thanks to

Martin Dockendorff, Luke McEllin, Maja Blesic, Maria Mavridaki, Thomas Wolf, Mateusz Wozniak, Arianna Curioni, Elisa Wiedemann, Nejra Rizvanovic, Alper Demircan, Kassandra Friebe, Thomas Ganzetti, Ieva Lukosiunaite, Shaheed Azaad, Natalie Sebanz and all other current and former members of the Social Mind and Body Lab. A special shoutout also to Györgyné Finta, Andrea Jenei, Candasch Acar, and Fanni Takatsy for making the lab and department run so smoothly.

Finally, my thanks go to my friends and family. Without your support I would not be sitting here writing these words. Thank you for always believing in me and for being there in the good and in the bad times.

Table of Content

1. Introduction

Changing the world through our actions — to exert *agency* — is constitutive for our life. Our life, however, would be void if we were only able to exert agency on our own. There would be no handshakes, no bands, no parties, no conversations, no infrastructure, no governments, no research institutions, and the human species would be either extinct by now or still be living in caves. Fortunately though, humans possess the capacity to exert *joint agency* as well, i.e. we are able to coordinate our actions with others to achieve shared goals. This capacity for *joint action* enables us not only to great our friends with a handshake, make music with others or dance together. It is thought to provide the very foundation for the evolution of human culture, language, and cognition (Knoblich & Sebanz, 2006; Tomasello et al., 2005b; Vygotsky, 1978).

While the relevance of joint action for human life is without doubt, explaining what constitutes our ability to act jointly is a challenging scientific endeavour that has occupied research across various disciplines in cognitive science. From a psychological perspective, engaging in joint action requires dedicated cognitive, perceptual, and motor capacities that enable individuals to plan and control their actions in relation to the actions of their interaction partners and in relation to the joint action outcomes they produce together (Knoblich et al., 2011). Prior research has established that these capacities involve a set of specialized coordination mechanisms that operate while joint actions are being performed (see Vesper & Sebanz, 2016 for a comprehensive overview). However, before engaging many of these mechanisms, co-actors must have access to dedicated mental representations that encode task-relevant features of the joint action in the first place, such as co-actors' *goals* and the *actions* and *outcomes* to be produced in order to reach them (Vesper et al., 2010). Despite an impressive range of empirical and theoretical research contributions over the last decades (see Sebanz & Knoblich, 2021; Vesper & Sebanz, 2016 for reviews), a comprehensive account of the content

and structure of these mental representations — that I shall refer to as *joint task representations*[1] — is still lacking. The work I present in this dissertation aims to offer such an account by investigating what joint task representation encode, how they are structured, and how their structure supports successful coordination of joint actions.

In the following, I will first situate the approach of the current dissertation in the broader field of joint action research to motivate why an understanding of joint action requires an investigation of the content and structure of joint task representations. Building on previous research, I will then present three hypotheses about possible structures of joint task representations and lay out how these structures support coordination of joint actions. Based on this synthesis I will then lay out the research aims that guided my own empirical work.

1.1. Approaches to joint action

Broadly defined, joint action refers to "any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the environment" (Sebanz et al., 2006, p. 70). This broad definition views a wide range of social situations as joint actions. They range from spontaneous alignment of clapping rhythms in a group of spectators over mundane activities such as carrying furniture together to highly specialized skills such as an army corps executing a tactical manoeuvre. As these examples highlight, joint actions can unfold across multiple time scales, involve varying numbers of co-actors and require interpersonal coordination at different levels of action specification. These range from the coordination of higher-level plans and decisions about whether to act together, and what goals to pursue, to precise mutual adjustments of bodily movements in space and time

---

[1] Throughout the dissertation I sometimes use the term *joint action plan* or simply *joint action representation* instead of *joint task representation*. I use these terms to refer to the same concept and use them interchangeably.

(Pacherie, 2012). It is therefore not surprising that answers to the question of what constitutes our ability to act jointly are as multifaceted as the subject itself.

1.1.1. Philosophical accounts of joint action

Philosophical accounts approach joint action by analysing the kind and structure of mental states that demarcate *intentional* joint action from acting alone or merely in parallel with others (Bratman, 1992; Gilbert, 2009; Pacherie, 2013; Searle, 1990; Tuomela, 2006). They all view the presence of *shared intentions* as a necessary precondition to engage in joint action. However, they hold different positions on how to best capture their collective nature that would distinguish them from individual intentions. Some accounts locate the collective aspect of shared intentions at the level of the subject holding the intention (Gilbert, 2009), others at the level of their content (Bratman, 1992), while some accounts postulate that having a shared intention reflects a distinct mode or attitude of having an intention qua member of a group, a so-called we-intention or an intention in the "we-mode" (Gallotti & Frith, 2013; Searle, 1990; Tuomela, 2006). The notion of shared intentions as the denoting aspect of joint action has sparked insightful research on the role of moral obligations and commitments in joint action (Michael & Pacherie, 2015), and on the role of language for establishing common-ground between co-actors (Clark, 1996). It has also guided empirical research on the phylogenetic and ontogenetic development of joint action (Call, 2009; Carpenter, 2009; Tomasello et al., 2005a).

Despite the fruitfulness of these research programs, philosophical accounts of joint action centred on shared intentions, suffer from a central limitation: Their analysis of joint action remains limited on stages that precede actual engagement in joint action by focussing on the coordination of relatively high-level, future-directed action plans. Consequently, philosophical accounts of joint action fall short in explaining how co-actors eventually come to execute a joint action once they formed a shared intention to do so (Pacherie, 2013; Tollefsen & Dale, 2012).

1.1.2. Dynamical accounts of joint action

A very different approach to joint action that is primarily concerned with the online execution of joint actions is offered by researchers working in the framework of coordination dynamics (Marsh et al., 2009; Oullier et al., 2008; Schmidt et al., 2011; Schmidt & Richardson, 2008). Rather than postulating distinct mental states and internal cognitive processes that make people proficient to engage in joint action, these accounts focus on the description and explanation of interpersonal coordination patterns that emerge dynamically while social interactions unfold in real-time. According to these accounts, interpersonal coordination emerges through the informational coupling between individuals, which — following general laws of self-organizing systems — leads them to spontaneously synchronize or entrain their actions in time (e.g., Richardson et al., 2007).

Dynamical approaches to joint action offer powerful tools for quantifying the unfolding dynamics of interpersonal coordination and provide parsimonious explanations for emergent forms of coordination in social interactions. However, they remain mainly focused on the spatiotemporal coupling of continuous, usually rhythmic bodily movements between people that arise spontaneously, often without people explicitly planning or intending to do so. As such, they remain limited in their reach to explain instances of joint action that require people to plan and control discrete actions in relation to the actions of their co-actors or in relation to distinct goals co-actors are trying to achieve (Knoblich et al., 2011).

1.1.3. Cognitive accounts of joint action

While philosophical accounts miss spelling out what it takes for people to eventually carry out a joint action once they formed a shared intention to do so, dynamical accounts fall short in capturing goal-directed aspects of interpersonal coordination. Cognitive accounts of joint action fill this gap by spelling out the mental representations that translate higher-level plans to act jointly into discrete actions and outcomes co-actors need to produce in the here and now to

4

reach their goals. The formation of these representations, here referred to as *joint task representations*, are thought to underlay the immediate preparation for an upcoming joint action and provide the bedrock for subsequent coordination mechanisms operating during its actual performance (Pacherie, 2012; Sebanz & Knoblich, 2021; Vesper et al., 2010).

Instead of focussing on shared intentions, cognitive accounts view *shared goals* as the central ingredient to explain people's capacity to act jointly[2]. Rather than referring to an abstract plan of doing something together, shared goals refer to concrete outcomes to be achieved as the common effect of the coordinated action contributions of multiple individuals (Butterfill, 2012; Pacherie, 2013; Sacheli et al., 2015). Thus, joint task representations are thought to capture the joint action outcomes co-actors are trying to achieve together, while specifying the action contributions co-actors need to execute in order to reach them. As such, joint task representations provide organizing structures that prepare the coordination of goal-directed joint actions by specifying the necessary contributions that co-actors need to execute to reach a shared goal. Furthermore, they provide control structures for ongoing coordination during the execution of a joint action by governing action monitoring and prediction processes (Knoblich et al., 2011; Vesper et al., 2010). However, how co-actors construct representations of shared goals and how these representations structure peoples action planning and control processes remains a subject of ongoing theoretical discussion in joint action research (Butterfill, 2015; Keller et al., 2016; Pacherie, 2013; Pesquita et al., 2018; Pezzulo et al., 2025; Sacheli et al., 2018; Sebanz & Knoblich, 2009; Sinigaglia & Butterfill, 2022; Vesper et al., 2010). In the following section, I will lay out three theoretical proposals of the content and structure of joint task representations informed by these discussions, spell out how these different structures support interpersonal coordination in joint actions, and provide an overview of previous research findings in their support.

---

[2] Throughout the dissertation I use the terms *shared goals*, *joint goals* and *joint action goals* interchangeably.

1.2. Proposed structures of joint task representations

1.2.1. Minimal structure: JOINT = ME + X

A first sketch of a possible structure of joint task representations was provided by Vesper et al. (2010). With the aim to formulate a minimal representational architecture for joint action, Vesper and colleagues (2010) proposed that rudimentary forms of joint action could be supported by task representations that merely represent the desired outcome of the joint action and an individual's own contribution needed to achieve it, with the only additional requirement being that the individual knows that the outcome cannot be achieved by acting alone. I.e., joint task representations may simply encode the joint outcome to be achieved, an individual's own contribution to it and some placeholder representing an unspecified additional contribution by some external agent or force. This minimal structure can be expressed by the formula JOINT = ME + X. Figure 1-1 depicts a graphical illustration of the minimal structure of joint task representations.
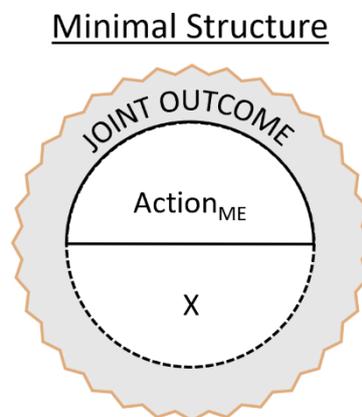
Figure 1-1. Graphical illustration of the minimal structure of joint task representation. Adapted from Vesper et al. (2010).

Such a minimal structure of joint task representations could support rudimentary forms of goal-directed coordination between co-actors by providing a common perceptual reference point to which each individual co-actor plans and monitors their individual task contributions. E.g. to clean up the department kitchen table after lunch with colleagues, each person involved

could simply take care of their own dishes, monitoring only if the shared goal of an empty table is reached, without having a detailed representation of others' contributions to the joint outcome. Furthermore, as argued by Vesper et al. (2010) this minimal structure of joint task representations may also underlay proto forms of joint actions non-human primates are able to take part in (e.g., Melis et al., 2006).

Direct empirical tests of the minimal structure have been scarce. However, a range of studies provide evidence for one of its main underlying assumptions, namely that co-actors represent their own contributions to a joint action in relation to outcomes that depend on additional contributions external to oneself. E.g., Loehr and Vesper (2016) showed that if people learned to play simple piano melodies as part of a musical duet performed together with an accompanist, they show a drop in performance when asked to reproduce the learned melody alone without the accompanying auditory feedback of their counterpart. This finding indicates that task performance in joint action contexts is predominantly guided by representations of joint rather than individual action outcomes. This is also supported by findings of Loehr et al. (2013). Investigating how pianists monitor the unfolding of musical duets played together with another person, Loehr et al. (2013) showed that duetting pianists are especially sensitive to performance errors that alter jointly rather than individually determined outcomes of their musical duet, such as the harmony of jointly produced chords.

1.2.2. Parallel structure: JOINT = ME + YOU

The minimal structure offers a parsimonious way of representing a joint task that may suffice to guide joint task performance under favourable conditions. However, in many cases, performing a joint task seems to require more than just caring about one's own part in it. Most often, joint task performance will benefit if co-actors have some grasp of the actions performed by their task partners as well. As such it seems highly likely that joint task representations take

on more elaborated structures that take the actions of other co-actors into account as well. Indeed, there is ample empirical support for this assumption.

The foundation for the idea that others' actions can figure into people's own action planning and control processes was laid out by common coding accounts of perception and action (Hommel et al., 2001; Jeannerod, 1999; Prinz, 1990, 1997). Extending the idea that actions are represented in terms of their sensory consequences (c.f., ideomotor theory; Greenwald, 1970; James, 1890), these accounts proposed that own and others' actions are represented in a common representational format that underlies the perception, execution and imagination of actions alike. This implied that perceiving or anticipating the action of another person activates the same representations people would use to execute similar actions themselves, a view that has gained direct support from neuroscientific findings on action mirroring (see Rizzolatti & Sinigaglia, 2016 for a comprehensive review). By implying that own and others' actions get represented in a commensurable format, the common coding principle provides a representational domain that allows people to represent others' actions in terms of their own action capacities. This is thought to provide a representational platform that enables the integration of others' actions into people's own action planning and control processes (Sebanz & Knoblich, 2009).

Initial evidence that people integrate others' actions into their own action planning and control processes has been provided by research on task sharing (c.f., Dolk & Prinz, 2016; Wenke et al., 2011 for reviews and critical discussion). These studies investigate how performing independent parts of a task next to another co-actor modulates individual task performance compared to when performing one's own part of the task alone. In short, these studies have shown that people are affected by response features (Sebanz et al., 2003; Tsai et al., 2006), stimulus-response rules (Atmaca et al., 2008, 2011; Sebanz et al., 2005), loci of attention (Böckler et al., 2012; Welsh et al., 2005) and action-outcome mappings (Pfister et al., 2014) of their co-actors, despite them being irrelevant to and even interfering with people's

own part of the task. These findings have been interpreted as a profound tendency of people to form task representations that *co-represent* others' part of a task next to one's own, even if coordination between co-actors towards a shared goal is not actually required.

Crucially, subsequent research has shown that this basic tendency to co-represent others' tasks and actions plays a functional role in joint task contexts that involve interpersonal coordination towards shared goals. In a series of studies, Kourtis et al. (2013, 2014) showed that preparation for an upcoming joint action, requiring two co-actors to coordinate reaching movements in space and time, recruits neural resources that extend those underlying the preparation of people's individual reaching movement performed alone. Instead, neural signatures of joint action preparation were found to resemble those underlying the preparation of both reaching movements performed bimanually, suggesting that people planned both their own and their co-actors task contribution as if performing both parts themselves. Relatedly, Vesper et al. (2013, 2014) showed that co-actors adjust performance parameters of their own contribution to a joint action outcome based on prior information about the upcoming action of their co-actors, suggesting that co-actors are able to integrate information about the upcoming task contributions of their partners to plan and adjust their own actions (see Schmitz et al., 2017; Meyer et al., 2013 for related findings).

The evidence for common coding and co-representation of one's own and others' task contributions suggests that, in many cases, joint task representations encode not just an individual's own contributions to a joint action outcome but also those of their co-actors. But how could joint task representation that encode both own and others' task contributions be structured? One possibility is that joint task representations encode own and others' task contributions as distinct subplans, such that each individual represents their own contribution to the joint outcome in parallel to the individual contributions taken care of their co-actors (c.f., Bratman, 1992; Butterfill, 2015; Pacherie, 2012). That is, joint task representations may encode the joint outcome to be achieved, an individual's own contribution to it and the individual

contributions expected from the other co-actors involved in the joint action. This parallel structure of joint task representations can be captured by the formula JOINT = ME + YOU. Figure 1-2 depicts a graphical illustration of the parallel structure of joint task representations.

Parallel Structure

Figure 1-2. Graphical illustration of the parallel structure of joint task representation.

How would this parallel structure support interpersonal coordination in joint action? Encoding others' expected task contributions in addition to one's own, the parallel structure enables co-actors to predict and monitor how own and others' individual task contributions unfold in space and time (Sebanz & Knoblich, 2009). This is thought to be achieved by the common coding of one's own and others' actions, which enables people to apply internal models in their own motor system to simulate and predict the timing and outcomes of the upcoming and ongoing actions of their co-actors (Wilson & Knoblich, 2005; Wolpert et al., 2003). The parallel structure implies that people generate such predictions separately for their own and their co-actors' task contributions, enabling people to monitor if their own and their co-actors' individual performance matches the respective subplans each co-actor is taking care of in bringing about the shared goal. This allows co-actors to monitor if everyone sticks to their predetermined role in a joint action and to adjust the timing and completion of their own task contributions to those of others if necessary.

1.2.3. Relational structure: JOINT = RELATIONS(ME + YOU)

While offering a possible way to integrate others' task contributions with own action planning and control processes, the parallel structure of joint task representations still seems to be out of reach to capture crucial characteristics of some types of joint actions. While enabling people to predict and monitor individual task contributions of self and others, the parallel structure seems to fall short in capturing the interdependency between own and others' individual task contributions that are characteristic for many forms of joint actions. E.g., balancing a tray of champagne glasses with another person requires *my* actions to be directly contingent on *your* actions and vice versa to meet relational constraints between *our* actions that keep the tray at level. As this example highlights, the achievement of a shared goal is often dependent on specific *relations* between individual task contributions of self and others that capture the task performance of the group. A set of empirical findings indicates that joint task representations may capture these relational aspects of joint task performance as well.

Initial evidence that co-actors form more than separate representations of own and others' task contributions has been provided by research on group imitation (see Cracco, 2025; McEllin et al., 2018 for review). These studies showed that executing a joint action with another co-actor is facilitated by observing one's own and one's co-actor's parts being modelled by another dyad compared to observing only one's own part being modelled by a single actor (Tsai et al., 2011) or by observing both parts of the joint action being modelled by single individual acting bimanually (Ramenzoni et al., 2014). These findings indicate that joint task performance benefits from the opportunity to observe relations between own and others individual task contribution and suggest that joint actions get represented in relation to group-level rather than individual-level performance.

Further evidence that co-actors represent relations between their individual contributions to a joint task comes from a set of studies showing that interfering effects of observing another

person executing an action that is incongruent to an action concurrently planned or performed by oneself (*visuomotor interference effect*, c.f., Brass et al., 2000; Kilner et al., 2003) are diminished (Clarke et al., 2019; Formica & Brass, 2024; Rocca et al., 2023), absent (Sacheli et al., 2018, 2019) or even reversed (van Schie et al., 2008) in joint action contexts that require co-actors to perform complementary actions to achieve a shared goal. These findings could be explained by assuming that joint task representations encode how own and others individual task contributions must relate to each other to produce the joint action outcomes that correspond to co-actors' shared goal.

Direct evidence for this claim has recently been provided in a study by Kourtis et al., (2019). In this study, two co-actors were asked to synchronize arm reaching movements that resulted in different postural configurations of their palms facing either in the same or in different directions. Co-actors were found to initiate their actions faster and more synchronized when they received prior information about the relation between their individual actions, i.e., whether they were about to perform the same or different arm reaching movements. Notably, this was even the case, if they did not know which individual action to prepare themselves. This finding indicates that co-actors formed task representations that encoded not only each other's individual task contributions but also how they relate to each other at the level of the group.

The findings reviewed above suggest another way in which own and others task contributions could be embedded within structured representations of co-actors' joint tasks. Instead of representing own and others' task contributions as parallel subplans that must be completed individually to achieve a shared goal, joint task representations could encode primarily how joint action outcomes are achieved by the aggregate action of the group, that is by agent-neutral features of co-actors' joint task performance that capture group-level relations between individual action contributions of self and others (c.f., Pezzulo et al., 2025). I.e., joint task representations could encode the joint outcome to be achieved and the relations between individual task contributions that must be met to reach them. This relational structure can be

captured by the formula JOINT = RELATIONS(ME + YOU). Figure 1-3 depicts a graphical illustration of the parallel structure of joint task representations.
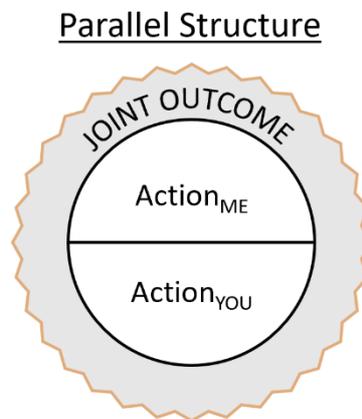
Relational Structure
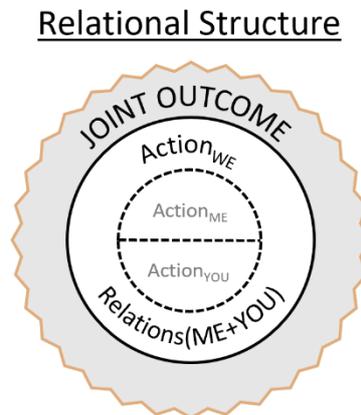
Figure 1-3. Graphical illustration of the relational structure of joint task representations.

How could this relational structure support interpersonal coordination in joint action beyond the parallel structure? The relational structure would offer a parsimonious way to predict and monitor the joint outcomes of co-actors' coordinated task contributions. Instead of running separate predictive models for own and others' individual task contributions in parallel and integrating their respective outcomes to evaluate whether the joint action goes as planned (c.f., Keller et al., 2016; Pesquita et al., 2018), the relational structure would allow co-actors to predict the outcomes of co-actors combined actions by means of a single model of their joint task performance (Knoblich & Jordan, 2003; Pezzulo et al., 2017).

Furthermore, the relational structure would enable own action selection and control processes to be directly contingent on expected or perceived task contributions of other co-actors. This could enable co-actors to flexibly adjust their own actions to those of their co-actors by constraining action alternatives to complementary responses (Candidi et al., 2015; Sacheli et al., 2018; Sartori & Betti, 2015), by disambiguating expected task contributions of self and others (Kourtis et al., 2019), and by allowing co-actors to correct others' mistakes through compensatory responses (Musco et al., 2025; Sacheli et al., 2021, 2022).

Finally, the relational structure could enable dynamical role allocations during joint action. Representing a joint task from a group-level perspective, the relational structure is thought to encode agent-neutral invariants of the required task contributions that realize the shared goal. As such, the relational structure offers a representational level at which agent-specific task contributions of self and others are not yet specified but can be dynamically deduced and redistributed given situational constraints and current task demands (c.f., Gordon et al., 2023).

1.3. Research aims

Much of the previous work on the representational foundations of joint action has been focussed on the conditions and consequences of co-representing others' shares of a joint task next to one's own. Consequently, the notion of co-representation still receives centre stage in cognitive accounts of joint action (c.f., Miss et al., 2022; van der Wel et al., 2021). The work reviewed in the previous section indicates, however, that cognitive accounts of joint action limited on co-representation fall short in capturing a set of important findings in joint action research. The relational structure of joint task representation described in the previous section could thus offer a vital extension for cognitive accounts of joint action that captures the intuition and empirical indications that there is something more to joint action than a YOU and a ME, namely an irreducible collective mode of acting as a WE (Gallotti & Frith, 2013).

However, direct evidence for the relational structure of joint task representations is still sparse and empirical findings that are in line with the proposal that co-actors form task representations that encode group-level relations between individual task contributions of self and others remain limited to a small number of studies reviewed above (see also della Gatta et al., 2017). Against this background, the aim of the present dissertation was to expand research on the content and structure of joint task representations by putting the relational structure of joint task representations to empirical test. To that end, we conducted four empirical research

projects that were all designed to assess ***whether joint task representations encode relations between individual action contributions of self and others***. In the following, I will provide an overview of the individual research projects comprising this dissertation and describe how each of them addresses this overarching research question.

1.3.1. STUDY I: Do relations between co-actors' individual task contributions shape imitation between groups?

The first study, presented in Chapter 2, approached the content and structure of joint task representations by investigating anticipated imitation effects in social interactions between groups.

The starting point for this study is the idea that people use anticipated actions evoked in other people to initiate and control their own contributions to a social interaction (Kunde et al., 2018). An empirical demonstration of this principle is the finding that being predictively imitated by another person facilitates the execution of to-be-imitated responses, a phenomenon known as *anticipated imitation* (Pfister et al., 2013). Anticipated imitation effects demonstrate that overlap or conflict between features of a planned response and features of an anticipated response evoked in another person shapes task performance in social contexts, providing insights into the content and structure of the representations guiding action selection and control in social interactions.

Study I extended these ideas to social interactions between groups to probe the content and structure of joint task representations. The basic idea behind the study is the following: If co-actors represent group-level relations between their individual contributions to a joint task, anticipating another group of co-actors imitating these relations should facilitate task performance, while anticipating another group of co-actors executing different relations between their individual task contributions should hinder task performance.

Study I tested this prediction in an online experiment in which participants performed a stimulus-response task either with or without another co-actor while their task responses were predictively imitated by either one or both member(s) of another dyad. By manipulating congruency between the number of responses performed by participants' own group and the number of imitative responses performed by another group, Study I tested whether relations between co-actors' individual task contributions shape anticipated imitation effects in social interactions between groups.

1.3.2. STUDY II: Is action-outcome learning sensitive to relations between co-actors' individual task contributions?

The second study, presented in Chapter 3, approached the content and structure of joint task representations by investigating action-outcome learning in synchronous joint actions.

Starting from the basic assumption that goal-directed action requires people to learn and represent relationships between their actions and the outcomes they produce, Study II set out to investigate how co-acting individuals learn and represent relationships between their joint actions and the joint action outcomes they produce together. Drawing from research on ideomotor learning (c.f., Moeller & Pfister, 2022 for a comprehensive overview), Study II tested how the repeated experience of contingencies between synchronous motor responses of two co-actors and ensuing sensory outcomes leads co-actors to link jointly produced action outcomes to features of their preceding responses. The basic idea behind the study is the following: If co-actors represent group-level relations between their own and their partner's task contributions, they may link the production of jointly produced action outcomes to group-level features of their joint task contributions rather than to individual-level features of their own task contributions viewed in isolation.

To test this prediction, Study II deployed a two-stage action-outcome learning task (c.f., Elsner & Hommel, 2001) adapted to a synchronous joint action context. In a learning phase,

two co-actors had to synchronize individual button press responses to produce a series of distinct two-tone chords that were mapped to individual-level features of each co-actor's own response contribution, as well as to group-level features of co-actors' joint response configurations. In a subsequent test phase, the same two-tone chords were presented as imperative stimuli that required both co-actors to respond to the previously produced tone outcomes with a stimulus-response mapping that either preserved or reversed the individual and/or the joint action-outcome mappings of the preceding learning phase. By manipulating compatibility between learning and test phase mappings at the group-level, Study I tested whether action-outcome learning in synchronous joint action is sensitive to group-level relations between co-actors' individual response contributions.

1.3.3. STUDY III and IV: Do representations of relations between co-actors' individual action outcomes modulate imitative response tendencies?

The third and fourth study, presented in Chapter 4.1. and 4.2., approached the content and structure of joint task representations by investigating if acting towards a shared goal that specifies relations between the outcomes of own and others' individual task contributions allows co-actors to overcome tendencies to imitate each other.

Study III and IV originate from a central empirical paradox in research on joint action. On the one hand, an extensive body of research shows that people possess a profound tendency to imitate the actions they observe in others (Heyes, 2011). On the other hand, jointly acting towards a shared goal requires co-actors frequently to coordinate actions that are incongruent but complementary to each other so that tendencies to imitate each other's actions often become detrimental for joint task performance (Sartori & Betti, 2015). Interestingly, a growing number of studies has shown that joint action partners can resolve this tension: If co-actors coordinate their actions towards a shared goal, their tendency to imitate each other's actions is reduced (Clarke et al., 2019; Rocca et al., 2023) or dissolves completely (Sacheli et al., 2018, 2019).

However, *why and how* acting towards a shared goal reduces imitative response tendencies still remains elusive.

Against this background Study III and Study IV tested the hypothesis that shared goals reduce imitative response tendencies by specifying relations between action outcomes of own and others' individual task contributions. The basic idea behind this hypothesis is the following: If joint task representations encode relations between own and others' individual contributions to a joint action outcome, observing the task contribution of a co-actor should prime the selection of complementary task contributions that satisfy the relational constraint between them imposed by the shared goal.

To test this prediction, Study III and Study IV deployed a sequential card-selection task (c.f., Ondobaka et al., 2012) in which participants took turns with a virtual co-actor selecting one of two action targets to achieve either individual or joint action goals. By manipulating congruency between low-level movement goals (selecting the left or the right of their targets) and between higher-level action goals (selecting the target with the higher or lower value) of both co-actors, the studies assessed participants' imitative response tendencies at different hierarchical levels of action representation. By instructing participants to either work towards individual or joint action goals in this task, both studies tested whether the introduction of a joint action goal that specified relations between the outcomes of co-actors' individual task contributions is able to reduce imitative response tendencies at different levels of action representation.

2. STUDY I:

Anticipated Imitation Between Groups

2.1. Introduction

An extensive body of research shows that humans possess a marked tendency to imitate the behaviours of others, fulfilling crucial instrumental functions for social affiliation, observational learning, action prediction and interpersonal coordination (Heyes, 2011; Lakin et al., 2003; Wilson & Knoblich, 2005). Given its functional relevance, research on imitation has long served as a window into the cognitive representations and mechanisms that shape and support social interaction.

Prominent mechanistic explanations of our propensity to imitate the actions of others derive from ideomotor accounts of human action control (Brass & Heyes, 2005a; Wohlschläger et al., 2003). These accounts build on the idea that actions are controlled by representations of their anticipated sensory outcomes, thus assuming a shared representational system for action and perception (Greenwald, 1970; Hommel, 2013; Hommel et al., 2001; Prinz, 1997; Shin et al., 2010). In the context of social interactions, a critical implication of ideomotor accounts is that actions performed by *other* people can directly affect the execution of one's *own* actions: As the actions performed by other people will most often resemble the sensory outcomes of our own actions, observing or anticipating others' actions is thought to activate corresponding action representations in the observer.

Ample empirical evidence for this assumption is provided by two related phenomena. First, a large body of research on *automatic imitation* demonstrates that action execution is facilitated by the concurrent observation of similar (i.e., congruent) and impaired by the concurrent observation of dissimilar (i.e., incongruent) actions performed by other agents (see Cracco et al., 2018 for a recent meta-analysis). These findings indicate that perceptual

representations of actions triggered by *action observation* activate corresponding motor representations in the observer, facilitating or interfering with the execution of congruent or incongruent action plans respectively.

Second, research on *anticipated imitation* has provided converging evidence that not only external observation but also the *internal anticipation* of other people's actions can facilitate or interfere with the execution of congruent or incongruent action plans respectively (Müller, 2016; Pfister et al., 2013, 2017; Weller et al., 2019). In particular, these studies demonstrated that being predictively imitated by another agent facilitates performance of to-be-imitated motor responses compared to motor responses that are predictively followed by counter-imitative (i.e., incongruent) responses of another agent (see Pfister et al., 2025 for review). Resembling findings from response-effect compatibility studies (e.g., Ansorge, 2002; Keller & Koch, 2006; Kunde, 2001; Pfister & Kunde, 2013), these findings convincingly show that *anticipated* action outcomes — including predictable responses of other people (c.f., Kunde et al., 2018) — play a causal role in controlling one's actions, facilitating or interfering with motor plans that share compatible or incompatible features with its anticipated outcomes respectively.

2.1.1. From dyadic imitation to imitation of multiple agents

Most research on imitation is focussed on dyadic social interactions in which individuals observe or anticipate the actions of a single agent (see Figure 2-1, Panel A). This focus, however, neglects the fact that many of the social interactions people take part in involve interactions with larger groups of people. E.g., teachers must monitor the behaviour of multiple students in parallel, soccer players must take the actions of multiple teammates and opponents into account to plan their next move, and conductors must orchestrate the actions of a whole ensemble of different musicians at once. Therefore, recent research efforts have started to extend research on dyadic imitation by investigating how observing or anticipating the actions

of multiple agents influence individuals' action planning and control processes (see Cracco, 2025 for review).
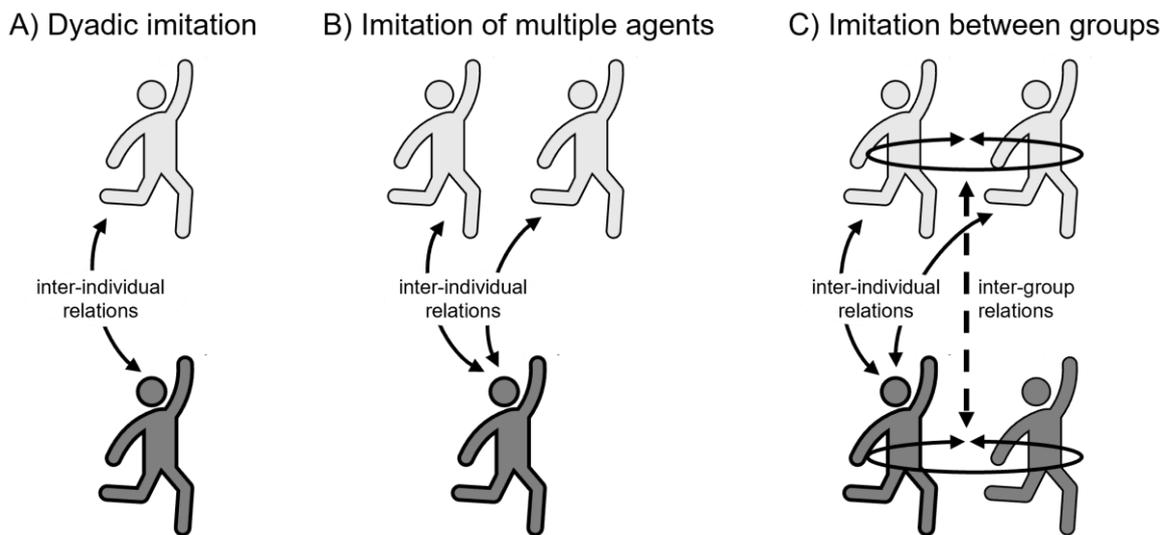
Figure 2-1. Illustration of different congruency relations between own and others' actions in social interactions between A) two individual agents, B) an individual agent and multiple other agents, and C) two opposing groups of jointly acting individuals. Solid arrows represent inter-individual relations between actions of single agents. The dotted arrow represents inter-group relations between actions performed by one group and those performed by the other.

Investigating how automatic imitation effects scale with the number of observed actions, Cracco and colleagues (2015) showed that observing *identical* actions performed by more than one agent produces stronger automatic imitation effects, which were found to increase asymptotically with the number observed actions (Cracco & Brass, 2018b), ceiling at group sizes involving up to ten agents (Cracco et al., 2022). Moreover, studies investigating effects of observing *different* actions of multiple agents have shown that concurrently observed actions can be reliably decoded from brain activity in the observer's motor system (Cracco et al., 2019), elicit concurrent imitative response tendencies (Cracco & Brass, 2018a), and activate conflicting action representations in the observer (Cracco et al., 2019, 2021). Taken together, these studies indicate that people represent the actions of multiple observed agents simultaneously in their motor system and show that basic action planning processes can be influenced by the observation of multiple agents at once.

In contrast to the growing body of studies investigating automatic imitation of multiple agents, less is known about the behavioural effects of being imitated by more than one agent. An initial investigation of this topic by Galang et al. (2024) showed no evidence that anticipated imitation effects are influenced by the number of imitators, with recent findings by Neszmélyi and Pfister (2024) showing similar results.

2.1.2. From imitation of multiple agents to imitation between groups

Extending research on dyadic imitation, the research reviewed in the preceding section focussed on situations in which *individuals* observe or anticipate the actions of multiple agents (see Figure 2-1, Panel B). However, when multiple people come together, we may encounter a group of people not just as individuals, but often as part of a group ourselves. I.e., we often observe or anticipate the actions *of* multiple agents, while interacting *with* other agents at the same time (see Figure 2-1, Panel C). E.g., as a couple in a ballroom dance class, we may try to copy the joint moves performed by another dance couple; or as part of a tennis double, we may anticipate the next moves of our two opponents, while coordinating the running paths on our half of the pitch.

Thus, an interesting question to ask is how being part of a group may shape effects of observing or anticipating another group of agents. As a caveat to this question, one may wonder why observing or anticipating the actions of a group of people should have different effects when acting as part of a group oneself compared to when acting merely as a single individual. One hypothesis could be that effects of action observation or anticipation are solely driven by *inter-individual relations* between the actions performed by oneself and the actions performed by all other individuals in the interactive scenario (c.f., Figure 2-1, Panel C, solid arrows). As such, the two-groups scenario would not be much different than the scenario in which an individual observes or anticipates the actions of multiple other agents, and imitation effects may simply scale with the total number of agents in the scene.

Crucially though, research findings in the domain of *joint action* suggest that individuals performing a task together with other co-actors (i.e., as part of a group) form action representations that integrate their co-actors' actions into *joint task representations* that encode relations between one's own and one's co-actors' actions emerging at the level of their group-level performance (sometimes referred to as "we-representations"; c.f., Kourtis et al., 2019; Marschner et al., 2024; Sacheli et al., 2018; Sinigaglia & Butterfill, 2022). Importantly, these representational peculiarities of joint as compared to individual action, suggest an alternative hypothesis about how effects of action observation and anticipation might be modulated in social interactions between groups (c.f., McEllin et al., 2018; Milward & Sebanz, 2016). When people perform a task together with other co-actors while observing or anticipating the actions of another group of jointly acting individuals, not only *inter-individual relations* between own and others' actions may affect people's action planning and control processes but also *inter-group relations* that emerge between the group-level performance of one group relative to that of the other group (see Figure 2-1, Panel C, dashed arrow).

Testing this possibility, Tsai et al. (2011) conducted an experiment in which participants performed a task alongside a confederate co-actor in which they imitated the movements of two hand stimuli depicted on a computer monitor in front of them. In different conditions, participants performed the task either without (individual response) or together with the confederate (joint response), in response to observing either one or both hands moving on screen. Interestingly, results showed that response times were fastest when the number of movements observed on screen matched the number of imitative responses performed by participants and their confederate, respectively. That is, *individual* responses were faster when imitating movements of one vs. two hand(s), while *joint* responses were faster when imitating movements of two vs. one hand(s) (see Essa et al., 2019 and Ramenzoni et al., 2014 for similar findings). Interestingly, this effect emerged only if the observed hand stimuli were displayed as belonging to another dyad but not when displayed as belonging to a single individual. This

finding shows that in interactions between groups, imitation effects are indeed shaped by inter-group rather than inter-individual action relations, affecting how observed actions are mapped to actions performed in response to them. Crucially, this finding seems to indicate that actions performed together with others are coded qualitatively different than actions performed alone, namely in relation to group-level rather than individual-level performance.

2.1.3. The present study

The findings by Tsai et al., (2011) provide initial evidence that acting as part of a group changes how we process and respond to the actions we observe in others. But does this justify the conclusion that actions performed together with others are planned and represented qualitatively different than actions performed alone? One objection could be that the group-level congruency effects found by Tsai et al. (2011) are caused by *action observation*, i.e., by the external perception of individual or joint actions, priming the subsequent execution of individual or joint action plans respectively. As such, they may show that individual and joint action plans become *exogenously* activated by *observing* individual or joint actions in others. Yet, they can provide only limited insights into the *endogenous* mechanisms that would drive the activation of individual or joint action plans in a goal-directed, self-initiated way.

Crucially, the central hypotheses of ideomotor theories is that goal-directed actions are initiated and controlled by *anticipations* of their perceivable outcomes — that is, by future rather than currently perceived events (Kunde, 2001a). Applied to social interactions between groups, tapping into the anticipatory nature of human action planning would thus require to study effects of *anticipating* rather than *observing* the actions of another group of agents. This was precisely the aim of the present study.

To this end, we converted the experimental design by Tsai et al. (2011) to an anticipated imitation task. Rather than responding to observed actions of one or two agents either alone or together with another co-actor, in our experiment, participants performed a stimulus-response

24

task either alone or together with another agent, while their responses where predictively imitated by either one or two agents of an opposing dyad. Figure 2-2 provides an overview of our experimental setup and design.

A)

B)

Imitative Response
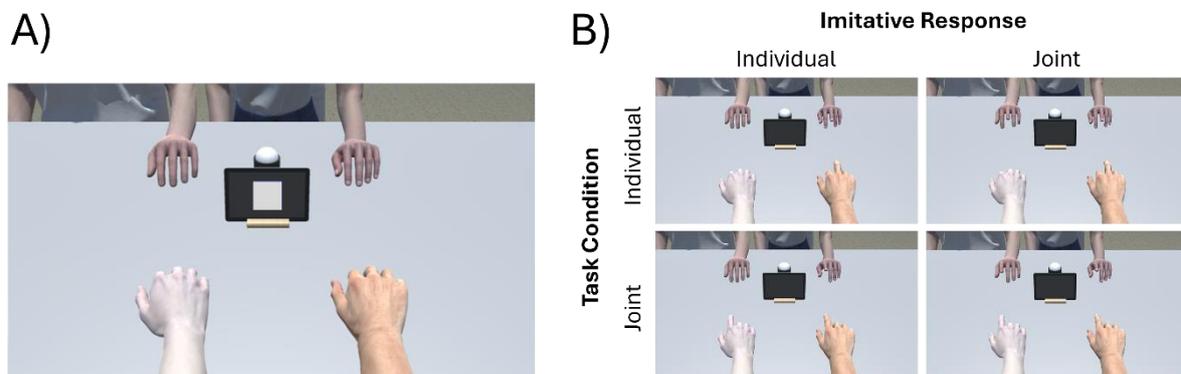
Figure 2-2. Experimental setup and design. A) Participants performed a two-alternative forced choice task in a virtual environment depicting four animated avatar hands. Participants task was to lift the index or middle finger of their right hand in response to shape stimuli (square or circle) appearing on screen. One of the two hands displayed in the lower half of the scene was introduced as representing participants' own hand which always mirrored participants' finger lift responses as they performed them. The hand next to the participant's hand was introduced as belonging to a virtual partner performing the task alongside participants. The two hands displayed above participant's and their partner's hand were introduced as belonging to two virtual agents observing participant's and their partner's finger lift responses with the aim to imitate them. B) In different blocks, participants performed the task either without (individual task condition) or together with their virtual partner (joint Task condition), while either one (individual imitation) or both (joint imitation) of the opposing agents imitated participants' finger lift responses.

Under the assumption that anticipated responses of other people get employed to initiate and control one's own contribution to a social interaction (Kunde et al., 2018) our experimental setup allows us to test whether actions performed together with others are indeed planned and represented differently than actions performed alone. If actions performed together with others are planned and represented in relation to group- rather than individual-level performance (group-level hypothesis), performing an action *together* with another agent (joint action) should be facilitated by anticipating an imitative response by another dyad of jointly acting agents (joint imitation) but be impaired by anticipating an imitative response by a single member of another dyad (individual imitation). At the same time, performing an action *without* another agent (individual action) should instead be facilitated by anticipating an imitative response by

25

a single member of another dyad (individual imitation) but be impaired by anticipating an imitative response by another dyad of jointly acting agents (joint imitation). In contrast, if actions performed together with others remain represented in relation to one's individual-level performance alone (individual-level hypothesis), facilitatory effects of anticipating imitative responses by other agents should be expected to simply scale with the number of imitative responses performed by all other agents. That is, performing an action may benefit from performing it together with another co-actor vs. alone and from being imitated by two vs. one member of another group. Figure 2-3 illustrates the predictions of the group-level and individual-level hypothesis respectively.
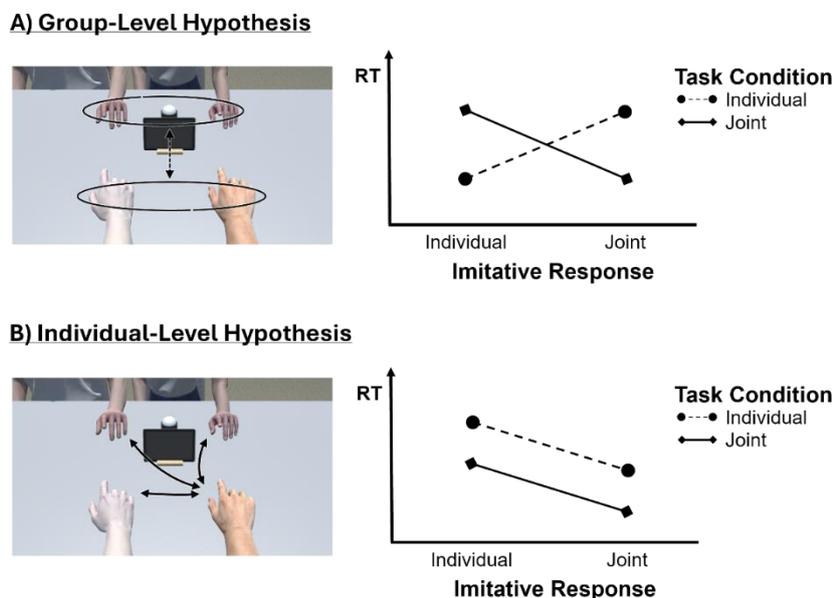
Figure 2-3. Contrasting hypothesis and resulting predictions for the current study. A) According to the group-level hypothesis, actions performed together with others are represented in relation to group-level performance (illustrated by the two ellipses). Acting together with another person (Joint Action) should then be facilitated by anticipated imitation of another dyad (Joint Imitation), while acting alone (Individual Action) should be facilitated by anticipated imitation of another individual agent (Individual Imitation). Participants' task performance should thus be driven by numerical congruency between the actions performed by their dyad and the actions performed by the imitating dyad, i.e., by congruency relations between groups rather than between individuals (illustrated by the dashed arrow). B) According to the individual-level hypothesis, actions performed together with others remain represented in relation to individual-level performance. Task performance should then be expected to scale with the number of imitative responses by all other agents in the interactive scene, i.e., by additive effects of congruency relations between individuals (illustrated by the solid arrows). Note: RT = response time as a measure of task performance (lower values indicate better performance).

2.2. Methods

2.2.1. Participants

In total, $N = 250$ participants took part in the experiment. The sample size was determined based on an a priori power analysis and pragmatic concerns regarding resource constrains. We wanted to ensure that our study would be sufficiently powered to detect a within-subjects interaction effect of size $\eta_p^2 = .03$ in a 2 x 2 repeated measures ANOVA with 80% power at an alpha level of .05. The effect size estimate was based on a series of pilot studies. To calculate the required sample size to find such an effect, we treated the two-way interaction effect as a paired-sample difference of difference scores between factor levels, transforming our effect size estimate of $\eta_p^2 = .03$ to dz ~ 0.18 in a paired samples t-test (c.f., Langenberg et al., 2023). Conducting an a priori power analysis with these input parameters in G*Power (Faul et al., 2007) we calculated a required sample size of $N = 245$. Assuming that we would need to remove a number of participants from our sample due to our preregistered exclusion criteria (see below), we eventually decided to collect data from $N = 250$ participants. From the final sample 15 participants were eventually excluded which led to a final sample size of $N = 235$ participants (mean age = 30.0 years, standard deviation = 5.9 years; 42 % female, 58 % male).

Participants were recruited on the online research platform Prolific (prolific.com). All participants were right-handed, had normal or corrected-to-normal vision, resided in the UK and were fluent in English. All participants gave their informed consent before participation and were compensated with 3 GBP for completing the experiment in approximately 20 minutes. Ethical approval for the study was granted by the Psychology Ethics Committee of Humboldt University Berlin.

2.2.2. Apparatus and stimuli

Participants performed the experiment online in a browser on their own personal computer. Experiment files and data were stored on a JATOS server (Just Another Tool For Online Studies, Lange et al., 2015) hosted by Humboldt University Berlin. The experimental task was built in Unity and was controlled by a custom made C# script. The task was set in a virtual 3D environment depicting four animated avatar hands resting on the surface of a white table (see Figure 2-2, Panel A). One of the two avatar hands in the lower part of the scene was introduced as representing participant's own hand that would mirror their responses in the experimental task. The other hand in the lower part of the scene was introduced as belonging to a virtual partner that would perform the experimental task together with the participant. The two avatar hands at the top of the scene were introduced as belonging to two virtual agents that would observe the participant's and their partner's finger lift responses in order to imitate them. Participants responded by operating the "N" and the "M" key on their keyboard with the index and the middle finger of their right hand respectively. Imperative stimuli instructing participants' responses were a circle and a square displayed in grey. Shape stimuli, trial-level instruction and feedback prompts were presented on the depiction of a small tablet monitor displayed in the center of the scene. All avatar models in the scene were purchased from the Unity Asset Store. Movement animations of the avatar hands were self-created with UMotion Pro, a Unity package for animating rigged avatar models. The duration of the movement animations (from resting state to the respective finger being up) was set to five frames, corresponding to ~ 165 ms assuming a display rate of 30 frames per seconds.

2.2.3. Procedure and design

Participants were instructed to respond to the presentation of two shape stimuli by lifting the index or middle finger of their right hand from their response keys respectively. The experiment started with a familiarization phase in which participants first trained the task alone

with only their avatar hand being displayed in the scene, followed by a series of training trials performed together with the virtual co-actor in which only participants' avatar hand and that of their virtual co-actor were displayed[3]. After performing eight successful training trials in each of these trial series, participants proceeded to the main part of the experiment. Here participants were informed that in different blocks, they would perform the task either with (Joint Action) or without their virtual partner (Individual Action) and that either one (Individual Imitation) or both (Joint Imitation) of the opposing avatars would imitate their finger lift responses.

The trial timeline for the main part of the experiment is illustrated in Figure 2-4. Each trial started with a "Hold" prompt displayed on screen, instructing participants to hold down both of their response keys. As soon as participants complied, the "Hold" prompt disappeared after 500 ms, followed by a fixation cross presented for a duration of 1000 ms. If participants released one of their keys during this period, an error message ("Too early") was displayed for 1500 ms and the trial started all over. Otherwise, the fixation cross was followed by the presentation of one of the shape stimuli, displayed until participant's response or until the end
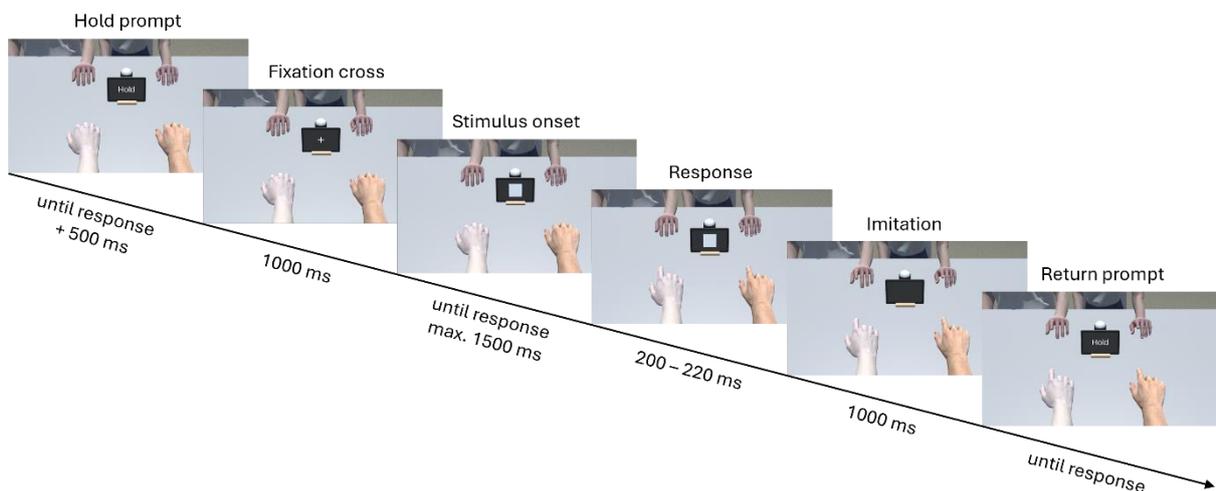
Figure 2-4. Trial timeline for the main part of the experiment. The example shows a trial in which participants responded to the shape stimuli together with their virtual co-actor (joint action) and where imitated by both members of the opposing dyad (joint imitation).

---

[3] The trial timeline in the familiarization phase corresponds to that described for the main part of the experiment, just without an imitative response of the opposing avatars, which were not yet visible in this phase of the experiment.

of a maximum response window of 1500 ms. As soon as participants lifted one of their fingers from their response keys within this window, a corresponding finger lift animation of their avatar hand was triggered. In Individual Action blocks, the avatar hand of the virtual partner remained in resting position (all fingers down) throughout the whole block. In Joint Action blocks, participants' responses also triggered a finger lift animation of their partner's avatar hand, which was played after a jittered delay ranging between 0 and 20 ms. If triggered, the partner's avatar hand always lifted the correct finger.

If participants responded correctly, either one (in Individual Imitation blocks) or both (in Joint Imitation blocks) of the opposing avatars lifted the same finger as participants. The respective animations were triggered after an interval ranging between 200 and 220 ms. The final outcome, i.e., two to four hands lifting the same finger, was displayed for at least 1000 ms after which another "Hold" prompt instructed participants to press and hold down both of their keys again. As soon as participants complied, their avatar hand lowered the respective finger again. Depending on the current block, the partner's avatar hand followed to lower their finger directly after and, after another delay ranging between 200 and 220 ms, either one or both of the two opposing avatars did so as well. The next trial started after an inter trial interval of 500 ms in which all four hands were displayed in their starting position again.

If participants lifted the wrong finger, an error message was displayed and a signaling lamp in the scene turned red for 1500 ms. Participants were informed that the red lamp would signal the opposing avatars to withhold their imitative response. The error message was followed by a "Hold" prompt, instructing participants to return to the starting position as described above. If participants failed to respond within the response window, their own and their partner's avatar hand, as well as those of the two opposing avatars remained in their resting position and a feedback message ("too slow!") was displayed for 1500 ms followed by the "Hold" prompt as described before.

30

To ensure that participants paid attention to the imitative responses of the opposing avatars, we implemented occasional catch trials, in which the imitating avatar(s) lifted their finger twice. If participants noticed this behavior, they had to release their remaining finger from their keyboard within a response window of 1500 ms. If participants failed to respond as described, a feedback message was displayed for 5000 ms, reminding participants what to do when noticing a double finger lift of the opposing avatar(s). The feedback message was then followed by the "Hold" prompt and related procedure as described above.

The study thus had a 2 x 2 fully within-subjects factorial design with Task Condition (Individual Action vs. Joint Action) and Imitative Response (Individual Imitation vs. Joint Imitation) as within-subjects factors. Participants performed 40 trials in each of the four blocks corresponding to the four experimental conditions resulting from our factorial design. Each block started with an instruction text, announcing whether participants were going to perform the block with or without their virtual partner and whether one or both of the opposing avatars were going to imitate their responses. Each block started with a series of training trials to familiarize participants with the structure of the block. The actual block started after two successful training trials. In blocks where only one of the opposing avatars would perform an imitative response, the side of the imitating avatar (i.e., the left or the right one) changed after the first half of the block (which was announced and preceded by a new series of training trials)[4]. Stimuli presentation (square vs. circle) was pseudo-randomized to account for balanced numbers of both stimuli in each half of each block. Four of the 40 test trials in each block were catch trials, two randomly allocated to the first and two randomly allocated to the second half of each block (the required response on catch trials was counterbalanced in both halves of each block). Stimulus-response mappings were counterbalanced between participants. Furthermore, we counterbalanced the position of participants' avatar hand within the scene, which was

---

[4] This was done to control for possible spatial compatibility effects while ensuring predictability of the imitative responses.

displayed either on the lower left or the lower right half of the scene. We also counterbalanced the position of the avatar imitating participants responses in the first and second half of the Individual Imitation blocks. Finally, block order was counterbalanced between participants using a balanced Latin square design with four possible block orders[5].

2.2.4. Analysis plan

Training trials were not included in our analysis. From the 40 test trials in each block we excluded trials with response omissions and trials in which participants responded by releasing both of their keys at once (1.1 % of all trials). From the remaining trials, we excluded trials with a response time (RT) that was more than 3 standard deviations away from participants' individual mean RT aggregated across all trials in the experiment (1.4 % of remaining trials). On the participant level, we excluded participants who failed to respond to any of the catch trials in at least one of the four blocks (affecting 15 participants) and those with an overall response error rate (ER) above 20% (affecting no further participant). For the remaining sample, participants' RT and ER were aggregated and analyzed by means of separate 2 x 2 repeated-measures ANOVAs with Task Condition (Joint Action vs. Individual Action) and Imitative Response (Individual Imitation vs. Joint Imitation) as within-subjects factors. Data was preprocessed and analyzed in R. ANOVAs were run with the *afex* package using type three sums of squares. Significant interactions were followed up by simple main effect analysis using the *emmeans* package. For all inference statistical analysis alpha level was set to .05.

---

[5] The four possible block orders were A) [Individual Action | Individual Imitation], [Individual Action | Joint Imitation], [Joint Action | Joint Imitation], [Joint Action | Individual Imitation], B) [Individual Action | Joint Imitation], [Joint Action | Individual Imitation], [Individual Action | Individual Imitation], [Joint Action | Joint Imitation], C) [Joint Action | Individual Imitation], [Joint Action | Joint Imitation], [Individual Action | Joint Imitation], [Individual Action | Individual Imitation], D) [Joint Action | Joint Imitation], [Individual Action | Individual Imitation], [Joint Action | Individual Imitation], [Individual Action | Joint Imitation].

2.2.5. Transparency and openness

Data collection happened in March 2025. Raw data and analysis scripts are accessible in an Open Science Framework project repository accessible under https://osf.io/jpxqr/?view_only=e4594c0947174580a65418b05ccf72df. Directed hypotheses, study design, sampling and analysis plan were preregistered in the Open Science Framework. The preregistration can be accessed online under https://osf.io/sf8ru/?view_only=bdd178c8ce3841299064a1927c8f5c0b. Any deviations from the preregistration are mentioned explicitly and marked as explorative analysis.

2.3. Results

2.3.1. Pre-registered analysis

The results of the preregistered analysis are displayed in Figure 2-5.
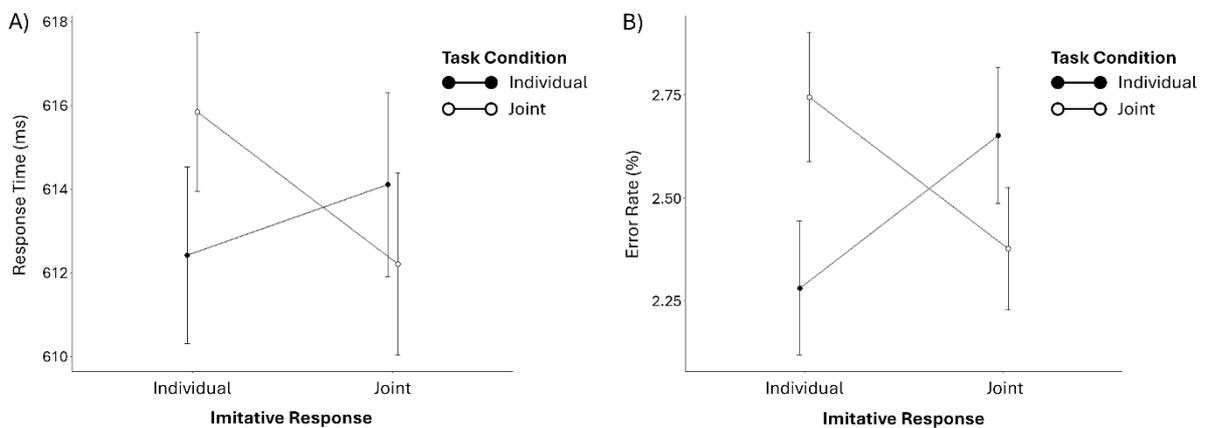
Figure 2-5. A) Mean response time and B) mean error rate across participants as a function of Task Condition (Individual Action vs. Joint Action) and Imitative Response (Individual Imitation vs. Joint Imitation). Error bars represent standard error of the mean corrected for within-subjects designs.

The RT analysis revealed no statistically significant effects. The main effects of Task Condition and Imitative Response were non-significant (both $F < 1$). Crucially, the predicted interaction between Task Condition and Imitative Response also did not reach significance, $F(1, 234) = 1.43$, p = .23, $\eta^2 = .006$. At least descriptively though, participants responded

33

slightly faster in Joint Action trials when imitated by both members of the opposing dyad (Joint Action | Joint Imitation: M = 612 ms, 95% CI [596, 628]) compared to when imitated by one member (Joint Action| Individual Imitation: M = 616 ms, 95% CI [599, 632]), while the reverse trend was observed in Individual Action trials (Individual Action | Joint Imitation: M = 614 ms, 95% CI [597, 631] vs. Individual Action | Individual Imitation: M = 612 ms, 95% CI [596, 629]).

The ER analysis showed no significant main effects of Task Condition or Imitative Response (both $F < 1$). Crucially though, the interaction between both factors reached significance, $F(1,234) = 4.68$, $p = .032$, $\eta_p^2 = .020$. Examining Panel B in Figure 2-5, the interaction was driven by fewer response errors in Joint Action trials when participants were imitated by both members of the opposing dyad (M = 2.38 %, 95 % CI [1.97, 2.78]) compared to when being imitated by a single member (M = 2.74 %, 95 % CI [2.33, 3.15]), with the opposite pattern observed for Individual Action trials (Individual Action | Joint Imitation: M = 2.65 %, 95 % CI [2.20, 3.10] vs. Individual Action | Individual Imitation: M = 2.28%, 95 % CI [1.87, 2.69]). Yet, follow-up tests of simple main effects did not reach significance: Main effect of Imitative Response on Joint Action trials, $t(234) = -1.56$, $p_{corrected} = .24$; main effect of Imitative Response on Individual Action trials, $t(234) = 1.70$, $p_{corrected} = .18$.

2.3.2. Explorative analysis

Our preregistered analysis showed a performance pattern that partially aligns with the predictions of the group-level hypothesis. While a group-level congruency effect was evident for participants' error rates, descriptive differences in response times — although in line with the group-level hypothesis — did not reach statistical significance. Based on these initial findings, we decided to run a follow-up analysis, exploring possible moderators of performance differences between our experimental conditions. To this end, we first combined participants RTs and ERs into inverse efficiency scores (IES = RT / [1 - ER]; Townsend & Ashby, 1978) to

get a combined measure of participants' task performance[6]. We then performed a median split of our sample based on participants' overall mean response time in the task. This allowed us to assess whether participants overall response speed could have moderated the group-level congruency effect indicated by our preregistered analysis. Analysing overall response speed as a possible moderator was motivated by the idea that participants who performed the task relatively fast may have selected their responses primarily in a stimulus-based rather than an effect-based action control mode (c.f., Waszak et al., 2005), i.e., by selecting their responses merely in compliance with the instructed stimulus-response mapping but without necessarily anticipating their distal effects in the environment (e.g., imitative responses performed by the avatar hands). As previous research findings have shown that compatibility effects between actions and their resulting consequences require an effect-based action control mode (Pfister et al., 2010; Zwosta et al., 2013a), possible interactions between Task Condition and Imitative Response could thus have been overshadowed by participants who performed the task in a stimulus-based action control mode instead.

To test this post-hoc hypothesis, we thus analysed participants' IES by means of a 2 (Task Condition: Individual Action vs. Joint Action) x 2 (Imitative Response: Individual Imitation vs. Joint Imitation) x 2 (RT Group: Fast Responder vs. Slow Responder) mixed ANOVA with Task

---

[6] Analyzing IES as a single measure of task performance is justified in the current study as error rates were relatively low while the converging patterns for error rates and response times provide no indications for a speed-accuracy trade off (Bruyer & Brysbeart, 2011).

Condition and Imitative Response as a within- and RT Group as a between-subjects factor. The results are depicted in Figure 2-6.
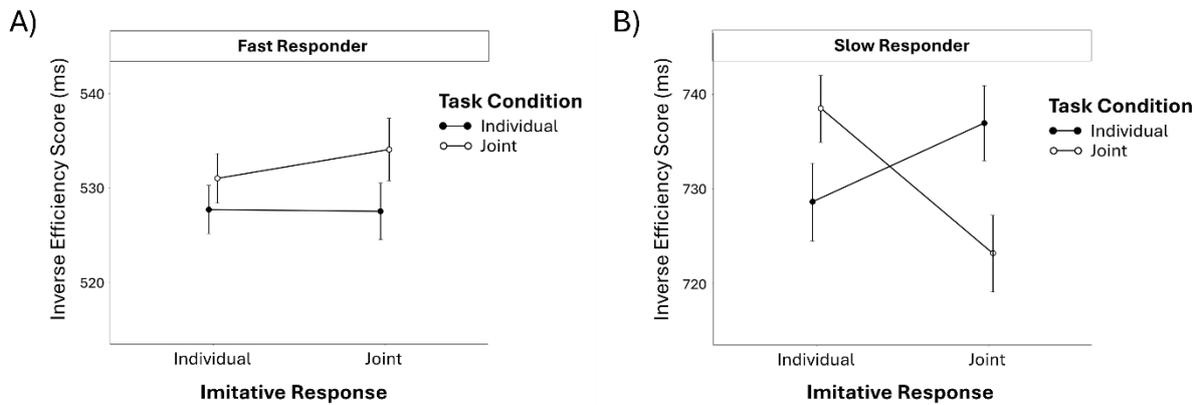
Figure 2-6. Inverse efficiency scores as a function of Task Condition and Imitative Response plotted separately for A) fast responder and B) slow responder. Error bars represent standard error of the mean corrected for within-subjects designs.

Not surprisingly, the analysis showed a main effect of RT Group, $F(1,233) = 323.5$, $p < .001$, $\eta_p^2 = .581$. The two-way interaction between Task Condition and Imitative Response approached but did not reach statistical significance, $F(1,233) = 3.81$, $p = .052$, $\eta_p^2 = .016$. Crucially though, the analysis showed that this interaction was moderated by participants' overall response speed, as indicated by a significant three-way interaction between Task Condition, Imitative Response and RT Group, $F(1,233) = 6.61$, $p = .011$, $\eta_p^2 = .028$. No other effects were significant (see Appendix A for the detailed result table). To follow-up on the three-way interaction, we analysed IES in each RT Group separately by means of two separate 2 (Task Condition) x 2 (Imitative Response) repeated-measures ANOVAs. Main effects of Task Condition and Imitative Response were non-significant in both analysis (see Appendix A for the detailed result tables). Yet, contrasting both analysis showed that the two-way interaction between Task Condition and Imitative response was significant in the slow responder group, $F(1,116) = 7.21$, $p_{\text{corrected}} = .016$, $\eta_p^2 = .059$, while it was non-significant in the fast responder group ($F < 1$). Following up on the interaction between Task Condition and Imitative Response in the slow responder group showed that the interaction was driven by lower IES (i.e., better

36

performance) in Joint Action trials when participants were imitated by both vs. one member(s) of the opposing dyad (Joint Action | Joint Imitation: $M = 723$ ms, 95 % CI [704, 743] vs. Joint Action | Individual Imitation: $M = 738$ ms, 95 % CI [718, 759]; $t(116) = 2.77$, $p_{corrected} = .013$), while the opposite trend emerged on Individual Action trials (Individual Action | Joint Imitation: $M = 737$ ms, 95 % CI [715, 759] vs. Individual Action | Individual Imitation: $M = 729$ ms, 95 % CI [707, 750]; $t(116) = -1.39$, $p_{corrected} = .33$).

2.4. Discussion

The present study assessed how being imitated by one (individual imitation) or two member(s) of an opposing dyad (joint imitation) affects to-be-imitated responses that are performed without (individual action) or together with another co-actor (joint action). Under the assumption that anticipated actions of other agents get employed to initiate and control one's own contributions to a social interaction (c.f., Kunde et al., 2018), the current study allowed us to test two opposing hypotheses about the structure and content of action representations guiding action performance in joint task settings. If actions performed together with others (i.e., joint action) are guided by task representations encoding group- rather than individual level performance (group-level hypothesis), task performance in the current study should be affected by congruency relations between the interacting groups: Joint action should be facilitated by anticipating joint vs. individual imitation, while individual action should be facilitated by anticipating individual vs. joint imitation. In contrast, if actions performed together with others are guided by task representations that remain centred on individual-level performance (individual-level hypothesis), task performance in the current study should be affected by congruency relations between individual agents: Task performance should increase the more agents mimic participants' individual responses.

Providing partial support for the group-level hypothesis, our findings showed that imitative responses of two vs. one member(s) of another dyad affected the accuracy of to-be-

37

imitated responses in opposite ways, depending on them being performed with or without another co-actor. In other words, we found evidence that participants' response accuracy was driven by congruency relations between groups rather than individuals: Not the total number of anticipated responses by all individual agents in the scene, but congruency between the number of actions performed by participants' own group (i.e., "us") and the number of anticipated responses performed by the opposing group (i.e., "them") affected how accurate participants' performed the task. However, limiting the conclusiveness of our preregistered analysis, evidence for a group-level congruency effect was restricted to analysis of participants response accuracy, while analysis of participants response times — although showing a trend in the same direction — revealed no statistically significant effects.

Interestingly though, results of our explorative analysis suggest that significant performance differences between our experimental conditions were partially obscured by a subset of participants who responded relatively fast on the task compared to others. Our explorative analysis showed that the group-level congruency effect on task performance was moderated by participants' overall response speed: Those participants who responded relatively slow on the task but not those who responded relatively fast were affected by congruency relations between groups and showed opposing effects of being imitated by both vs. one member of the opposing dyad depending on whether they performed the task with or without their co-actor. In contrast, participants who responded relatively fast on the task showed no indications of being affected by imitative responses of other agents in the scene.

A possible explanation for this limiting factor in our results could be that the fast responding participants performed the task primarily in a stimulus-based rather than an effect-based action control mode, i.e., by selecting their responses simply in compliance with the instructed stimulus-response mapping but not in order to trigger imitative responses in the other agents (c.f., Herwig et al., 2007; Waszak et al., 2005). Previous research has shown that response selection is generally faster in stimulus- compared to effect-based action control

38

modes (e.g., Keller et al., 2006; Waszak et al., 2005) and that congruency effects between actions and their outcomes depend on task instructions that foster effect-based rather than stimulus-based action control modes (Pfister et al., 2010; Zwosta et al., 2013a). Indeed, task instructions in the current study focussed on participants' primary task to lift their index or middle finger in response to shape stimuli appearing on screen but made it less explicit that participants should do so in order to trigger an imitative response in the other agents. This might have led many participants to zoom in on their primary task to select responses in reaction to the task-relevant stimuli, without integrating the responses of the other agents in the scene in to their own representation of the task.

The findings of the present study bear several indications. First, our findings inform research investigating how the behavioural changes we evoke in other people, i.e., the *social* consequences of our actions, get integrated into our own action planning and control processes. Previous research on so-called *sociomotor action control* (Kunde et al., 2018) has shown that people access their own contributions to a social interaction by anticipating the behavioural responses they evoke in their interaction partners. Whereas demonstrations of this principle remained limited to dyadic interactions so far (for exemptions see Galang et al., 2024; Neszmélyi & Pfister, 2024), our findings indicate that people can use anticipated actions of *multiple* other agents to initiate and control their own contributions to a social interaction.

However, our findings also point to limitations in how readily people integrate behavioural responses evoked in others into their own action planning and control processes. The finding that only some but not all participants were affected by imitative responses of the other agents in the scene indicates that integrating predictable actions of others into representations of one's own actions is not an automatic process but seems to be dependent situational factors that may comprise people's current task construal or their dominant mode of action control. Future research on action control in social interaction should thus assess the conditions and enabling factors that make people more or less pronounced to represent their

contributions to a social interaction in terms of the responses they evoke in others in a more systemic way (see Kunde et al., 2018 for a summary of initial findings).

Second and central to the aim of the present study, our findings inform research on the content and structure of joint task representations. The finding that participants' task performance was driven by congruency relations between groups indicates that participants formed task representations encoding group-level performance (i.e., "*WE* both lift our index finger") rather than individual level performance alone (i.e., "*I* lift my index finger"). As such, anticipating an imitative response by both members of the other group (i.e., "*THEY* will both lift their index finger") could prime access to the joint task representation of lifting one's finger together with one's co-actor, while anticipating an imitative response by only one member of the other group (i.e., "one of them will lift their index finger") could interfere with the joint task representation of lifting one's finger together with one's co-actor.

This finding extends previous research on imitation between groups that has led to similar theoretical conclusions (Essa et al., 2019; Ramenzoni et al., 2014; Tsai et al., 2011; see McEllin et al., 2018 and Milward & Sebanz, 2016 for discussion). While these studies showed that *external observation* of individual and joint actions can prime the execution of individual and joint actions respectively, our study goes one step further by showing that the *internal anticipation* of individual or joint action in others can produce similar effects. As such, our findings tap into the endogenous action planning processes that guide the initiation and control of joint actions in a goal-directed rather than stimulus-driven way, providing direct support for the idea that task representations encoding group-level performance (i.e., "we-representations") can play a role in joint action planning (Formica & Brass, 2024; Kourtis et al., 2019).

In summary, our study provides initial evidence that anticipated imitation effects can be modulated by congruency relations between interacting groups. This finding indicates that people can use anticipated actions of multiple other agents to initiate and control their own contributions to a social interaction and supports the idea that actions performed together with

others become represented in relation to group-level rather than individual-level performance. Limiting our findings, it remains to be determined why not all participants in the current study showed to be affected by predictable responses of other agents in the scene. Future research should explore task conditions that may foster the integration of others' anticipated actions into people's own action representations. For instance, task instructions that emphasize action-outcome contingencies over stimulus-response mappings in the present setup may enhance sensitivity to the imitative behaviour of others. In addition, the presence of physical rather than virtual co-actors could further promote their integration into people's action representations.

3. STUDY II:

Action-Outcome Learning in Synchronous Joint Action

3.1. Introduction

Joint actions, such as playing a piano duet, moving furniture together or greeting a friend with a fist bump, require people to purposefully coordinate their actions in the service of shared goals (Butterfill, 2018; Pacherie, 2013; Sebanz et al., 2006). Previous research indicates that successful performance of joint actions directed at shared goals hinges on dedicated mental representations specifying the desired outcomes of the joint action and the joint contributions of self and others that are needed to achieve them (Knoblich et al., 2011; Vesper et al., 2010a; see Sebanz & Knoblich, 2021 for a recent review). Yet, little is known about how people acquire these representations and what learning mechanisms underlie their formation. To fill this gap, the current study investigated how co-actors learn and represent novel instrumental relationships between their joint actions and the outcomes they produce together.

3.1.1. Action-outcome learning

Producing goal-directed actions, either alone or together with others, presupposes knowledge about instrumental relationships between actions and their resulting outcomes on the body and the environment (Dickinson & Balleine, 1994; Hommel, 2017; Wolpert & Flanagan, 2001). A parsimonious way to explain how people acquire such knowledge implicitly through sensorimotor experience is provided by associative accounts of goal-directed behaviour (Wit & Dickinson, 2009), most prominently by ideomotor theories (Greenwald, 1970; Hommel et al., 2001; Prinz, 1997; see Pfister, 2019 and Shin et al., 2010 for reviews). Ideomotor theories propose that the capacity to produce goal-directed behaviour relies on bidirectional action-outcome associations that are acquired through the recurrent experience of contingent

relationships between actions and their outcomes. Once established, bidirectional action-outcome associations provide a simple mechanistic explanation of how actions become causally initiated and controlled: By activating perceptual representations of action outcomes, either through internal mental anticipation or through external perception in the environment, their associated motor programs become directly activated and prepared for execution. Thus, ideomotor theories propose that actions are initiated and controlled through representations of their associated outcomes.

Evidence that bidirectional action-outcome associations underlay learning and control of individual motor behaviour comes from studies employing a well-established two-stage action-outcome learning task (cf. Elsner & Hommel, 2001). The task involves an initial learning phase, in which people experience novel action-outcome relationships by performing simple actions (e.g., left- and right-handed button presses) that contingently produce arbitrary outcomes in the environment (e.g., low and high tones). Whether people acquired bidirectional action-outcome associations during the learning phase is then probed in a subsequent test phase, in which the former action outcomes are presented as imperative stimuli, requiring participants to perform either free- or forced-choice responses. Numerous studies have demonstrated that the presentation of former action outcomes in the test phase primes execution of those responses that produced these outcomes in the preceding learning phase (Elsner & Hommel, 2001, 2004; Herwig et al., 2007; Hoffmann et al., 2009; Janczyk et al., 2022; Pfister et al., 2011; Sun et al., 2020; Watson et al., 2015). These findings support the idea that the recurrent experience of contingent relationships between actions and their outcomes leads to the acquisition of bidirectional action-outcome associations.

Bidirectional action-outcome associations thus provide a mechanistic explanation of how people acquire mental representations that enable the production of individual goal-directed actions. In the current study, we investigated how action-outcome learning takes place during

joint actions and could lead to the acquisition of dedicated *joint* action representations guiding interpersonal action coordination in the service of shared goals.

3.1.2. Representing joint actions

Many joint actions involve deliberate coordination between actions of multiple individuals that are collectively directed at the production of joint outcomes in the environment (Butterfill, 2012; Pacherie, 2013; Sebanz et al., 2006). Thus, in contrast to individual actions, joint actions seem to require that individuals learn and represent how their own actions *together* with those of their co-actors produce outcomes in the environment that reflect the *joint* effects of their coordinated action contributions.

These representational peculiarities of joint action are underlined by a range of empirical findings. By investigating how pairs of pianists monitor individual and joint performance during musical duets, Loehr and colleagues (2013) showed that co-actors are especially sensitive to performance errors altering outcomes that depend on their common contributions to the joint action (e.g., the harmony of jointly produced chords), while individual performance errors that leave these joint action outcomes unaffected are processed with lower priority. Furthermore, Loehr and Vesper (2016) showed that when people learn to play simple piano melodies as part of a musical duet performed with an accompanist, it is more difficult for them to produce the learned melody alone without the accompanying auditory feedback of their counterpart. Together, these studies indicate that actions performed in joint action contexts are predominantly guided by representations of joint rather than individual action outcomes.

This evokes the question of how co-actors represent instrumental relationships between their actions and the joint outcomes they produce together. One theoretical possibility proposed by Vesper et al. (2010) is that people engaging in joint actions rely on *minimal joint action representations* that merely specify the outcome of the joint action and an individual's own action contributions that are needed to achieve it, while only being aware that the outcome

44

cannot be achieved by acting alone — captured by the formula "*ME + X*". Thus, according to this proposal, individuals engaged in joint actions may link the production of joint action outcomes only to their own contributions to the joint action while their co-actors' contributions would not need to be specified in any detail.

While these minimal representational requirements capture a range of scenarios that qualify as genuine cases of goal-directed joint action (c.f., Knoblich et al., 2011; Vesper et al., 2010a), many empirical findings indicate that people often form more elaborate representations of their joint actions that also specify their co-actors' contributions to the joint action in more detail. Studies on shared task representations show that people tend to co-represent specific aspects of others' tasks and actions when acting alongside each other, which is indexed by modulations of people's individual task performance by specifics of their co-actors' tasks and actions. For example, people's individual task performance has been shown to be modulated by simple stimulus-response rules (Atmaca et al., 2008; Sebanz et al., 2003, 2005) and action-outcome mappings (Pfister et al., 2014; Sacheli et al., 2018, 2021) of their co-actors as well as by more elaborated aspects of others' tasks, such as their physical task constrains (Schmitz et al., 2017) or the order of their actions (Schmitz et al., 2018).

These findings imply that joint action representations can encode not only an individual's own contributions to the joint action but also those of their co-actors. Yet, the assumption of shared task representations leaves open the question of how co-actors integrate information about their own and their partners' action contributions into unified representations of their joint action performance (Butterfill, 2015; Keller et al., 2016; Knoblich & Jordan, 2003; Pesquita et al., 2018; Sebanz & Knoblich, 2009; Sinigaglia & Butterfill, 2022).

An interesting possibility is that co-actors form joint action representations that specify foremost what they are pursuing together as a group (i.e., as a "*WE*") rather than as separate interacting individuals (i.e., as "*ME+YOU*") (Butterfill, 2015; della Gatta et al., 2017; Gallotti & Frith, 2013; Kourtis et al., 2019; Pacherie, 2013; Tsai et al., 2011). Thus, instead of specifying

the individual action contributions of self and others separately and in parallel, joint action representations may primarily encode how actions are to be performed by the group as a whole.

Ample empirical support for the idea that co-actors form group-level representations of their joint actions is now provided by several lines of research. Studies investigating rationality principles of higher-level action planning in joint action contexts have shown that co-actors raise individual effort to maximize action efficiency at the level of the group (Török et al., 2019, 2020). Studies investigating how co-actors' experience agency in joint action contexts have demonstrated that people's judgments of control over joint actions are strongly affected by group-level task performance (Dewey et al., 2014; Loehr, 2018) and seem to reflect a sense of joint rather than self-agency (Bolt et al., 2016; Bolt & Loehr, 2017).

Further evidence for group-level action representations has been provided by studies on action mimicry. These studies showed that actions performed in synchrony or in turns with a partner are facilitated when co-actors observe the same actions performed by another dyad compared to when observing only individual parts of the joint action performed by a single actor (Ramenzoni et al., 2014; Tsai et al., 2011). Thus, performing actions as part of a joint action benefits more from observing another group modeling the joint action than from observing a single actor modeling only individual parts of it.

Lastly, studies on interpersonal coordination indicate that joint action representations can specify not only each co-actor's separate contributions to a joint action, but also how the individual contributions of self and others relate to each other at the level of the group. This is supported by findings showing that interference effects between observed and executed actions of two co-actors become modulated when both actions are performed as interrelated contributions towards a shared goal (Clarke et al., 2019; della Gatta et al., 2017; Sacheli et al., 2018). Furthermore, a recent study by Kourtis et al. (2019) showed that co-actors' action initiation and coordination performance benefits from prior information about pending joint

46

actions that merely specifies relations between co-actors' upcoming individual action contributions (e.g., whether co-actors will perform similar or different actions).

Taken together, the reviewed evidence for group-level action representations opens up the possibility that co-actors may link the production of joint action outcomes directly to group-level relations between their individual contributions to the joint action.

3.1.3. The present study

The purpose of the present study was to examine how people acquire joint action representations by investigating how co-actors come to represent novel instrumental relationships between their coordinated actions and the joint outcomes they produce together. To that end, we assessed how the recurrent experience of contingent relationships between joint actions and their resulting outcomes affects action-outcome learning in jointly acting individuals.

Based on the previous literature reviewed above, we contrasted two theoretical alternatives of how co-actors might link their coordinated actions to the joint outcomes they produce together. A first possibility is that action-outcome learning during joint action is merely sensitive to an individual's own contributions to the joint action. Thus, individuals engaged in a joint action may link the production of joint action outcomes merely to their own action contributions, but not to the contributions of their co-actors. This would indicate that action-outcome learning during joint action leads to the implicit formation of *minimal joint action representations*, merely specifying the joint action outcome and an individual's own contribution to it.

Alternatively, action-outcome learning during joint action may instead be sensitive to group-level relations between co-actors' individual contributions to the joint action. According to this possibility, individuals engaged in a joint action may link the production of joint action outcomes directly to spatial and/or temporal relations between their own and their co-actors'

47

contributions to the joint action that emerge at the level of their group-level performance. This would indicate that action-outcome learning during joint action leads to the implicit formation of *group-level representations* of the joint action. See Figure 3-1 for a visual illustration of the two theoretical alternatives.
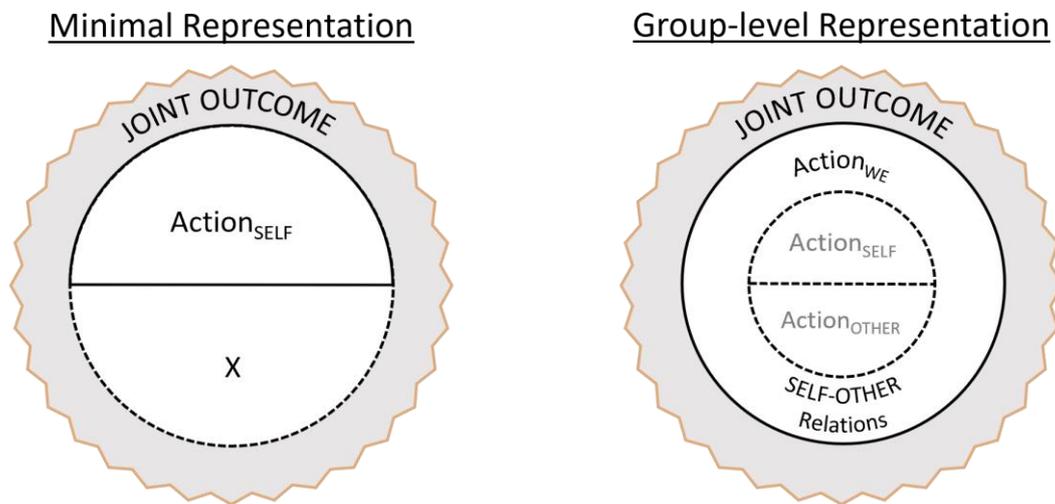
Figure 3-1. Illustration of two possible representational structures that may result from action-outcome learning in joint action contexts (see main text for explanation).

To contrast these alternatives, we set out to investigate action-outcome learning in *synchronous joint actions* that require multiple co-actors to synchronize their individual action contributions with one another to achieve a desired outcome in the environment. Synchronous joint actions provide a specifically interesting test case for action-outcome learning as they introduce contrasting predictions of how jointly acting individuals could represent instrumental relationships between their actions and perceived outcomes: When co-actors perform synchronized actions to produce a joint outcome (e.g., performing synchronized key strokes on a piano to produce a harmonic chord), each co-actor could attribute the perceived outcome to individual-level features of their own action contribution alone (e.g., *I* pressed a certain key with my left/right index finger), but also to relational group-level features of their own and their partners' action contributions taken together (e.g., *WE* pressed a certain configuration of keys with similar/different fingers). As such, synchronous joint actions should allow us to test

whether action-outcome learning in joint action contexts leads to the formation of minimal or group-level representations of the joint action.

Therefore, we adapted the individual action-outcome learning task by Elsner and Hommel (2001) to a synchronous joint action setting. In an initial learning phase, two co-actors produced a series of low and high two-tone chords by means of synchronized key presses on a joint response key layout with four horizontally aligned keys. In a subsequent test phase, both co-actors were instructed to respond to the former chord outcomes with respect to a stimulus-response mapping that was manipulated to either preserve or reverse the action-outcome relationships of the previous learning phase, with respect to both individual-level features of each co-actor's isolated responses *and* to relational group-level features of co-actors joint response configurations (see Figure 3-2). According to the minimal account, action-outcome learning should be only sensitive to individual-level features of co-actors' isolated response contributions. Following associative accounts of action-outcome learning, this should be reflected in a performance advantage in the test phase when the instructed stimulus-response mapping preserves the action-outcome relationship of the preceding learning phase on the level of each individual co-actor. In contrast, according to the group-level account, action-outcome learning should be sensitive to relational features of co-actors' joint response configurations. Following associative accounts of action-outcome learning, this should be reflected in a performance advantage in the test phase when the instructed stimulus-response mapping preserves the action-outcome relationship of the preceding learning phase on the level of the group.

49
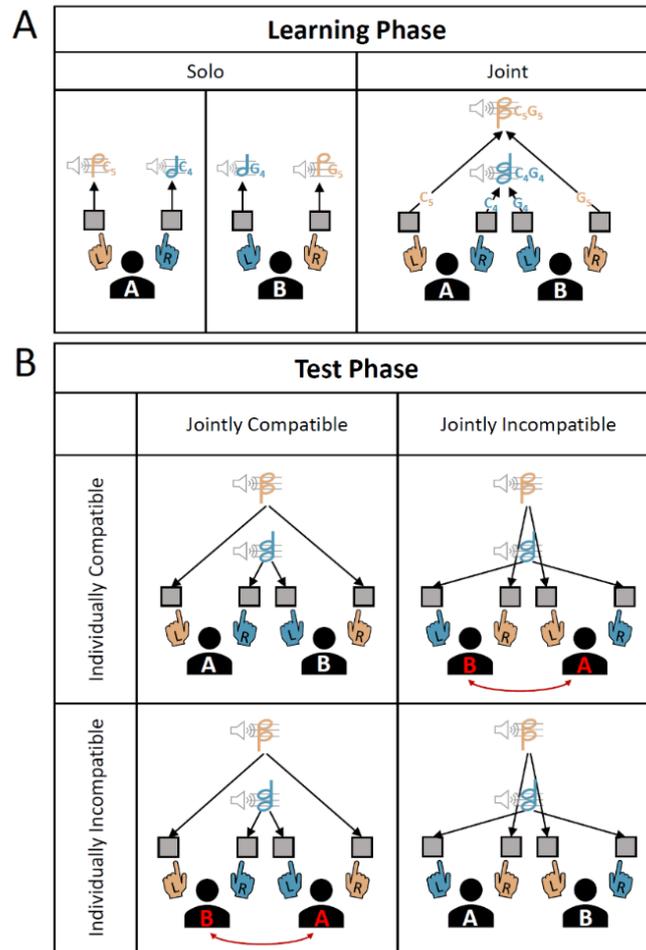
Figure 3-2. Experimental set-up and design. Panel A shows the learning phase. In the solo part of the learning phase (left), individual participants produced low and high tones by pressing the left (L) or right (R) button on their individual response box. In the joint part of the learning phase (right), pairs of participants performed synchronous responses to produce low- and high-pitched two-tone chords together. At the group level, co-actors produced the two chords by performing synchronized responses to the inner (marked in blue) or outer buttons (marked in orange) of their joint response button layout. At the individual level, each co-actor contributed to the chord outcomes by performing a response to the left or right button on their individual response box. Thus, the jointly produced chords could be represented as being contingent on the relational response configurations co-actors performed together (i.e., WE press the inner/outer buttons) or as being related to an individual participant's own response contribution (I press my left/right button) or both. Panel B shows the test phase. Pairs of participants responded in parallel to the previously produced chords according to a prescribed stimulus-response mapping. On the group level, the stimulus-response mapping was manipulated to either preserve (Jointly Compatible mapping) or reverse (Jointly Incompatible mapping) the action-outcome mapping of the previous learning phase with respect to co-actors' joint response configurations (inner/outer). The stimulus-response mapping was also manipulated at the individual level, so that it either preserved (Individually Compatible mapping) or reversed (Individually Incompatible mapping) the action-outcome mapping of the previous learning phase regarding each co-actor's individual response contributions (left/right). To manipulate joint and individual compatibility of the stimulus-response mappings orthogonally to each other, co-actors were instructed to switch their seating positions from learning to test phase in two of the four test phase conditions (indicated by the red arrows).

50

3.2. Methods

3.2.1. Participants

In total, eighty adult participants took part in the experiment, grouped into pairs that were randomly composed upon study sign-up. Four pairs were dropped from analysis meeting preregistered exclusion criteria so that the final sample included seventy-two participants (25 male, 47 female, $M_{Age} = 25.8$, $SD_{Age} = 4.5$) grouped into thirty-six pairs (17 same gender pairs, 19 mixed gender pairs)[7]. Recruitment took place via the online research participation system of Central European University (CEU) in Vienna, Austria. All participants gave written informed consent and were compensated with 10 Euro for participation. Ethical approval for the study was granted by the Psychological Research Ethics Board of CEU.

3.2.2. Apparatus and stimuli

The experiment was run in a quiet, well-lit room. Stimulus presentation and response recording was controlled by a custom-made script written in PsychoPy (Peirce et al., 2019), running on a Dell computer attached to a 24 inch LCD monitor with a refresh rate of 60 Hz. Each participant responded using an individual response box (The Black Box ToolKit; dimensions: 202 mm x 137 mm x 35 mm LWH) with four horizontally aligned response buttons. Only the two outer buttons of the response boxes were used for the experiment. If not specified differently, written instructions and text stimuli were presented in white letters against

---

[7] Previous studies deploying variants of our task in the domain of individual action observed medium to large effect sizes for individual compatibility manipulations in the test phase (Eder & Dignath, 2017; Elsner & Hommel, 2001, Exp. 1; Hoffmann et al., 2009, Exp. 1; Hommel et al., 2003; Wolfensteller & Ruge, 2011). A sensitivity analysis in G*Power (Faul et al., 2007) showed that the final sample size of $N = 72$ would have been sufficient to detect main effects of size $d = .67$ with 80% power at an alpha level of .05 (two-sided) in a 2x2 between-subjects factorial design. Thus, our final design and sample size should have been sufficiently powered to detect and replicate an individual compatibility effect, if the effect persists in joint action contexts. We take this as a reasonable starting point for justifying our sample size decision, as our preregistered predictions were targeted at main effects in our final 2x2 between-subjects factorial design.

black background on the computer monitor at a viewing distance of approximately 60 cm. Auditory stimuli comprised four synthesized organ notes differing in pitch (C_4, G_4, C_5, G_5)[8] and were presented via stereo speakers (Genius SP-HF 180) placed to the left and right of the computer monitor at a volume of approximately 60 dB.

3.2.3. Procedure

The experiment was divided into a learning and a test phase and lasted about 60 minutes. The experimenter was present throughout the whole session and monitored the procedure from outside the participants' view.

Learning phase

The learning phase comprised two parts. The first part was performed by each pair member alone while the other waited outside the laboratory room (*solo part*). The second part was performed by both pair members together (*joint part*).

Solo part. For the solo part of the learning phase, participants were seated centrally at the long side of a table facing the computer monitor with a single response box (labelled with "A" or "B" respectively) placed in front of them. One pair member received the "A"-labelled box (referred to as Participant A), the other the "B"-labelled box (referred to as Participant B). Participants were instructed to produce a series of high and low tones by pressing the left and the right button on their response box with the index finger of their left and right hand, respectively. The mapping between responses and tones was instructed by an illustration of the response box highlighting the respective buttons and labelling them with "high" and "low" respectively (see Figure 3-3, Panel A). Within pairs, participants always received the reversed mapping compared to their partner (e.g., Participant A: left → high tone, right → low tone;

---

[8] Organ tones were chosen due to their sharp attack and decay. All tones were synthesized with the digital audio workstation software LMMS (https://lmms.io/) using open-source sound fonts.

Participant B: left → low tone, right → high tone; c.f. Figure 3-2). The mappings were counterbalanced across pairs. One pair member produced C-notes (low tone: C_4, high tone: C_5) while the other pair member produced G-notes (low tone: G_4, high tone: G_5). Thus, for both pair members the pitch difference between low and high tones was one octave.

At the beginning of each trial, participants were instructed to decide which tone to produce next by pressing the assigned button on their response box. Participants could choose freely but were instructed to produce a balanced amount of high and low tones throughout the experimental phase. After indicating their decision, participants' choice was centrally displayed as a text prompt (e.g., "high tone") for 1000 ms. Then, a counter appeared on screen, counting in an interval of 500 ms from three down to a "GO!" prompt that remained on screen for 500 ms. Participants were instructed to issue their respective response in synchrony with the onset of the "GO!" prompt. If participants responded in line with their indicated decision and if their response fell within a response window of 500 ms around the onset of the "GO!" prompt, the respective tone was played for a duration of 600 ms. If participants' responses fell outside an additional asynchrony window of 250 ms around the onset of the "GO!" prompt, a feedback message reminding them to respond in synchrony with the onset of the "GO!" prompt that was
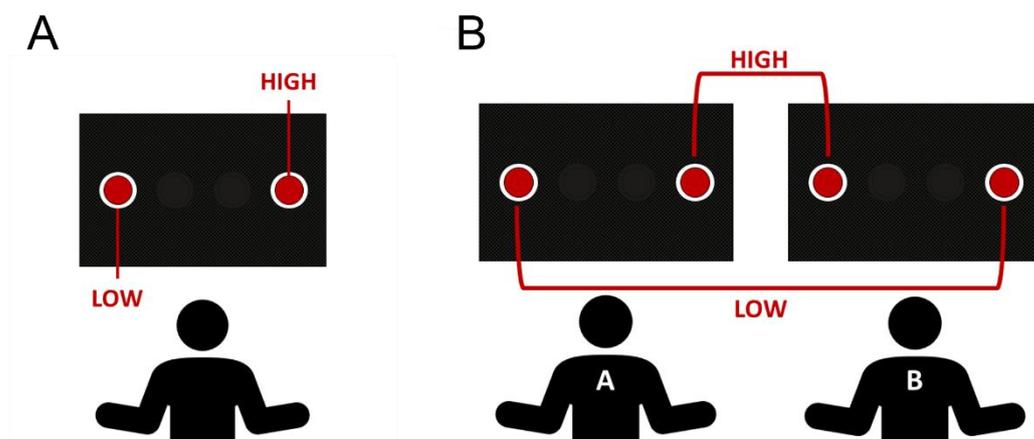
Figure 3-3. Illustrations used to instruct participants about their action-outcome mappings in the solo part of the acquisition phase (A) and in the joint part of the acquisition phase (B). In the test phase, stimulus-response mappings were instructed with a picture corresponding to panel B.

presented on screen after the tone had played. If participants responded not in line with their indicated decision or gave no response within the response window, no tone was played, and an error message was displayed in the centre of the screen for 1500 ms. Throughout each trial, a visual illustration of participants' response box was displayed at the bottom of the computer screen, mirroring participants' responses. As soon as participants produced a response within the response window, the respective response button in the display was highlighted until the end of the trial. The next trial started after a blank screen, displayed for 500 ms. Break messages appeared after every ten successful trials (trials in which participants produced a tone) and informed participants about the ratio of high and low tones produced so far. The solo part of the learning phase ended after each pair member had performed forty successful trials.

Joint part. For the joint part of the learning phase, the two pair members were seated side by side at the long side of the table, both facing the centrally placed computer monitor. Participant A sat on the left side, Participant B sat on the right. Participants' individual response boxes were placed between markers on the table in front of them, 20 cm apart from each other. They were instructed to produce high and low two-tone chords together by each pressing one of their response buttons in synchrony with their partner. The target chords were two organ chords made up from one C- and one G-note and differed in their relative pitch by means of one octave (lower chord: C_4G_4, higher chord: C_5G_5). The mappings between participants' individual responses and tone outcomes remained the same as in the solo part of the learning phase, so that one pair member was responsible for playing the C- and the other was responsible for playing the G-note of a chord. Consequently, producing the target chords required pair members to produce different spatial response configurations together: Pairs jointly responded to the inner two buttons (requiring Participant A to press their right and Participant B to press their left button) or the outer buttons (requiring Participant A to press their left and Participant B to press their right button) of their joint response button layout (see Figure 3-2). The mappings

54

between participants' responses and chord outcomes were instructed by an illustration, showing the response boxes of both pair members side by side, highlighting the relevant response buttons, and labelling them with "low" and "high" respectively (see Figure 3-3, Panel B). The rationale behind using this instruction was to leave it to the participants whether to code their responses at the group level (inner/outer) or at an individual level (left/right).

The trial timeline for the joint part of the learning phase is illustrated in Figure 3-4, Panel A. At the beginning of each trial, pairs were instructed to jointly decide which chord to produce next. Therefore, one pair member was randomly assigned to verbally propose the next chord to
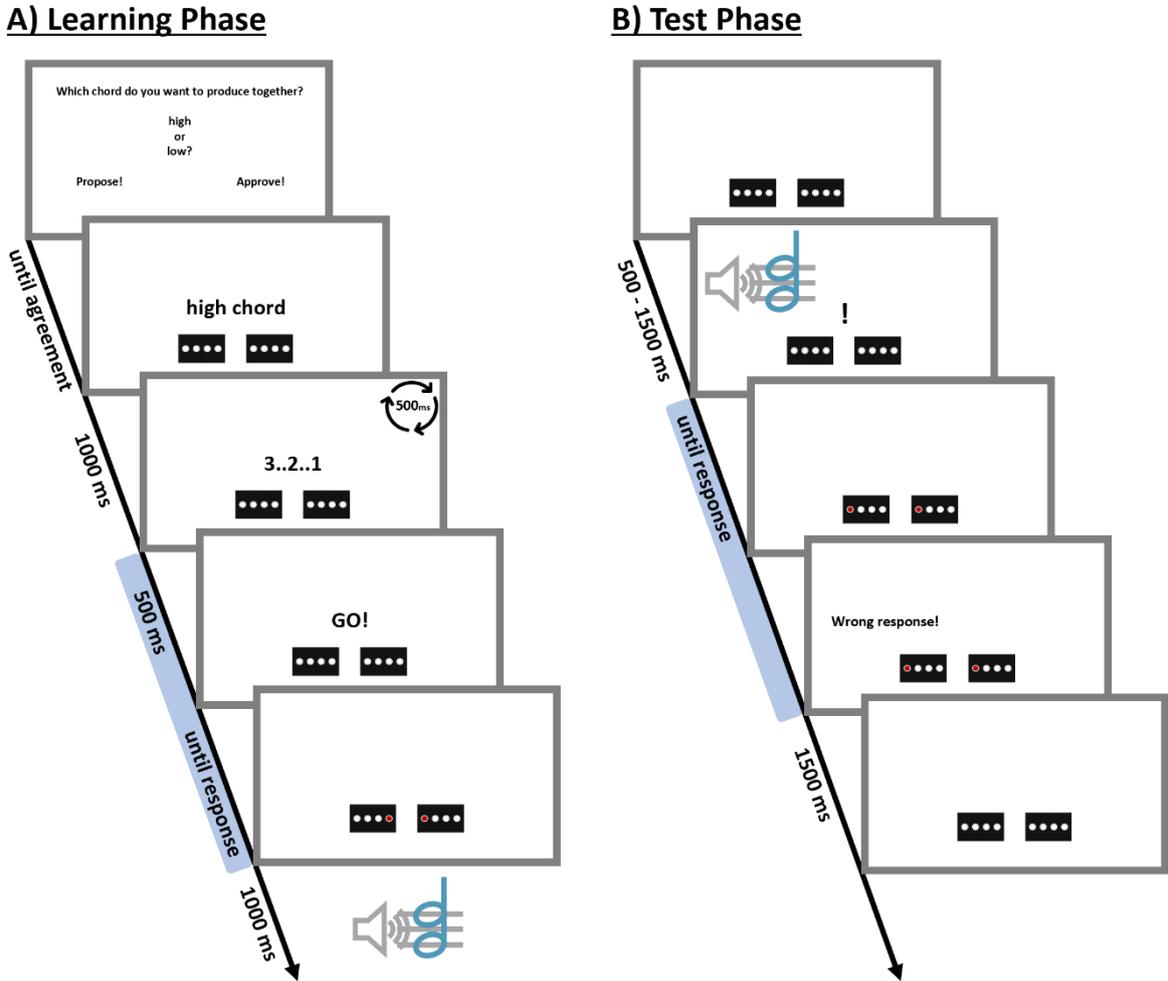
Figure 3-4. Illustration of the trial timeline for the joint part of the learning phase (A) and for the test phase (B). A) shows a valid trial in the joint learning phase. B) shows a test phase trial in which one participant produced a response error. Blue bars represent 1000 ms response windows in which responses were recorded in both phases. The trial timeline for the solo part of the learning phase was identical to A) despite minor differences (see main text).

their partner who could either approve the proposal or propose the other alternative instead. The assignment was instructed by means of the prompts "Propose!" and "Approve!" printed on the side of the screen corresponding to participants' seating positions and was balanced across trials. Pairs were instructed to produce a roughly balanced amount of low and high chords throughout the experimental phase. As soon as participants found an agreement, the experimenter registered their choice and proceeded the trial manually[9].

First, pairs' choice was centrally displayed as a text prompt on screen (e.g., "high chord") for 1000 ms. Then, a counter appeared on screen, counting in an interval of 500 ms from three down to a "GO!" prompt that remained on screen for 500 ms. To produce the chords, participants had to respond as synchronously as possible at the onset of the "GO!" prompt[10]. The response window opened 500 ms before the onset of the "GO!" prompt and closed 500 ms after. The first registered response by one of the pair members that fell within this response window, and that matched the required response configuration, triggered the tone assigned to that response. If the second pair member's response also fell within the response window and within an asynchrony window of 250 ms with respect to the first pair member's response, the tone mapped to the second response was triggered too. The respective two-tone chord resulting from both tones played simultaneously was then played for a duration of 1000 ms.

If one pair member produced a response that did not match the required response configuration, the trial was ended, either stopping a previously triggered tone of their partner

---

[9] Initial proposals were approved by the second pair member in over 99% of all trials.

[10] Using a countdown in the learning phase had the aim to facilitate synchronized responding of both partners by making the onset of the GO! prompt predictable. This procedural aspect of the learning phase deviates from common procedures of the two-stage action-outcome learning task deployed in previous studies. An anonymous reviewer suggested that the synchrony instruction in the learning phase could have led participants to respond to a 500 ms delay after stimulus onset in the subsequent test phase because they were trained to do so in the learning phase. We deem this possibility rather unlikely because, in contrast to the GO! prompt in the learning phase, both timing and type of the test phase stimuli were randomized (see details below). Furthermore, as the timing demands in the learning and the test phase were identical across experimental conditions, they shouldn't be able to serve as an alternative explanation of predicted performance differences between experimental conditions.

or preventing registration of any further response. In this case, an error message ("Wrong response!") was displayed after the offset of the "GO!" prompt at the side of the participant who had performed the error. If one pair member did not produce a response within the response window, previously triggered tones of their partner were stopped, and an error message ("Missing response!") was displayed at the side of the participant who had failed to respond. If no response was registered within the response window, the error message was displayed on both sides of the screen. If both pair members produced correct responses within the response window but failed to meet the asynchrony demand between their responses, the firstly triggered tone was stopped immediately and the feedback message "Be more synchronous!" appeared centrally on screen. Feedback and error messages were displayed for 1500 ms.

Throughout each trial, an illustration of the two response boxes displayed side by side was shown at the bottom of the computer screen, mirroring participants' responses. As soon as a pair member produced a response within the response window, the respective response button in the display was highlighted until the end of the trial. The next trial started after a blank screen, displayed for 500 ms. There was as short break after every twenty successful trials (trials in which pairs produced a chord together) in which participants were informed about the ratio of high and low chords produced so far. The joint part of the learning phase ended if pairs had performed eighty successful trials.

Test Phase

The test phase started directly after the joint part of the learning phase with no break in between. Depending on the experimental condition, pairs either remained in their current seating position or were instructed to swap seats, each of them taking their individual response box with them to the new seating position (see Figure 3-2). Pairs were then instructed to respond to the auditory presentation of the same high and low two-tone chords they had produced in the preceding learning phase as quickly and as accurately as possible in accordance with a fixed

stimulus-response mapping displayed on screen[11]. The stimulus-response mapping was instructed through an illustration similar to the illustration used in the joint part of the learning phase. It depicted the two response boxes of both pair members side by side, highlighting the respective response buttons, and labelling them with "high" and "low" respectively (see Figure 3-3, Panel B). Given this instruction participants could construe their responses at the group-level (inner/outer), at an individual level (left/right) or both.

Furthermore, pairs were instructed to withhold any response when chords were presented together with the display of a red "X" on the computer screen (no-go trials)[12]. Depending on the experimental condition, the stimulus-response mappings required pairs to respond either with the same inner/outer response configuration they had performed together to produce the two chords in the preceding learning phase (Jointly Compatible mapping) or not (Jointly Incompatible mapping). At the same time, the stimulus-response mapping required each individual participant to respond either with the same left/right response they had performed individually to produce the two chords in the preceding learning phase (Individually Compatible mapping) or not (Individually Incompatible mapping) (see Figure 3-2).

The trial timeline for the test phase is illustrated in Figure 3-4, Panel B. Each trial started after a variable inter trial interval that ranged between 500 and 1500 ms with the presentation of the higher or the lower chord, played for a duration of 1000 ms. The chord was presented together with a white exclamation mark (go trials) or a red "X" (no-go trials) displayed centrally on screen. The response window was open for the duration of the chord. Participants received error feedback for incorrect or missing responses displayed for 1500 ms on the side of the pair member who had produced the error. Throughout each trial, an illustration of the two response boxes displayed side by side was shown at the bottom of the computer screen, mirroring

---

[11] Pair members were not explicitly instructed to synchronize their responses with each other.

[12] No-go trials served as catch trials to keep participants' attention centred on the screen.

participants responses. As soon as a pair member produced a response within the response window, the respective response button in the display was highlighted until the end of the trial.

Pairs performed 50 trials that were divided into 5 sub-blocks à 10 trials. Within each sub-block, the number of high and low chords was balanced, and two trials (one high chord and one low chord trial) were no-go trials. Trial order within each sub-block was randomized. The first sub-block served as training, after which the instruction slide depicting the instructed stimulus-response mapping was shown again. The remaining trials proceeded with no breaks in between.

3.2.4. Design

All pairs performed two test phase blocks, one time with a jointly compatible and one time with a jointly incompatible stimulus-response mapping. Before performing the second test phase block in the remaining experimental condition, pairs repeated the joint part of the learning phase for a second time in the same seating positions as before. The order of the two test-phase blocks was counterbalanced across pairs. Individual Compatibility of the stimulus-response mappings was manipulated between pairs, so that one half of the pairs performed both test phase blocks with an Individually Compatible and the other half of the pairs with an Individually Incompatible stimulus-response mapping.

As specified in our preregistration, we eventually limited our design to pairs' first test phase block only, as preliminary analysis of participants test phase performance with Joint Compatibility as a within-subjects factor revealed significant block order effects (see Appendix B2). This left the study with a 2 (Individual Compatibility: Individually Compatible vs. Individually Incompatible) x 2 (Joint Compatibility: Jointly Compatible vs. Jointly Incompatible) fully between-subjects factorial design.

3.2.5. Data analysis

Learning phase

For the *solo part* of the learning phase, trials with response omissions were removed prior to analysis (4.4% of all trials). From the remaining trials, error rates ($ER_{LearnSolo}$, relative frequency of trials in which participants failed to produce a tone) and the ratio of high and low tones produced on successful trials were aggregated for each participant. For the *joint part* of the learning phase, trials in which both pair members omitted responses were removed prior to analysis (0.4% of all trials). From the remaining trials, error rates ($ER_{LearnJoint}$, relative frequency of trials in which pairs failed to produce a chord together)[13] and the ratio of high and low chords produced on successful trials were aggregated for each pair.

To assess the possibility of potential differences between participants in the four test phase conditions regarding their learning phase performance, error rates and outcome ratios in the solo and joint part of the learning phase were analysed as a function of the experimental conditions participants performed in the subsequent test phase. The ratio of high tones produced in the solo part of the learning phase and the ratio of high chords produced in the joint part of the learning phase were compared against chance by means of one-sampled *t*-tests, separately for each of the four test phase groups resulting from the factorial design. Error rates in the solo and in the joint part of the learning phase were compared between the four test phase conditions by means of an ANOVA with Individual Compatibility (Individually Compatible vs. Individually Incompatible) and Joint Compatibility (Jointly Compatible vs. Jointly Incompatible) as between-subjects factors.

---

[13] This includes trials with missing responses by one and wrong responses by one or both pair members as well as trials in which pair members responded not synchronously enough.

Test Phase

The first ten trials of the test phase served as training trials to familiarize participants with the trial procedure and were therefore excluded from analysis. For the remaining test trials, go- and no-go trials were separated and response omissions on go-trials were removed (3,9% of all go-trials). For no-go trials, the frequency of erroneous responses was calculated for each participant. For go-trials, error rates (ER_{Test}, relative frequency of wrong responses) and mean response times (RTs) on correct trials were calculated for each participant. As specified in our preregistration, we also calculated a combined measure of participants response performance on go-trials accounting for speed-accuracy trade-off by combining error rates and mean RTs into inverse efficiency scores (IES), calculated as mean $RT/(1-ER_{test})$ (Bruyer & Brysbaert, 2011). Furthermore, asynchronies between valid go-trial responses of both pair members were aggregated for each pair by calculating pairs' mean absolute response asynchronies (|ASY|) on valid go-trial.

As analysis of error rates in the joint part of the learning phase revealed significant differences between the four test phase groups regarding their learning phase performance (see Appendix B1), we included error rates of the joint part of the learning phase ($ER_{LearnJoint}$) as a covariate in the statistical analysis of participants' test phase performance. Hence, all test phase measures were analysed as a function of the four test phase conditions by means of an ANCOVA with Individual Compatibility (Individually Compatible vs. Individually Incompatible mapping) and Joint Compatibility (Jointly Compatible vs. Jointly Incompatible mapping) as between-subjects factors and error rates of the joint part of the learning phase ($ER_{LearnJoint}$) as a covariate[14].

---

[14] For analysis of participants' individual performance measures in the test phase (IES, RTs and ER_{Test}), pair-level error rates in the joint part of the learning phase were treated as an individual-level variable for each pair member.

3.2.6. Open science statement

Sample size, data exclusion criteria, analysis plan and directed hypotheses were preregistered in the Open Science Framework. The preregistration can be accessed under https://osf.io/vk348. The inclusion of $ER_{LearnJoint}$ as a covariate in the analysis of participants' test phase performance was not preregistered, as differences between the four test phase groups regarding their learning phase performance was unexpected.

3.3. Results

Results for the learning phase can be found in Appendix B1. The results for the test phase analysis are depicted in Figure 3-5.

Analysis of participants' IES (Figure 3-5, Panel A) revealed a significant main effect of the Joint Compatibility factor, $F(1,67) = 9.23$, $p = .003$, $\eta_p^2 = .121$. Yet, opposite to our initial predictions, participants' response performance was less efficient (implying higher IES) with a Jointly Compatible mapping (estimated marginal mean [EMM] = 667 ms, 95%CI [634 ms, 701 ms]) compared to a Jointly Incompatible mapping (EMM = 592 ms, 95%CI [557 ms, 627 ms]). There was also a significant effect of the covariate ($ER_{LearnJoint}$), $F(1,67) = 9.42$, $p = .003$, $\eta_p^2 = .123$, indicating that participants responded less efficiently in the test phase the more errors they had made together with their partner in the joint part of the learning phase ($r(70) = .474$, $p < .001$). Neither the main effect of the Individual Compatibility factor, nor its interaction with the Joint Compatibility factor, were significant (both $F < 1$). Further explorative analysis of the Individual Compatibility factor by means of an independent samples Bayesian t-test showed that the IES data was $BF_{01} = 9.2$ times more likely under the null hypothesis predicting the

absence of a performance advantage in the individually compatible compared to the individually incompatible test phase conditions.
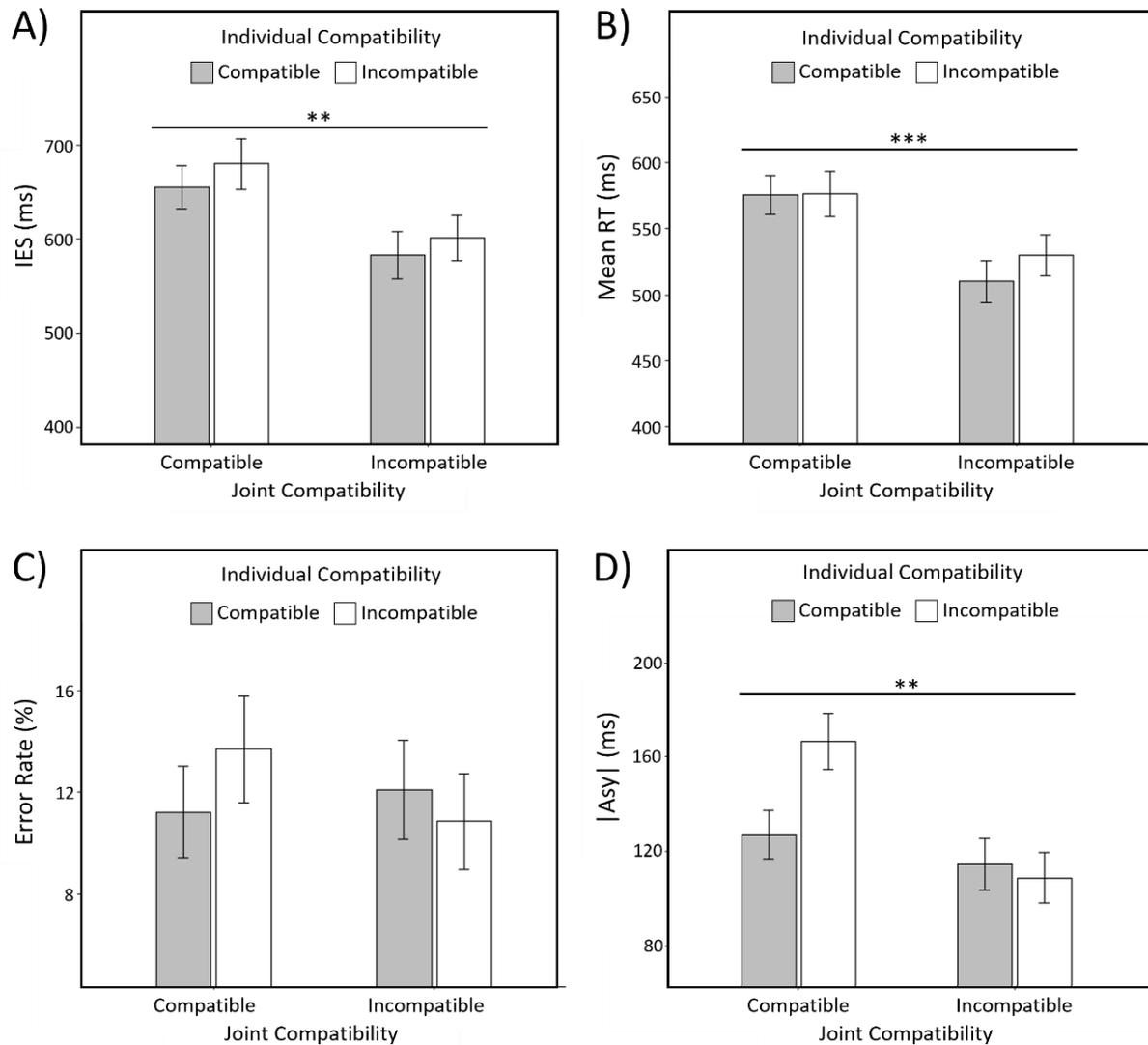
Figure 3-5. Condition means of participants' IES (A), mean RTs (B) and error rates (C) and of pairs' mean absolute response asynchronies (|Asy|) (D) in the test phase. Condition means represent estimated marginal means adjusted for inclusion of ERLearnJoint as a covariate in the comparison. Error bars represent the respective standard errors of the mean. Significant main effects are marked with asterisks (** $p < .01$; *** $p < .001$).

The analysis of participants' mean RTs is displayed in Figure 3-5, Panel B. Again, there was a significant main effect of the Joint Compatibility factor, $F(1,67) = 12.6$, $p < .001$, $\eta_p^2 = .159$, but — as for IES — the effect was in the opposite direction as predicted. RTs in the Jointly Compatible conditions (EMM = 576 ms, 95%CI [555 ms, 597 ms]) were slower compared to the Jointly Incompatible conditions (estimated marginal means = 519 ms, 95%CI [498 ms, 542

ms]). There was also a significant effect of the covariate ($ER_{LearnJoint}$), $F(1,67) = 4.87$, $p = .03$, $\eta_p^2 = .067$, reflecting slower response times in the test phase the more errors participants had made together with their partner in the joint part of the learning phase ($r(70) = .358$, p = .002). Again, neither the main effect of the Individual Compatibility Factor, nor its interaction with the Joint Compatibility factor were significant (both $F < 1$).). Further explorative analysis of the Individual Compatibility factor by means of an independent samples Bayesian t-test showed that the RT data was $BF_{01} = 7.4$ times more likely under the null hypothesis predicting the absence of a response time advantage in the jointly compatible compared to the jointly incompatible test phase conditions.

Analysis of participants' error rates on go-trials of the test phase (Figure 3-5, Panel C) revealed no significant main or interaction effects (all $F < 1$). Only the effect of the covariate ($ER_{LearnJoint}$) on participants' test phase error rates was significant, $F(1,67) = 5.01$, $p = .03$, $\eta_p^2 = .07$, reflecting higher error rates in the test phase the more errors participants had made together with their partner in the joint part of the learning phase ($r(70) = .374$, $p = .001$).

Analysing the frequencies of participants' erroneous responses on no-go trials revealed no significant main or interaction effects (all $F < 1$ after controlling for the effect of the covariate, $F(1,67) = 3.168$, $p = .08$, $\eta_p^2 = .05$).

The analysis of mean absolute asynchronies between both pair members responses (Figure 3-5, Panel D), revealed a significant main effect of Joint Compatibility, $F(1,31) = 10.31$, $p = .003$, $\eta_p^2 = .25$, with larger absolute asynchronies in the Jointly Compatible conditions (EMM = 147 ms, 95%; CI [132 ms, 162 ms]) compared to the Jointly Incompatible conditions (EMM = 112 ms; 95%CI [96 ms, 128 ms]). The covariate ($ER_{LearnJoint}$) had no significant effect on absolute asynchronies between pair members responses in the test phase ($F < 1$). The main effect of the Individual Compatibility factor was non-significant, $F(1,31) = 2.38$, p = .13, $\eta_p^2 = .071$; the interaction between the Joint and the Individual Compatibility factor approached but was shortly off statistical significance, $F(1,31) = 3.81$, p = .06, $\eta_p^2 = .109$. Further explorative

correlation analysis of participants' mean absolute response asynchronies relative to their partners' responses revealed a positive relationship with participants' mean RTs, $r(1,35) = .29$, $p = .015$[15], indicating that higher asynchronies between pair members responses were associated with overall slower responses of both pair members.

3.4. Discussion

The present study tested two possible accounts of how co-actors might learn and represent novel instrumental relationships between synchronized action contributions and the joint outcomes they produce together. According to a *minimal account* of joint action representations, action-outcome learning should only be sensitive to each co-actor's individual contributions to the joint action, which should have been reflected in a modulation of participants' test phase performance by means of our individual compatibility manipulation. According to a *group-level account* of joint action representations, action-outcome learning should be sensitive to relations between co-actors individual contributions to the joint action, which should have been reflected in a modulation of participants' test phase performance by means of our joint compatibility manipulation. Our results showed that joint but not individual compatibility manipulations of the instructed stimulus-response mappings affected participants' response performance in the test phase. However, the result pattern we observed was not in line with the directed predictions we derived from associative accounts of action-outcome learning.

3.4.1. Action-outcome learning on the individual level

Against the prediction we derived from of the minimal account, we found no evidence that participants' test phase performance benefitted from a stimulus-response mapping that was compatible with the preceding learning phase regarding individual-level features of each co-

---

[15] Degrees of freedoms were adjusted to the number of pairs.

actor's isolated response contributions (left/right). This indicates that presentation of the two-tone chords in the test phase was unlikely to prime learned associations with individual-level features of participants' isolated response contribution. This would imply that participants did not represent the chord outcomes in the learning phase to be contingent on their individual action contributions alone.

We propose that the lack of an individual compatibility effect can be attributed to the peculiarities of the synchronous joint action setting investigated in the present study. First, during the joint part of the learning phase, participants produced distinct chord outcomes that were contingent on the *compound* of two synchronous response contributions of both co-actors. This may have obscured contingent relationships between perceived outcomes and the isolated response contributions of each individual co-actor. The finding that co-actors did not acquire action-outcome associations on the individual level may thus be explained by *overshadowing*, an effect that has been observed in research on classical conditioning (e.g., Pavlov, 1927) and human contingency learning (e.g., Dickinson et al., 1984; Shanks, 1989). Research in these domains has shown that learners tend to become insensitive to contingency relations between two consecutive events (e.g., an individual action and a perceived outcome) if the preceding event occurs in compound with another potential predictor (e.g., the action of another co-actor). This interpretation would also be in line with recent studies providing direct evidence for limitations of associative action-outcome learning in individual action contexts when the number of action and outcome possibilities increase beyond a limited set of simple one-to-one mappings (Flach et al., 2006; Watson et al., 2015).

Second, many studies in the domain of individual action indicate that learning and retrieval of associations between outcomes and low-level response features is not a quasi-automatic process but appears to be modulated by intentional and attentional factors (c.f., Herwig & Waszak, 2009; Kiesel & Hoffmann, 2004; Pfister, 2019; Vogel et al., 2020; Zwosta et al., 2013). These studies indicate that people link perceived outcomes to *selective* features of

66

their actions that are determined by top-down interpretational processes defining how produced actions are currently encoded (see also Ansorge & Wühr, 2004). Thus, the missing evidence for action-outcome learning on the individual level may indicate that participants encoded their responses in the learning phase not simply as individual left/right responses.

As we attribute the lack of an individual compatibility effect in our study to the peculiarities of synchronous joint action, future studies may extend our research to other joint action scenarios in which individual and joint action contributions and their respective outcomes are more clearly distinguishable in space and time. This would be the case in sequential joint actions in which co-actors produce individual actions / action outcomes in turns with each other to produce more distal joint action outcomes over time (e.g., Sacheli et al., 2018). Thus, an interesting question for future research would be whether action-outcome learning remains sensitive to individual-level response features of co-actors isolated response contributions in other joint action contexts that deviate from the special case of synchronous joint action investigated in the present study.

3.4.2. Action-outcome learning on the group level

Turning to the prediction we derived from the group-level account, we did not find evidence that participants' test phase performance benefitted from a stimulus-response mapping that was compatible with the preceding learning phase regarding group-level features of co-actors' joint response contributions (inner/outer). In contrast, we observed an unexpected effect in the opposite direction, reflecting a performance advantage for participants who received jointly *incompatible* stimulus-response mappings in the test phase. This unexpected *reversed joint compatibility effect* speaks against the hypothesis that presentation of the two-tone chords in the test phase led to an automatic activation of associated group-level features of co-actors' joint response contributions. This finding indicates that the mechanisms proposed by simple

associative accounts of goal-directed action (e.g., ideomotor theories) are not directly extendable to incorporate group-level features of joint action performance.

Nevertheless, the reversed joint compatibility effect still requires an explanation of why co-actors were reliably affected by alterations of group-level relations between their own and their partner's contributions to the joint action from learning to test phase. This finding implies that the action-outcome representations participants acquired in the learning phase must have been sensitive to group-level relations between their own and their partner's contributions to the joint action in some way.

A possible explanation of the reversed joint compatibility effect can be derived from theories postulating that acquired representations of action-outcome relationships are not stored as rigid bidirectional associations formed in long-term memory but as propositional knowledge structures represented in current working memory (Custers, 2023; Mitchell et al., 2009; Seabrooke et al., 2016; Sun et al., 2022). According to these accounts, propositional representations of action-outcome relationships acquired in the learning phase would have influenced participants' test phase performance, not by priming responses associated with a former action outcome on a trial-by-trial level, but by modulating how participants translated the task instructions at the start of the test phase into task-relevant stimulus-response rules held in procedural working memory during the ensuing task (i.e., task sets, c.f., Brass et al., 2017; Hazeltine & Schumacher, 2016; Monsell, 2003; Rogers & Monsell, 1995).

Thus, the reversed joint compatibility effect may have arisen because participants — as a function of their previous learning experience — performed the test phase in the respective conditions with different task-sets in mind, encoding the ambiguously instructed stimulus-response mappings at the start of the test phase either in relation to their group-level performance (i.e., if *WE* hear a high/low chord, *WE* press our inner/outer buttons) or merely in relation to their individual-level performance instead (i.e., if *I* hear a high/low chord, *I* press my left/right button).

68

Specifically, pairs who received *jointly compatible* test phase instructions might have been inclined to encode the instructed stimulus-response mappings in relation to their group-level performance, as a group-level construal of the task instructions could be most easily reconciled with their previous group-level performance in the learning phase. In contrast, pairs who received *jointly incompatible* test phase instructions may have been reluctant to encode the instructed stimulus-response mappings in relation to their group-level performance, as a group-level construal of the task instructions stand in conflict with their previous group-level performance in the learning phase. Due to this *conflict during task-set formation* (c.f., Monsell et al., 2001), participants receiving jointly incompatible test phase instructions may have reverted to encode the instructed stimulus-response mappings in relation to their individual-level performance instead to ensure efficient response performance in the upcoming task.[16]

Following this explanation, the reversed joint compatibility effect could stem from the fact that implementing stimulus-response rules encoded on a group level (i.e., in the jointly compatible test phase conditions) raises higher cognitive demands compared to implementation of stimulus-response rules encoded merely in relation to people's individual-level performance (i.e., in the jointly incompatible test phase conditions).

---

[16] Notably, this idea would be in line with Vallacher and Wegner's (1985, 1987, 2012) theory of action identification which formulates basic principles of how people conceptualize their actions. The theory assumes that people can conceptualize their actions at different hierarchical levels of abstractions, ranging from lower-level interpretations related to specific movements they perform to higher-level interpretations related to more distal ends of their actions. At which level people construe their actions is thought to be determined by three interconnected principles, stating that 1) a prepotent level of action identification is maintained unless 2) an action can be conceptualized at a higher level, creating a tendency to change to that higher level or 3) an action cannot be performed in terms of the prepotent level of action identification, creating a tendency to revert to a lower level. Applied to our study, the theory predicts that co-actors would tend to conceptualize their actions at the highest level of abstraction afforded by the task (i.e., in terms of their group-level performance) until it would create problems for efficient task performance, at which point co-actors would tend to construe their actions at a lower level instead (i.e., in terms of their individual-level performance).

First, a possible reason for this could be that the implementation of stimulus-response rules encoded on a group level may require an additional processing step that specifies people's response contribution not only at the group level but at the individual level as well (e.g., *WE* press our inner buttons, so *I* press my left/right button). This assumption of a hierarchical specification of joint action representations would be in line with current theoretical models of joint action planning implicating a cascading processing hierarchy that proceeds from higher-level action representations related to group-level performance to lower-level action representations related to the individual-level action contributions of the separate co-actors (Candidi et al., 2015; Keller et al., 2016; Pacherie, 2012; Pesquita et al., 2018; Sacheli et al., 2018; Sinigaglia & Butterfill, 2022; Zapparoli et al., 2022). Therefore, participants' may have performed worse in the jointly compatible compared to the jointly incompatible test phase conditions because they accessed individual-level representations of their own response contributions only indirectly, mediated through a higher-order action representation at the group level. In contrast, participants in the jointly incompatible test phase conditions would have performed better because they were able to access individual-level representations of their own response contributions directly without requiring an additional processing step.

Second, implementing stimulus-response rules encoded on a group level may also involve further specification of individual-level response features of other co-actors' contributions to the groups' response as well (e.g., *WE* press our inner buttons, so *I* press my left/right button and *YOU* press your right/left button). This would be suggested by research findings demonstrating that people tend to co-represent others' tasks and actions (Atmaca et al., 2008; Kourtis et al., 2013; Novembre et al., 2014; Schmitz et al., 2018; Sebanz et al., 2003, 2005) as long as they are perceived as co-acting partners contributing to a joint task (Kourtis et al., 2010; Meyer et al., 2011; Sacheli et al., 2018, 2021). Thus, participants may have performed worse in the jointly compatible compared to the jointly incompatible test phase conditions because they faced additional processing costs associated with the specification of their partner's response

70

contributions that would have been absent in participants in the jointly incompatible test phase conditions who encoded the stimulus-response rules merely in relation to their own individual-level performance.

3.4.3. Conclusion

Our study investigated how co-actors acquire joint action representations by learning novel instrumental relationships between synchronous joint actions and the outcomes they produce. While our results reveal limitations of purely associative accounts of action-outcome learning in synchronous joint action contexts, they can be explained by propositional accounts and align with a growing body of evidence that joint action is supported by dedicated mental representations encoding group-level relations between co-actors' joint action contributions. Taken together, our study informs current theorizing on the mechanistic nature of action-outcome learning in individual and joint action settings and provides novel evidence that co-actors tend to form group-level representations of their joint actions if afforded by given task constraints.

4. STUDY III and IV:

Modulations of Imitative Response Tendencies

in Joint Task Settings

4.1. STUDY III: Do individual and joint action goals modulate imitative response tendencies?

4.1.1. Introduction

To engage in social interactions, people need to interpret the actions they observe in others and produce appropriate and timely responses to them. Understanding how these processes work in concert and enable interaction partners to coordinate their actions is a central question for research on joint action (Sebanz et al., 2006). A common assumption in this field of research is that people process others' actions through the lens of their own action capacities, i.e., that they reuse cognitive resources deployed for action execution to simulate and predict the actions they perceive in others (Wilson & Knoblich, 2005; Wolpert et al., 2003). This idea has gained ample empirical support from neuroimaging studies showing that action observation and action execution recruit similar brain circuits (for a review see e.g., Rizzolatti & Sinigaglia, 2016) as well as from behavioural research showing that people tend to imitate the actions they perceive in others (Heyes, 2011; c.f., Genschow & Cracco, 2025 for an exhaustive collection of relevant research findings). This imitative response tendency is indexed by the well-replicated finding that actions are initiated faster while observing similar (i.e., congruent) compared to dissimilar (i.e., incongruent) actions of somebody else (imitative congruency effect; see Cracco et al., 2018 for a meta-analysis). Prominent theoretical accounts propose that this imitative congruency effect is the result of a direct matching process between observed and self-executed

movement features, suggesting that action observation automatically and obligatorily activates corresponding motor representations in the observer (e.g., Heyes, 2011; Iacoboni et al., 1999; Rizzolatti et al., 2001).

Direct matching between perceived and self-executed actions can support social interactions by allowing interaction partners to predict and align their actions (Knoblich et al., 2011). However, it also poses challenges for social interactions that require interaction partners to perform dissimilar yet complementary actions to achieve their goals (Sartori & Betti, 2015). For instance, handing someone a glass of champagne requires well-coordinated, yet different ways of grasping the glass by both interaction partners. Arguably, in these situation, automatic imitation of observed movements would interfere with task performance, raising the important question of whether and how our cognitive system can modulate imitative response tendencies in a flexible and context appropriate way (Brass & Heyes, 2005a).

A possible solution to this problem is offered by goal-directed theories of imitation. According to these accounts, imitation is not the result of an automatic matching process between observed and self-executed movement features, but rather an interpretative process in which individuals infer and reproduce the *goals* underlying others actions (e.g., Bekkering et al., 2000; Csibra, 2007; Gattis et al., 2002; Wohlschläger et al., 2003; c.f., Heyes, 2013 for discussion). Central to these accounts is the assumption that actions can be represented at different, hierarchically organized levels of abstraction, ranging from overarching *task goals* (e.g., building a tower) over more proximal *action goals* (e.g., placing a brick onto another) to specific *movement goals* (e.g., lifting one's right arm) (c.f., Grafton & Hamilton, 2007; Kilner et al., 2007; Pacherie, 2008; Uithol et al., 2012; Vallacher & Wegner, 2012). Goal-directed accounts of imitation assume that observers prioritize the reproduction of higher-level goals over lower-level movement patterns, such that tendencies to imitate low-level movement features of observed actions can become overruled by interpretations of observed actions in terms of their higher-level task and action goals (Ondobaka & Bekkering, 2012, 2013).

73

The influence of action goals on imitation

While a growing body of evidence supports goal-directed theories of imitation, whether — and if so how — higher-level goal representations modulate imitative response tendencies still remains poorly understood. Preliminary evidence has been provided by studies showing that interference effects between perceived and executed *movement goals* are susceptible to imitative congruency relations between own and others higher-level *action goals* (Massen & Prinz, 2007; Ondobaka et al., 2012, 2013). E.g., a seminal study by Ondobaka et al. (2012) showed that imitative congruency effects between own and other's movement goals (e.g., arm reaching movements to the left or the right) emerge only if co-actors perform their movements to achieve congruent action goals (e.g., selecting similar action targets), but not if co-actors perform their movements to achieve incongruent action goals (e.g., selecting dissimilar action targets). Moreover, performing actions towards congruent compared to incongruent action goals was shown to benefit participants' response performance overall, indicating that people have a tendency to imitate higher-level action goals of others independent of the means they choose to reach them.

These findings indicate a central role of action goals in shaping imitative congruency effects during social interactions and suggest a hierarchical control mechanism for imitative behaviour in which congruency relations between observed and self-executed actions at higher levels (e.g., action goals) modulate imitative congruency effects at lower-levels of action representation (e.g., movement goals). However, this conclusion has been challenged by recent studies that failed to provide conclusive evidence for a modulating role of action goals on imitative congruency effects between own and others lower-level movement goals (Cole et al., 2018; Janczyk et al., 2016). As such, the extent and conditions under which higher-level action goals become able to modulate imitative congruency effects between lower-level movement goals still remains a subject of ongoing debate (see also, Chiavarino et al., 2013; Leighton et al., 2010; Longo et al., 2008; Wild et al., 2010).

74

The influence of joint task goals on imitation

Further evidence for the notion that higher-level goal representations modulate imitative response tendencies in social interactions comes from research on joint action (see McEllin et al., 2018 and Sebanz & Knoblich, 2021 for reviews). As mentioned above, joint action often requires people to perform actions that are incongruent yet complementary to those of their co-actors (Sartori & Betti, 2015), making joint action an interesting test case for the flexibility of imitative response tendencies. Initial evidence for modulations of imitative response tendencies in joint action contexts has been provided by a range of studies showing that explicit instructions and implicit requests to *complement* rather than to imitate observed actions of an interaction partner can reduce or even reverse imitative congruency effects between own and other's movement goals (Bardi et al., 2015; Betti et al., 2019; Newman-Norlund et al., 2007; Ocampo et al., 2011; Ocampo & Kritikos, 2010; Poljac et al., 2009; Sartori et al., 2012, 2013; van Schie et al., 2008). Going beyond research on the role of individual action goals in modulating imitative congruency effects described above, these findings suggest a special role for goal representations that encode the joint outcomes of co-actors' complementary action contributions, i.e., what they are achieving together as a group (see Butterfill, 2012 and Sacheli et al., 2015). We refer to this level of goal representation as *joint task goals*.

Crucially, direct support for the notion that joint task goals shape imitative response tendencies has been provided by recent studies showing that imitative congruency effects between own and others movement goals are indeed reduced (Clarke et al., 2019; Rocca et al., 2023; Schmitz et al., 2023) or even absent (Sacheli et al., 2018, 2019) when own and others' movements are performed as interrelated contributions towards a joint goal. To explain these findings, theoretical models of joint action propose that representations of joint task goals lead co-actors to integrate representations of their own and their partners' individual action contributions into hierarchical joint action plans that encode relational constrains between co-

actors' individual action and/or movement goals (Candidi et al., 2015; Clarke et al., 2019; Pesquita et al., 2018; Sacheli et al., 2018; Sinigaglia & Butterfill, 2022; Zapparoli et al., 2022).

The present research

In light of the reviewed findings, we set out to further investigate if and how higher-level goal representations modulate imitative response tendencies during social interactions. Our motivation to do so was two-fold. First, we thought to re-assess previous research presenting mixed evidence for the notion that individual action goals modulate imitative congruency effects between own and others' lower-level movement goals. To this end, we conducted two conceptual replications of the study by Ondobaka et al. (2012) in which we manipulated imitative congruency relations between the actions of two co-actors at the level of their individual action and movement goals. Second, building on previous indications that joint task goals reserve a dedicated role for coordinating co-actors' individual contributions to a social interaction, we further asked whether and how *joint* as opposed to *individual* task goals add to the hierarchical control of imitative response tendencies in social interactions. To do so, in both experiments, we further instructed different groups of participants to either work towards individual or joint task goals, testing whether joint task goals might further modulate imitative congruency effects between co-actors action and movement goals.

In two experiments, participants performed a sequential target selection task together with a virtual partner, modelled after Ondobaka et al. (2012). On each trial, participants first observed a virtual partner (in the following referred to as the co-actor) selecting the higher or the lower of two playing cards (Experiment 1) or line drawings (Experiment 2) and where required to select one of two playing cards or line drawings themselves in response. Participants target selection was based on condition specific task instructions that required participants to select either the higher or the lower of their targets respectively (we refer to this as participants' action goal). Participant's and their co-actor's action targets were presented side by side, and

76

participants selected their respective target on a given trial by pressing a left or a right key on their keyboard to select the left or the right of their targets respectively (we refer to this as participants' movement goal). In both experiments, we manipulated imitative congruency between co-actors' action goals (i.e., whether participants had to select the higher or the lower of their targets after observing their co-actor selecting the higher or lower of their targets) and between co-actors' movement goals (i.e., whether participants had to select the left or the right of their targets after observing their co-actor selecting the left or the right of their targets)

Deviating from the original study by Ondobaka et al. (2012), congruency between co-actors' action goals was manipulated between- rather than within-subjects to rule out possible conflict or carry over effects between the different task instructions. Moreover, presenting the targets of both co-actors side by side deviates from the studies by Cole et al. (2018) Janczyk et al. (2016) and Ondobaka et al. (2012) in which co-actors' action targets were presented opposite to each other, so that the their absolute spatial locations overlapped. In these setups, imitative congruency effects at the level of co-actors movement goals (e.g., faster selection of left targets after observing one's partner responding to their left compared to their right target) have been attributed to effects of attentional disengagement from target locations cued by previous partner responses (referred to as *social inhibition of return,* see Cole et al., 2019 for review). By presenting co-actors' action targets side by side, our setup eliminates spatial overlap between co-actors' absolute target locations, ensuring that imitative congruency effects between co-actors' movement goals can be attributed to automatic imitation (which are the focus of the present study) rather than to purely attentional mechanisms.

Crucially, further extending the experimental design by Ondobaka et al. (2012), task instructions for selecting an action target on a given trial was manipulated between two differed groups of participants (we refer to this as participants' task goal). While one group of participants was explicitly instructed to always match or mismatch the action goal of their co-actor by selecting the higher or the lower of their targets (individual task goal condition),

77

another group of participants was instead instructed to complement the card selection of their co-actor to achieve a joint outcome that specified specific relations between their individually selected action targets (joint task goal condition).

Based on previous research by Ondobaka et al. (2012, 2013), for the individual task goal condition, we expected to find imitative congruency effects at the level of co-actors' individual action goals as well as at the level of their lower-level movement goals. More specifically, in line with the idea of an hierarchical control mechanism for imitative behaviour, congruency relations between co-actors' individual action goals should modulate imitative congruency effects between co-actors' movement goals. Furthermore, based on research suggesting a dedicated role of joint task goals in modulating imitative response tendencies in social interactions, we tested whether imitative congruency effects between co-actors' individual action and movement goals would be further modulated by the joint task goal to complement another person's action in the service of a joint goal. In particular, based on the idea that joint task goals modulate the impact of individual action and movement goals on imitative behaviour, we hypothesized that imitative congruency effects at the level of co-actors' individual action and movement goals would be reduced or even absent in the joint task goal condition.

4.1.2. Experiment 1

Methods

Transparency and openness. Data for Experiment 1 was collected in April 2023. Raw data, analysis scripts, stimuli, and experiment scripts are accessible in an Open Science Framework project repository (https://osf.io/wc96q/?view_only=dd5091015d744e81a0295afdb0050d49). The study design, hypotheses, sampling plan and data exclusion criteria were preregistered in the Open Science Framework. The preregistration is accessible under https://osf.io/5dn47/?view_only=04880f3bbccf4e3ba8c7bda9f5eca11f.

Participants. Two considerations determined our target sample size. First, we powered the experiment to replicate the action goal congruency effect reported by Ondobaka et al. (2012) in a between-subjects design. Therefore, Experiment 1 was powered to detect a medium to large sized effect of $d = 0.6$ in a two-samples t-test with 80% power at an alpha level of .05 (two-tailed). Conducting an a priori power analysis in G*Power (Faul et al., 2007) with these input parameters, we calculated a minimum sample size of 45 participants per group. Note that 45 participants per group also yields sufficient power to detect a crossed interaction pattern between action and movement goal congruency in our experimental design (see simulation results by Brysbaert, 2019). Second, since we were not aware of any published research that could guide power calculations for a possible modulation of imitation effects between the individual and the joint task goal condition, we decided to adopt the above number of participants per group for the joint task goal condition as well, leading to a total target sample size of $N = 180$. Conducting a sensitivity analysis in G*Power showed that this sample size would be sufficient to detect an interaction effect between the two between-subjects factors in our design of $d = 0.42$ with 80% power at an alpha level of .05. To account for potential exclusion of participants, we eventually decided to collect data from 200 participants, split into 50 participants per group of our 2 x 2 between-subjects experimental design (see below). Data of eight participants were removed (see preregistered exclusion criteria), so that our final sample size included $N = 192$ participants (mean age = 30.8, SD = 9.2; 49 % female, 51 % male). Participants were recruited on Prolific (https://www.prolific.com). All participants had normal or corrected-to-normal vision and hearing and were fluent in English. All participants gave written informed consent and received monetary compensation for their participation (2 GBP for a total duration of approx. 12 minutes). Ethical approval for the study was granted by the Psychological Research Ethics Board (PREBO) of Central European University.

Stimuli and apparatus. Participants completed the experiment online in a browser on their own personal computer. Experiment files and data were stored on a commercial server provider for psychological experiments (https://www.pavlovia.org). Stimuli presentation and response recording was controlled by a custom made script written in the jsPsych JavaScript framework (de Leeuw et al., 2023).

Participants were presented with four horizontally aligned playing cards, displayed against a green background (see Figure 4-1). The cards were drawn from a subset of a standard card deck comprising the numerical card values of the clubs' suit in the range from two to nine. Two of the cards were assigned to a virtual partner and were presented on the left half of the screen. They always showed one out of six possible card combinations from the card set, containing card values in the range from three to eight with a numerical distance between cards larger than two (i.e., [3|6], [3|7], [3|8], [4|7], [4|8], [5|8]). The other two cards were assigned to
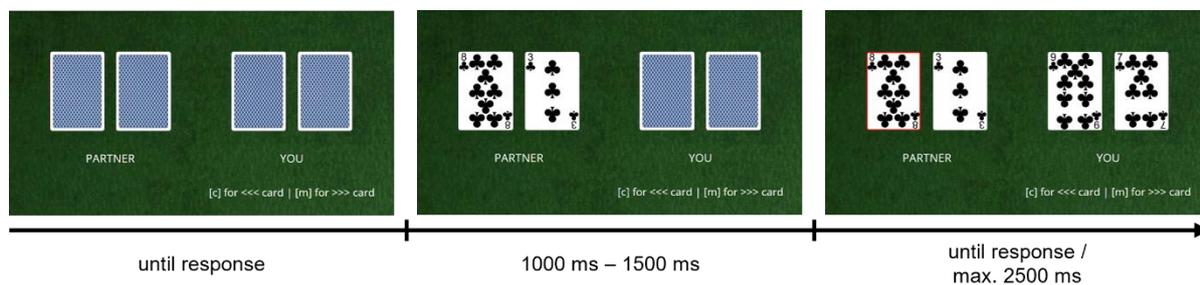
Figure 4-1. Trial timeline and stimuli example for Experiment 1. In the individual task goal condition, participants were instructed to either match or mismatch the card selection of their co-actor (indicated by the red frame) by selecting the higher or the lower of their cards. I.e., in the example above, half of the participants had to select the higher of their cards (clubs nine) while the other half had to select the lower of their cards (clubs seven) after observing their co-actor selecting the higher of their cards (clubs eight). In the joint task goal condition, participants were instead instructed to select one of their cards that complemented an ascending or a descending number sequences together with the card selected by their co-actor. Here, one half of the participants was instructed to complement an ascending sequence when their partner selected the higher of their cards and to complement a descending sequence when their partner selected the lower of their cards (i.e., selecting clubs nine as the successor of clubs eight in the example above), while the other half was instructed to complement an ascending sequence when their partner selected the lower of their cards and to complement a descending sequence when their partner selected the higher of their cards (i.e., selecting clubs seven as the predecessor of clubs eight in the example above). Note that selecting the correct card in each condition could require participants to select the card at the same or the opposite relative spatial position as their co-actor (i.e., selecting the left or the right of their cards after the co-actor selected the left of their cards).

the participants and were presented on the right half of the screen. Importantly, they always showed the next higher and the next lower card value relative to the card selected by the co-actor on a given trial (e.g., if the co-actor selects clubs' eight, the two cards of the participants were clubs' seven and clubs' nine). Participants selected the left or the right of their assigned cards by pressing the "c" (left) and the "m" (right) key on their keyboard respectively. Figure 4-1 depicts an example arrangement of the four playing cards on an individual trial.

Procedure and design. Each trial started with the presentation of all four playing cards faced down. Participants started a trial by pressing the space bar. First, the two cards of the co-actor were revealed. After a variable delay (1000 ms - 1500 ms), one of the co-actor's cards was highlighted with a red frame that indicated their card selection and the two cards of the participant were revealed. Participants then had to select one of their cards as quickly as possible within a maximum response window of 2500 ms. If participants gave a response, their selected card was highlighted by a red frame for 500 ms and participants received feedback (displayed for 2000 ms) informing them whether they had selected the correct card ("correct"), the wrong card ("error") or had failed to respond within the response window ("too slow"). The next trial started after a blank screen displayed for 500ms. The trial sequence is illustrated in Figure 4-1.

The rule that specified which card participants had to select on a given trial differed between four experimental groups. In the individual task goal condition, participants were instructed to always "match" or to always "mismatch" the card selection of the co-actor by selecting the higher or the lower of their cards. Thus, half of the participants in the individual task goal condition had to adopt the same (i.e., congruent) action goal as the co-actor (e.g., if the co-actor selected the higher of their cards, participants had to select the higher of their cards too) while the other half had to adopt the opposite (i.e., incongruent) action goals as the co-actor (e.g., if the co-actor selected the higher of their cards, participants had to select the lower of their cards instead).

81

In contrast, in the joint task goal condition, participants were instructed to complement the card selection of the co-actor by selecting one of their cards that completed an ascending or a descending number sequences together with the card selected by the co-actor. In particular, half of the participants in the joint task goal condition were instructed to complement an *ascending* number sequence when the co-actor selected the *higher* of their cards and to complete a *descending* number sequence when the co-actor selected the *lower* of their cards. The other half of participants in the joint task goal condition was instead instructed to complete an *ascending* number sequence when the co-actor selected the *lower* of their cards (e.g. if the co-actor selects the clubs three card, participants had to select the clubs four card) and to complete a *descending* number sequence when the co-actor selected the *higher* of their cards (see Figure 4-1 for examples)[17].

Importantly, the instructions for the two groups in the joint task goal conditions ensured that one group had to select the higher (or lower) of their cards when the co-actor selected the higher (or the lower) of their cards, while the other group had to select the higher (or lower) of their cards when the co-actor selected the lower (or higher) of their cards. Thus, although not relevant for performing the task as instructed, participants in the joint task goal condition had to select one of their cards by *implicitly* adopting the same (i.e., congruent) or opposite (i.e., incongruent) action goal as the co-actor. As such, the individual and the joint task goal condition were procedurally and perceptually identical and only differed in terms of either explicitly emphasizing participants' *individual* action goals (selecting the higher or the lower of their cards) or rather emphasizing co-actors' *joint goal* (complementing the card selection of the co-actor to build ascending or a descending number sequences).[18] In addition, to highlight the joint

---

[17] Note that the two cards of the participant always showed the next higher and the next lower card compared to the card selected by the co-actor. This ensured that selecting one of the two cards completes an ascending card sequence while selecting the other card completes a descending number sequence.

[18] Note that, as such, participants could have principally performed the individual task goal condition by following the instructions of the joint task goal condition and vice versa. As

outcome of co-actors' individual card selections, only the selected card of the partner and the selected card of the participant remained on screen during feedback display in the joint task goal condition, while in the individual task goal conditions all four cards remained on screen until the end of the trial.

For all experimental groups, congruency between movement goals was manipulated by varying the relative spatial position of the assigned cards, so that selecting the correct card on a given trial required participants to select either the left or the right of their cards after observing their co-actor selecting the left or the right of their cards respectively.

Thus, the study had a 2 (Task Goal: Individual vs. Joint) × 2 (Action Goal Congruency: Congruent vs. Incongruent) × 2 (Movement Goal Congruency: Congruent vs. Incongruent) mixed factorial design, with Task Goal and Action Goal Congruency manipulated between and Movement Goal Congruency manipulated within subjects. After going through a self-paced instruction phase explaining the trial structure and the condition specific card selection rules, participants performed eight training trials that were randomly drawn from a set of forty-eight possible card arrangements (counterbalanced for Movement Goal Congruency). If participants responded correctly on more than four trials during training[19], they proceeded to the main part of the experiment and performed forty-eight test trials in their assigned experimental condition appearing in fully randomized order (Movement Goal Congruency was counterbalanced across trials).

Data analysis. Training trials were excluded from analysis. Test trials with response omissions were removed prior to analysis (0.9% of all trials). For the remaining test trials, error rates (ER, relative frequency of wrong card selections) and response times for correct responses

---

this design feature makes carry over effects between different blocks manipulating Task Goal within subjects highly likely, Task Goal was manipulated as a between-subjects factor.

[19] In both experiments reported, participants with inadequate training performance had to quit the experiment and were compensated with 1 GBP for their time spent on the study. The final sample size only included participants who completed the whole experiment.

(RTs) were analysed by means of separate 2 (Task Goal: Individual vs. Joint) × 2 (Action Goal Congruency: Congruent vs. Incongruent) × 2 (Movement Goal Congruency: Congruent vs. Incongruent) mixed ANOVAs with Task Goal and Action Goal Congruency as between-subjects factors and Movement Goal Congruency as within-subjects factor. For all inference statistical analysis, the alpha level was set to .05. Significant interactions were followed up by analysis of simple main effects that were Bonferroni corrected for multiple comparisons.

To supplement the results of our preregistered frequentist analysis, we also conducted an exploratory Bayesian ANOVA using JASP (JASP Team, 2024). This allowed us to calculate Bayes factors to assess the relative evidence in the data for different factor models. To quantify the evidence for individual main and interaction effects of the design factors, we report model-averaged inclusion Bayes factors ($BF_{Inclusion}$), which measure the strength of evidence for including a particular factor or interaction in a model compared to models where that factor or interaction is excluded (see van den Bergh et al., 2020). Additionally, we calculated Bayes factors for individual model comparisons, assessing the relative evidence for different combinations of factors in explaining participants' performance. Therefore, we report Bayes factors (BF_{01}) quantifying the relative evidence of individual factor models against the best fitting model of our design factors (cf. van den Bergh et al., 2020). We used the default specification of prior distributions for parameter values in each model offered in JASP (cf. van den Bergh et al., 2020) set to its default options (*r* scale for fixed effects = .5). To assess the robustness of the results, we repeated the analysis with two different prior specifications (*r* scale for fixed effects = .2 and *r* scale for fixed effects = .8). Prior probabilities for individual models were drawn from a uniform distribution.

Results

The results of the RT analysis of Experiment 1 is depicted in Figure 4-2.
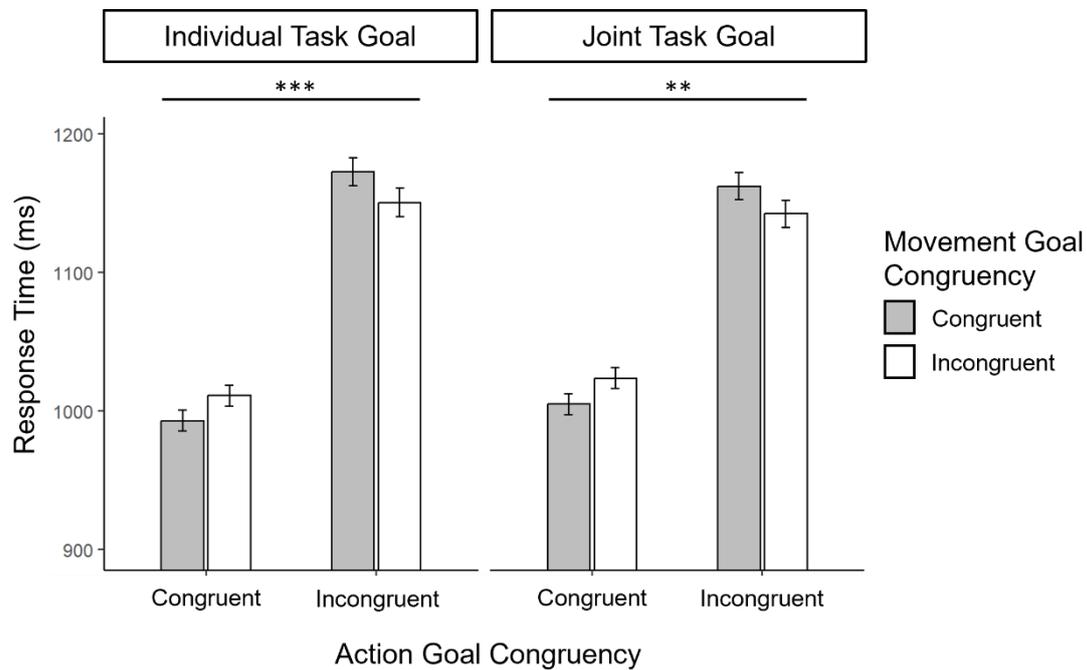
Figure 4-2. Mean response times in Experiment 1 as a function of Task Goal, Individual Goal Congruency and Response Congruency. Error bars represent standard errors of the mean (SEM) corrected for within subjects designs (Morey, 2008). Asterisks mark simple main effects between factor levels reaching significance after Bonferroni-correcting alpha levels for multiple comparisons. *** $p < .001$, ** $p < .01$.

The RT analysis revealed a significant main effect of Action Goal Congruency, $F(1, 188) = 23.1$, $p < .001$, $\eta_p^2 = .11$, $BF_{Inclusion} = 3413.9$, with faster RTs for participants in the action goal congruent (mean = 1008 ms, 95% CI [965 ms, 1050 ms]) compared to the action goal incongruent conditions (mean = 1156ms, 95% CI [1113 ms, 1200 ms]). Importantly, the interaction between Action Goal Congruency and Task Goal was non-significant ($F < 1$, $BF_{Inclusion} = 0.527$), thus providing no evidence that the main effect of Action Goal Congruency was moderated by participants' Task Goal (Individual vs. Joint). Yet, results showed a significant interaction between Action Goal Congruency and Movement Goal Congruency, $F(1, 188) = 10.1$, $p = .002$, $\eta_p^2 = .05$, $BF_{Inclusion} = 14.43$. Follow-up analysis showed that this interaction was driven by faster RTs in movement goal congruent trials (mean = 998 ms, 95%

CI [988 ms, 1009 ms][20]) compared to movement goal incongruent trials (mean = 1017 ms, 95%

CI [1007 ms, 1028 ms]) when co-actors' action goals were congruent to each other, $F(1,97) =$

6.17, $p_{Bonf} = .04$, $\eta_p^2 = .06$, and descriptively slower RTs in movement goal congruent trials

(mean = 1168 ms, 95% CI [1154 ms, 1182 ms]) compared to movement goal incongruent trials

(mean = 1147 ms, 95% CI [1133 ms, 1160 ms]) when co-actors' action goals where incongruent

to each other, $F(1,93) = 4.45$, $p_{Bonf} = .08$, $\eta_p^2 = .05$. The absence of a significant three-way

interaction between Action Goal Congruency, Movement Goal Congruency and Task Goal (F

< 1, $BF_{Inclusion} = 0.319$), showed no evidence that this interaction pattern was moderated by

participants' Task Goal (Individual vs. Joint). All other effects were non-significant (all $F < 1$).

In line with the frequentist analysis, Bayesian model comparison revealed that the factor

model including main effects of Action and Movement Goal Congruency as well as their

interaction provided the best fit to the data ($BF_M = 9.2$). Importantly, this model was preferred

over the model that further included the interaction between Task Goal and Action Goal

Congruency ($BF_{01} = 3.96$), quantifying substantial evidence against a moderating influence of

Task Goal on the effect of Action Goal Congruency. Similarly, the model including only main

effects for Action and Movement Goal Congruency as well as their interaction, was strongly

preferred over the model that included further interactions up to the three-way interaction

between all design factors ($BF_{01} = 91.89$), thus quantifying strong evidence that the interaction

pattern between Action and Movement Goal Congruency was not moderated by participants'

Task Goal. A table with all Bayesian model comparisons is presented in Appendix C1. Using

different prior specifications for parameter values in each model produced qualitatively similar

result patterns (result tables for Bayesian model comparisons with different prior options can

be found in the supplement).

---

[20] Confidence intervals for within-subjects contrasts are adjusted for within-subjects designs (Morey, 2008).

Results of the ER analysis were largely congruent with the RT analysis and there was no indication for a speed-accuracy trade-off. The detailed results of the ER analysis are presented in Appendix C1.

Discussion

The results of Experiment 1 revealed three main findings. First, we found an imitative congruency effect at the level of co-actors' action goals: Participants responded faster if their action goal (selecting the higher or the lower of their cards) was congruent (vs. incongruent) to the action goal of their co-actor. Second, we found a significant interaction between action and movement goal congruency: Congruent (vs. incongruent) movement goals (e.g. participants selecting their left card after the co-actor selected their left card) only led to faster responses when co-actors action goals were congruent but not when they were incongruent to each other. These two findings replicate those of Ondobaka et al. (2012) and support the notion that imitative response tendencies are primarily determined by congruency relations between co-actors' higher-level action goals which modulate the influence of congruency relations between co-actors' lower-level movement goals. Thus, our findings support goal-directed theories of imitation by showing that interpretations of observed actions in terms of their higher-level action goals modulate imitative congruency effects between lower-level representations of own and others' actions.

Third, however, we found no evidence that the effects of action and movement goal congruency were dependent on participants' task goal to either match or mismatch the card selection of the co-actor (individual task goal condition) or to complement the card selection of the co-actor to achieve a joint goal (joint task goal condition). More specifically, we found no indications that instructing participants to complement the card selection of their co-actor to achieve a joint outcome helped participants to overcome effects of imitative congruency between their own and their co-actor's individual action and movement goals. This finding

speaks against the hypothesis that joint task goals add to the hierarchical control of imitation by modulating imitative congruency effects between co-actors' individual action and/or movement goals.

Yet, it is possible that the missing influence of task goal instructions on imitative congruency effects in Experiment 1 could be explained by overall task difficulty in the joint task goal condition. In particular, we reasoned that participants in the joint task goal condition could have faced difficulties implementing the task instructions to build ascending and descending number sequences together with their co-actor and reverted to perform the task with the individual task goal of matching or mismatching the card selection of their co-actors instead. Reconceptualization of task goals in the face of increasing task difficulty is a central assumption of action identification theory which suggests that increasing difficulties to perform an action under a current task conceptualization leads people to conceptualize their task in alternative ways that ensure more efficient performance (Vallacher & Wegner, 1987, 2012).

Based on this speculation, we conducted a second experiment, closely replicating Experiment 1 with a different set of stimuli as action targets, that should reduce overall task difficulty and hence might facilitate participants' encoding of their joint goal in the joint task goal condition.

4.1.3. Experiment 2

Experiment 2 provided a close replication of Experiment 1 with a different set of target stimuli: Instead of playing cards with different numerical values, participants and their co-actor selected between line drawings that differed in their relative height. While Experiment 1 required participants to process abstract symbolic information to discriminate between action targets (numerical differences between symbolic card values), target selection in Experiment 2 required participants to process concrete perceptual information instead (perceptual height differences between line drawings). Thus, action selection in Experiment 2 could build on

modal representations of target stimuli which are thought to be processed more efficiently in contrast to amodal representations involved in the processing of more abstract symbolic information (Kaup et al., 2024).

Changing the modality of stimulus representations should thus enable more efficient processing of magnitude relations between action targets, lowering overall task difficulty in Experiment 2 compared to Experiment 1. We hypothesized that this might help participants in the joint task goal condition to select their targets in relation to the joint task goal of complementing their co-actors target selection, which could make possible modulations of joint task goals on imitative congruency effects more pronounced.

Methods

Transparency and openness. Data for Experiment 2 was collected in April 2023. Raw data, analysis scripts, stimuli, and experiment scripts are accessible in the Open Science Framework project repository referred to above. Hypotheses, study design, sampling and analysis plan of Experiment 2 were identical to Experiment 1, thus following the same preregistration protocol referred to above.

Participants. A new sample of 200 Prolific users took part in Experiment 2, randomly assigned to the same experimental conditions as in Experiment 1. From this sample, twelve participants met the preregistered exclusion criteria and were dropped from analysis, leading to a final sample size of $N = 188$ (mean age = 28.3, SD = 8.0; 36 % female, 64 % male).

Stimuli and apparatus. In Experiment 2, instead of playing cards, participants were presented with four horizontally aligned line drawings (see Figure 4-3). The line drawings were drawn from a set of eight line drawings with equal width but different relative heights. From the lowest to the highest line drawing of the stimulus set, line height increased by one unit equal to the size of the lowest line. From this set, two lines were assigned to the co-actor and were presented on the left half of the screen. They always depicted one out of six possible line

combinations from the stimulus set containing only lines from the second lowest to the second highest and only those line combinations with a height difference of at least two units equal to the height of the lowest line. The other two lines were assigned to the participants and were presented on the right half of the screen. Importantly, they always showed the next higher and the next lower line relative to the line selected by the co-actor on a given trial. Participants selected the left or the right of their assigned lines by pressing the "c" (left) and "m" (right) key on their keyboard respectively. Figure 4-3 depicts an example arrangement of the four line drawings on an individual trial.
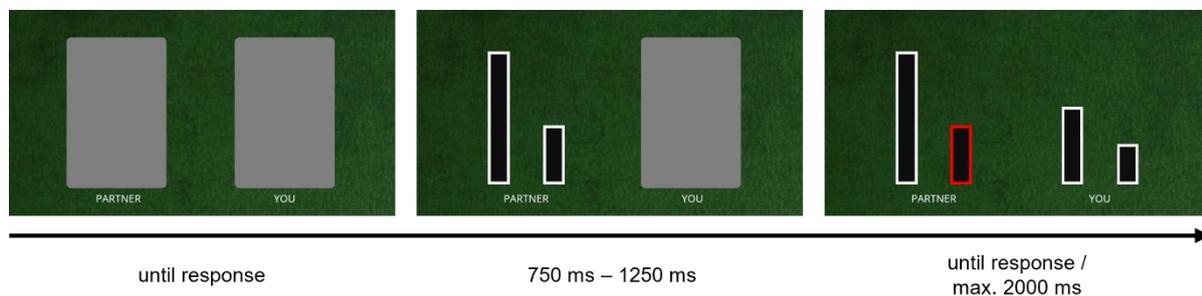
Figure 4-3. Trial timeline and stimuli example for Experiment 2. Task instructions for the different experimental conditions were analogous to those of Experiment 1.

Procedure and design. The procedure and design of Experiment 2 were identical to that of Experiment 1, despite the following differences. At the beginning of each trial, two grey boxes covered the lines assigned to the participant and their co-actor. After starting a trial by pressing the space bar, the left box disappeared and revealed the two lines of the co-actor. The co-actor's selection was indicated by a red frame surrounding one of their lines appearing 750 ms to 1250 ms after stimulus onset. Together with the selection of the co-actor, the right box disappeared and revealed the two lines assigned to the participant. Participants then had up to 2000 ms to select one of their lines. Timing of response feedback was identical to Experiment 1. The trial procedure is depicted in Figure 4-3.

Parallel to Experiment 1, the rule that specified which line participants had to select on a given trial differed between four experimental groups. In the individual task goal condition, one

group of participants was instructed to always "match", while another group of participants was instructed to always "mismatch" the line selection of the co-actor by selecting the higher or the lower of their lines. In the joint task goal condition on the other hand, one group of participants was instructed to select the line that produced an *ascending* (a descending) sequence with the line selected by the co-actor when the co-actor selected the *higher* (the lower) of their lines, while another group of participants was instructed to select the line that produced an *ascending* (a descending) sequence with the line selected by the co-actors when the co-actor selected the *lower* (the higher) of their lines. As in Experiment 1, in all four groups, selecting the correct line could require participants to adopt the same (congruent) or the opposite (incongruent) movement goal as the co-actor (i.e., selecting the left or the right of their lines). Thus Experiment 2 had the same 2 (Task Goal: Individual vs. Joint) × 2 (Action Goal Congruency: Congruent vs. Incongruent) × 2 (Movement Goal Congruency: Congruent vs. Incongruent) mixed factorial design as Experiment 1, with Task Goal and Action Goal Congruency manipulated between and Movement Goal Congruency manipulated within subjects.

Data analysis. The data analysis procedure was identical to Experiment 1. 1.2 % of all trials in the test phase were removed due to response omissions.

Results

The results of the RT analysis of Experiment 2 is depicted in Figure 4-4.
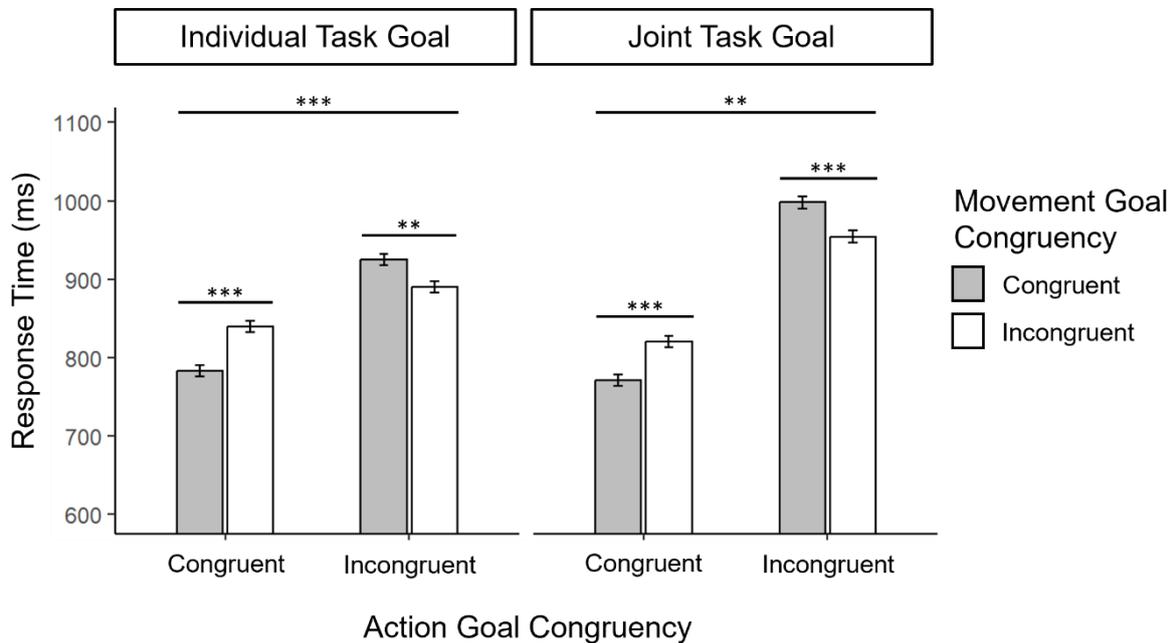
Figure 4-4. Response times in Experiment 2 as a function of Task, Inter-Individual Goal Congruency and Inter-Individual Response Congruency. Error bars represent standard errors of the mean (SEM) corrected for within subjects designs (Morey, 2008). Asterisks mark simple main effects between factor levels reaching significance after Bonferroni correcting alpha levels for multiple comparisons. *** $p < .001$, ** $p < .01$.

Analysis of participants' RTs revealed a significant main effect of Action Goal Congruency, $F(1, 184) = 33.7$, $p < .001$, $\eta_p^2 = .15$, $BF_{Inclusion} = 207687$, with faster RTs in the congruent action goal (mean = 803 ms, 95% CI [771 ms, 836 ms]) compared to the incongruent action goal conditions (mean = 942 ms, 95% CI [907 ms, 976 ms]). Yet, as in Experiment 1 the interaction between Action Goal Congruency and Task Goal was still off statistical significance, $F(1, 184) = 3.11$, $p = .08$, $\eta_p^2 = .02$, $BF_{Inclusion} = 0.81$, indicating no conclusive evidence for a moderating effect of Task Goal on the Action Goal Congruency effect. If anything, the RT difference between the action goal congruent and action goal incongruent conditions was descriptively larger in the Joint Task Goal condition (absolute mean difference = 181 ms, 95% CI [86.4 ms, 259 ms]) compared to the Individual Task Goal condition (absolute mean difference = 96.4 ms, 95% CI [2.31 ms, 191 ms]). Yet, similar to Experiment 1, results

of Experiment 2 showed a significant interaction effect between Action Goal Congruency and Movement Goal Congruency, $F(1, 184) = 76.9$, $p < .001$, $\eta_p^2 = .29$, $BF_{Inclusion} = 4.4 \times 10^{12}$. Follow-up analysis of simple main effects showed that this interaction was due to a significant effect of Movement Goal Congruency in the Action Goal Congruent conditions, $F(1,96) = 51.4$, $p_{Bonf} < .001$, $\eta_p^2 = .35$, with faster RTs in trials with congruent movement goals (mean = 777 ms, 95% CI [766 ms, 787 ms]) compared to incongruent movement goals (mean = 830 ms, 95% CI [819ms, 840ms]) and a significant but *reversed* effect of Movement Goal Congruency in the Action Goal Incongruent conditions, $F(1,90) = 27.9$, $p_{Bonf} < .001$, $\eta_p^2 = .24$, with slower RTs in trials with congruent movement goals (mean = 961 ms, 95% CI [951 ms, 971 ms]) compared to incongruent movement goals (mean = 922 ms, 95% CI [912 ms, 932 ms]). The interaction between Action and Movement Goal Congruency was not further qualified by Task Goal ($F < 1$, $BF_{Inclusion} = 0.278$ for the three-way interaction term).

In line with the frequentist analysis, Bayesian model comparison revealed that the factor model including main effects of Action and Movement Goal Congruency as well as their interaction provided the best fit to the data ($BF_M = 11.8$). This model was favoured over the model that further included the interaction between Task Goal and Action Goal Congruency ($BF_{01} = 1.91$), indicating weak evidence against a moderating influence of Task Goal on the Action Goal Congruency effect. The best fitting model was also favoured over the model that included all further interactions up to the three-way interaction between all three design factors ($BF_{01} = 25.8$), indicating strong evidence against the hypothesis that the interaction between Action and Movement Goal Congruency was further moderated by participants Task Goal. A table with all Bayesian model comparisons is presented in Appendix C2. Using different prior specifications for parameter values in each model produced qualitatively similar result patterns (result tables for Bayesian model comparisons with different prior options can be found in the supplement).

Results of the ER analysis were largely congruent with the RT analysis. Yet, besides a significant main effect of Action Goal Congruency, $F(1, 184) = 8.51$, $p = .004$, $\eta_p^2 = .044$, and a significant interaction between Action Goal Congruency and Movement Goal Congruency, $F(1, 184) = 16.6$, $p < .001$, $\eta_p^2 = .083$, we also found a significant main effect of Task Goal, $F(1, 184) = 4.56$, $p = .034$, $\eta_p^2 = .024$, which was driven by higher error rates in the joint (mean = 4.7 %, 95% CI [3.9 %, 5.5 %]) compared to the individual task goal condition (mean = 3.4 %, 95% CI [2.6 %, 4.3 %]). The detailed results of the ER analysis are presented in Appendix C2.

Comparison between experiments

To evaluate whether the novel stimuli set used in Experiment 2 had the intended effect to enhance overall task performance compared to Experiment 1, we compared participants' average RTs and ERs between Experiment 1 and Experiment 2 by means of independent samples t-tests. RTs were significantly faster in Experiment 2 (mean = 870 ms, 95% CI [852 ms, 889 ms]) compared to Experiment 1 (mean = 1081 ms, 95% CI [1058 ms, 1104 ms]), $t(387) = 10.1$, $p < .001$, d = 1.04. Comparing ERs between Experiment 1 (mean = 3.3 %, 95 % CI [2.8 %, 3.8 %]) and Experiment 2 (mean = 4.1 %, 95 % CI [3.5 %, 4.6 %]) showed no significant difference, $t(387) = 1.93$, $p = .054$, d = 0.20.

Discussion

The results of Experiment 2 replicated the findings of Experiment 1 in a modified version of our task that involved perceptual rather than conceptual discrimination of action targets. We again found evidence for an imitative congruency effect at the level of action goals, which showed to modulate imitative congruency effects at the level of co-actors' lower-level movement goals. Yet, although modifying the task compared to Experiment 1 enhanced overall task performance in Experiment 2, we still found no evidence for a modulatory effect of Joint vs. Individual Task Goal instructions on imitative congruency effects between co-actors'

individual action and/or movement goals. Interestingly though, participants performing the task with the instruction to complement the card selection of their partner (Joint Task Goal condition) conducted more response errors than participants performing the task with the instruction to match/mismatch the individual action goal of their co-actor (Individual Task Goal condition). This finding seems to indicate that implementing the task instructions for selecting a response in the joint task goal condition appeared to be more difficult than selecting a response based on the instructions of the individual task goal condition.

4.1.4. General Discussion

In two experiments, we investigated whether and how higher-level goals to which own and others' actions are individually or jointly directed modulate effects of imitative congruency between observed and self-executed contributions to a social interaction. More specifically, we asked 1) whether imitative congruency between co-actors' individual action goals modulate imitative congruency effects between co-actors' lower-level movement goals and 2) if acting towards joint as opposed to individual task goals ads to the control of imitative response tendencies in social interactions.

Results of both experiments showed evidence for an imitative congruency effect at the level of co-actors' action goals which was found to modulate the impact of imitative congruency between co-actors' lower-level movement goals. Yet, our findings provided no evidence that instructing participants to act towards joint rather than individual task goals modulated imitative congruency effects any further. In the following, we will discuss the implications of these findings in more detail.

Regarding our first question, our findings provide additional evidence that imitative congruency effects between own and others' lower-level movement goals become modulated by congruency relations between own and others' higher-level action goals. Thus, our findings replicate earlier evidence for an interaction between higher-level action and lower-level

movement goals in shaping imitative congruency effects in social interactions (Ondobaka et al., 2012, 2013). As such, our findings contrast with recent findings by Cole et al. (2018) and Janczyk et al. (2016), who found no indications for a modulation of movement goal congruency effects by co-actors' higher-level action goals. However, in their task setups effects of movement goal congruency can be explained by attentional mechanisms (social inhibition of return, see Cole et al., 2019), which was ruled out in the current study by presenting co-actors' action targets side-by-side rather than opposing each other. This design choice ensured that effects of movement goal congruency in the current study can be attributed to automatic imitation of observed movement goals rather than to low-level attentional mechanisms. The contrasting findings between the current study and those of Cole et al. (2018) and Janczyk et al. (2016), could thus be explained by differing mechanisms that produce movement goal congruency effects in the respective task setups. This suggests the conclusion that automatic imitation of observed movement goals but not lower-level attentional mechanisms (i.e., social inhibition of return) appear to be susceptible to modulations of co-actors' higher-level action goals (but see, Schmitz et al., 2023).

Noteworthy, we replicated the original finding by Ondobaka et al. (2012), despite using an online test setting with a virtual co-actor, manipulating action goal congruency between- rather than within-subjects, arranging action targets of co-actors' side-by-side rather than opposite to each other, and deploying a novel set of stimuli that allow for fast perceptual discrimination between action targets. Taken together, our findings thus generalize those of Ondobaka et al. (2012) and corroborate the hypothesis of an hierarchical control mechanism for imitative response tendencies in which higher-level representations of own and others' actions in terms of their individual action goals determine the influence of lower-level representations of own and others' actions in terms of their movement goals (Ondobaka & Bekkering, 2012, 2013).

Concerning mechanistic theories of imitation, our findings thus question the automaticity of people's tendency to imitate low-level movement features of observed actions (e.g., Heyes, 2011; Rizzolatti et al., 2001). More specifically, our findings speak against the notion that imitation effects rely on direct matching between low-level movement features of observed and self-executed actions, but suggest that low-level perception-action links become flexibly adapted to current task demands (cf., van Schie et al., 2008). Instead, our findings support goal-directed theories of imitation (Csibra, 2007; Wohlschläger et al., 2003) by showing that imitation effects are mainly determined by conceptual interpretations of own and others' actions in terms of their higher-level action goals rather than by features of the concrete movements people choose to reach them.

Turning to our second question, however, our study revealed no evidence that instructing participants to work towards joint rather than individual task goals further modulated imitative congruency effects emerging between co-actors' individual action and/or movement goals. In particular, instructing participants to produce joint action outcomes by complementing their co-actor target selection was not found to alter an evident tendency of participants to imitate the individual action goal of their co-actor, nor the modulating influence of action goal congruency on imitative congruency effects between co-actors' movement goals. This finding is noteworthy, as performing the task in the joint task goal condition as instructed did not require participants explicitly to adopt the same or opposite individual action goal as the co-actor (i.e., selecting the higher or the lower of their targets) but to complement their co-actors' target selection in the service of a joint goal instead (i.e., selecting the target that produces and ascending or descending sequence with the target elected by the co-actor). The finding of similar imitation effects between the individual and joint task goal condition thus seem to contrast with previous research reporting evidence for a modulating effect of joint task goals on imitative congruency effects between co-actors' individual action contributions (Clarke et al., 2019; Rocca et al., 2023; Sacheli et al., 2018). This raises the question, why imitative

congruency effects between co-actors' individual action and movement goals remained present in the current study even when participants were instructed to pursue a joint task goal together with their co-actor.

To explain modulations of imitative congruency effects by joint task goals, theories of joint action have suggested that representations of joint task goals lead co-actors to integrate separate representations of own and others' individual action contributions into *hierarchical joint action plans* that also encode complementary relations between co-actors' individual action contributions (Candidi et al., 2015; Marschner et al., 2024; Pesquita et al., 2018; Sacheli et al., 2018; Sebanz & Knoblich, 2021; Sinigaglia & Butterfill, 2022; Zapparoli et al., 2022). Our findings may thus indicate that participants in the joint task goal condition did not represent their own and their co-actors' actions as complementary contributions to a joint goal they were pursuing together but rather as separate action contributions guided by individual action goals of selecting the higher or the lower of their targets instead. Although staying tentatively, we reason that this could be explained by specific task characteristics of the current study, imbalanced task demands between the individual and joint task goal condition or by a more general bias to encode own and others' contributions to a joint action in terms of individual rather than joint goals.

Regarding task characteristics, one limitation of the current study is that our task involved computer-based interaction with a virtual co-actor in an online test setting. Yet, it is possible that adoption of joint task goals and integration of own and others' actions into hierarchical joint action plans guided by joint rather than individual goal representations requires real-life embodied interaction with human co-actors (cf., Ciardo et al., 2022; Sahaï et al., 2019 for empirical findings; but see Strasser, 2015, 2018 for discussion). Thus, although simulating an interactive task with a virtual, disembodied co-actor proved to produce significant imitation effects in the current study, it could have been insufficient to trigger further integration of own and others actions into hierarchical joint action plans.

98

Moreover, in the current task setup, participants were always required to adapt their actions to those of their co-actor but never vice versa. Arguably, this introduced a leader-follower hierarchy to the task that prevented a need for mutual adaptation and common coordination requirements between both co-actors. Yet, mutual adaptation, common coordination requirements and egalitarian role distributions have been identified as critical task characteristics that lead co-actors to develop a sense of acting together as a group (Bolt et al., 2016; Le Bars et al., 2020, 2022; Pacherie, 2014) or to feel truly engaged in a meaningful social interaction (Di Paolo & De Jaegher, 2012). It is thus possible that unidirectional coordination demands on site of the participants made them reluctant to adopt the instructed joint task goal of producing joint outcomes together with their co-actor and made them more likely to encode their own action contributions in terms of the individual task goal of matching or mismatching the card selection of their co-actor instead.

Finally, another explanation for the presence of imitative congruency effects between co-actors individual action and movement goals under both task goal instructions could be that implementation of the joint task goal instructions confronted participants with unnecessary high task demands that could be circumvented by adopting an individual task goal instead (i.e., matching or mismatching the action goal of the co-actor by selecting one's higher or one's lower target). As mentioned earlier, this interpretation would be supported by action identification theory (Vallacher & Wegner, 1985, 2012), proposing that people tend to conceptualize their actions in a given task context in a way that ensures most efficient task performance. Thus, participants in the joint task goal condition may have found it more easy to perform the task in an efficient way by discarding the joint task goal instruction of building ascending and descending target sequences and by adopting an individual task goal of matching or mismatching the target selection of their co-actors instead. This interpretation would also be in line with the finding of higher error rates in the joint compared to the individual task goal

condition in Experiment 2 which may reflect difficulties with initial or partial attempts to implement the instructions of the joint task goal conditions.[21]

This interpretation would suggest that representing own and others' contributions to a shared task as integrated complementary contributions towards a joint goal is not necessarily a default strategy for joint action planning but may be rather dependent on its utility for guiding efficient action performance in a given task context. Indeed, it has been proposed that joint action planning *minimally* requires co-actors to represent their joint goal as well as their own task contribution needed to achieve it while only being aware that the joint goal cannot be achieved by acting alone (Vesper et al., 2010). Such minimal joint action representations would thus not require that co-actors integrate representations of their own and their partners' individual action contributions into more elaborated hierarchical joint action plans that encode complementary relations between co-actors' individual task contributions. Yet, they might be often sufficient to guide efficient action performance in a given task. One possibility suggested by our findings could thus be that co-actors default to represent interactive tasks by means of action plans that remain limited on separate representations of their own and their partners' individual action contributions (i.e., individual action or movement goals), while additional efforts to further represent complementary relations between individual action contributions of self and others into hierarchical joint action plans may become only engaged when beneficial for current task performance. A promising direction for future research on joint action planning could thus be to identify conditions and enabling factors that provoke co-actors to switch from action plans centred at individual action and movement goals to more elaborated hierarchical

---

[21] As suggested by an anonymous reviewer, an alternative explanation of this finding could be that the joint task goal condition could have led to a parallel activation of joint and individual goal representations (Sartori et al., 2015), increasing cognitive demands in the joint compared to the individual task goal condition. Note that this interpretation would also be in line with our suggestion that participants may have coped with increased cognitive demands in the joint task goal condition by adopting an individual task goal instead.

representations of joint action that also encode complementary relations between co-actors' individual action and/or movement goals as well.

Conclusion

In summary, the present study provides complementing evidence that tendencies to imitate low-level movement features of observed actions are controlled by representations of own and others actions in terms of their higher-level individual action goals. Supporting goal-directed theories of imitation, this finding indicates that interpretations of observed actions in terms of their higher-level action goals can overrule tendencies to imitate others' low-level movement goas. Furthermore, our study indicates that instructions to pursue joint task goals to which observed and self-executed actions are collectively directed don't seem to overrule imitative congruency effects between co-actors' individual action and movement goals by default. This finding suggests that action planning in interactive task contexts might be dominated by action plans limited on co-actors' individual action and/or movement goals while further integration of own and others actions into hierarchical joint action plans that also encode complementary relations between them may become only engaged when afforded by current task demands.

4.2. STUDY IV: Automatic imitation of individual action goals in interactive task settings

4.2.1. Introduction

Action observation is essential to navigate our social environment. It underlies our ability to learn from others (Bandura, 1986), to imitate their behaviours (Iacoboni et al., 1999) and to coordinate our own actions with them (Sebanz & Knoblich, 2009). Understanding the cognitive processes that underlie the observation and interpretation of actions we observe in others is thus central to research on social cognition and social interaction. An important but unresolved question in these research areas concerns the role of the motor system in action observation (Csibra, 2007; Thompson et al., 2019; Wilson & Knoblich, 2005). Evidence shows that observing an action performed by another agent activates the observer's own motor system in a similar way as if the observer would perform the observed action themselves (for a review see Rizzolatti & Sinigaglia, 2016), which creates a tendency to imitate the actions we observe in others (Cracco et al., 2018; Heyes, 2011). Discord arises regarding appropriate explanations of this imitative action tendency and to what degree it is susceptible to contextual modulations. Informing this debate, the current study contrasted accounts of imitation focusing on low-level sensorimotor associations (e.g., Heyes, 2011) with accounts promoting a central role of higher-level action goals (e.g., Wohlschläger et al., 2003) to explain imitative action tendencies and link them to recent research investigating modulations of imitative action tendencies in joint action contexts (e.g., Clarke et al., 2019; McEllin et al., 2018; Sacheli et al., 2018).

Associative vs. goal-directed accounts of imitation

Theoretical explanations of imitative motor activation upon action observation can be classified into two opposing strands (c.f., Heyes, 2013). According to *associative accounts* of imitation (e.g., Brass & Heyes, 2005; Heyes, 2011; Iacoboni et al., 1999; Rizzolatti et al., 2001),

activation of corresponding motor representations upon action observation is the result of direct and stable excitatory connections between sensory and motoric representations of low-level movement features acquired through associative learning. Accordingly, these accounts propose that motor activation during action observation reflects an automatic and direct matching process between low-level sensory-motor representations that precedes any cognitively mediated interpretation of observed actions in terms of their goals.

In contrast, *goal-directed accounts* of imitation (Gattis et al., 2002; Wohlschläger et al., 2003) hold that activation of motor representations during action observation reflects predictive simulations of observed actions triggered by prior cognitive interpretations of the observed action in terms of their goals. These accounts assume that people first infer the most-likely goal of an observed action through cognitive processes outside the motor system. In a secondary step, the inferred goal is then thought to be forwarded to the observer's motor system that outputs a simulation of motor commands the observer would execute to achieve the extracted goal themselves. In many cases, the resulting activation of the observer's motor system is likely to match the motor commands the observed agent will perform to reach their goals, enabling accurate and timely predictions of how observed actions will unfold in space and time (Wilson & Knoblich, 2005; Wolpert et al., 2003).

Associative and goal-directed accounts of imitation both hold that people tend to imitate low-level movement features of observed actions in cases where the goal to which the action is most likely directed is closely linked to a particular pattern of motor commands, like in simple intransitive (i.e., non-object-directed) actions[22]. However, the two accounts make diverging

---

[22] Many studies on imitative response tendencies measure observers' responses to simple finger (Brass et al., 2000) or hand movements (Press et al., 2008) that lack a discernible purpose other than performing the movement per se. While such actions have been sometimes referred to as goal-less or non-goal directed actions (e.g., Chiavarino et al., 2013; Wild et al., 2010), we conceptualize such actions in the current paper as actions performed to reach a particular *movement goal*, describing physical features of the movements end state (e.g. "left finger up" or "hand open").

103

predictions concerning low-level movement imitation upon the observation of actions that are directed to higher-level goals that go beyond the mere realization of certain movement end states. Importantly, in these cases, the goal of an action can be represented at different, hierarchically nested levels of abstraction, ranging from low-level descriptions in terms of perceptual movement features and end states (e.g., reaching to the left; here referred to as *movement goals*), to higher-level descriptions in terms of the action's functional outcomes (e.g., picking one of two objects; here referred to as *action goals*) (c.f., Grafton & Hamilton, 2007; Ondobaka & Bekkering, 2012). Whereas the achievement of a movement goal is typically linked to a specific pattern of motor commands, the mapping between higher-level action goals and specific motor commands is less precise, as higher-level action goals can be usually achieved by a variety of motor commands that may be more or less appropriate in a given situation. Thus, in contrast to associative accounts, goal-directed accounts of imitation claim that motor system activation upon action observation does not necessarily correspond to the lower-level movement features of the observed action but should depend on the higher-level goals the action is thought to achieve.

Previous research has provided evidence supporting associative as well as goal-directed accounts of imitation. Support for associative accounts of imitation comes from studies on automatic imitation, showing that people tend to imitate observed movements even when they are irrelevant and interfere with one's own ongoing behavior, (e.g., Brass et al., 2000; Kilner et al., 2003; Stürmer et al., 2000; see Cracco et al., 2018 and Heyes, 2011 for review). These studies indicate that imitation of observed movements occurs involuntarily and automatically, which is in line with the idea of direct matching between observed and self-executed movement features. Furthermore, studies have shown that prolonged sensorimotor learning can modulate the size and direction of automatic imitation effects (Catmur et al., 2008; Catmur & Heyes, 2019; Cook et al., 2010; Heyes et al., 2005; Press et al., 2007). This indicates that involuntary tendencies to imitate observed movements are shaped by associative learning mechanisms.

104

Support for goal-directed accounts of imitation comes from studies assessing people's behavior when explicitly instructed to imitate others' actions. Here, studies have shown that when people are asked to imitate the action of another person, they are more likely to reproduce their higher-level action goal instead of the specific movements the person performed to reach it (Bekkering et al., 2000; Genschow et al., 2019; Leighton et al., 2010; Wohlschläger et al., 2003).

Interestingly, some studies further indicate that higher-level action goals also influences involuntary tendencies to imitate lower-level movement features of observed actions. In a seminal study, Ondobaka et al., (2012) found that people were only prone to imitate low-level movement features of another person when they were instructed to imitate their higher-level action goal, but not when being instructed to counter-imitate their higher-level action goal. In a recent replication of this study, we have further shown that instructions to counter-imitate another person's higher-level action goal can even reverse tendencies to imitate low-level movement features (Marschner et al., 2025). Moreover, in both studies, people had a general performance advantage when their task required them to adopt the same rather than a different action goal as their task partner. Supporting goal-directed accounts of imitation, these findings indicate a hierarchical interplay between imitation of higher-level action goals and imitation of lower-level movement features, where higher-level action goals control the impact and direction of low-level movement imitation (Ondobaka & Bekkering, 2012, 2013).

Imitative action tendencies in joint action contexts

The question of whether and how higher-level action goals shape imitative action tendencies becomes especially important in research on joint action (see McEllin et al., 2018 and Sebanz & Knoblich, 2021 for reviews). Joint actions refer to social interactions in which two or more people coordinate their actions in the service of a *joint goal*, i.e., social interactions structured around outcomes produced through coordinating action contributions of multiple co-

actors (Butterfill, 2012; Sacheli et al., 2015; Sebanz et al., 2006). Crucially, for many joint actions, imitative action tendencies become detrimental for smooth interpersonal coordination, as joint action often requires co-actors to coordinate dissimilar yet complimentary actions to achieve their joint goals (Sartori & Betti, 2015). Interestingly, a number of recent studies has found that involuntary imitation of observed movements is indeed attenuated (Clarke et al., 2019; Formica & Brass, 2024; Rocca et al., 2023; Schmitz et al., 2023) or even absent (Sacheli et al., 2018, 2019) when observed and self-executed actions are performed as interrelated contributions towards a joint action goal.

These findings are in line with goal-directed accounts of imitation because they indicate that involuntary tendencies to imitate observed movements become modulated when the observed action is interpreted as a contribution towards a higher-level goal to be achieved together with one's own action contribution. A possible explanation for this modulation is that representations of joint action goals lead co-actors to integrate separate representations of their own and their co-actors' individual action contributions into a unified joint action plan that imposes relational constraints between individual action contributions that are constitutive for the achievement of the joint goal (Clarke et al., 2019; Sacheli et al., 2018; Sinigaglia & Butterfill, 2022; Zapparoli et al., 2022). According to this idea, interpreting observed actions as contributions to a joint goal should trigger complementary action representations that satisfy relational constrains between individual task contributions of self and others.

The present research

Against this background, the current study set out to further investigate whether and how imitative action tendencies become modulated by higher-level action goals. Our study was motivated by two primary aims. First, we assessed whether higher-level action goals of an interaction partner affect imitative action tendencies despite being *irrelevant* for selecting one's own contributions to a shared task. Whereas previous studies showed that others' higher-level

action goals influence imitative action tendencies when people are *explicitly instructed* to imitate or counter-imitate the action goal of a task partner (Marschner et al., 2025; Ondobaka et al., 2012), evidence that imitative action tendencies can also become modulate by *task-irrelevant* action goals of others remains inconclusive so far. Whereas findings by Ondobaka et al. (2013) indicate that people tend to encode and imitate others' higher-level action goals despite not being required to do so, a recent replication attempt by Cole et al. (2018) found no effects of task-irrelevant action goals on imitative action tendencies.

Second, we aimed to further investigate the role of joint action goals in shaping imitative action tendencies in social interactions. Going beyond previous findings that representations of joint action goals can modulate involuntary tendencies to imitate lower-level movement features of observed actions (Clarke et al., 2019; Rocca et al., 2023; Sacheli et al., 2018), we asked whether joint action goals also shape imitative action tendencies that emerge between co-actors' individual-level action goals. Often, achieving a joint action goal (e.g., building a brick tower together with another person) does not only involve the coordination of different movements of both co-actors (e.g., grabbing and placing bricks at different locations and with different movement kinematics) but also the coordination of different higher-level action goals each co-actor achieves individually to contribute to the realization of the overarching joint action goal (e.g., selecting bricks of a different size or shape as one's co-actor). Thus, joint action goals add another layer to the hierarchical representation of own and others' actions. In this way, they may also shape how higher-level representations of own and others' actions in terms of their individual-level action goals influence each other.

To address these questions, we conducted three experiments using variations of Ondobaka et al.'s (2013) sequential card selection task. In all experiments, participants were paired with a virtual partner (in the following referred to as participant's co-actor) and selected one of two playing cards based on an imperative cue after observing their co-actor selecting one of two playing cards themselves (see Figure 4-5).

107
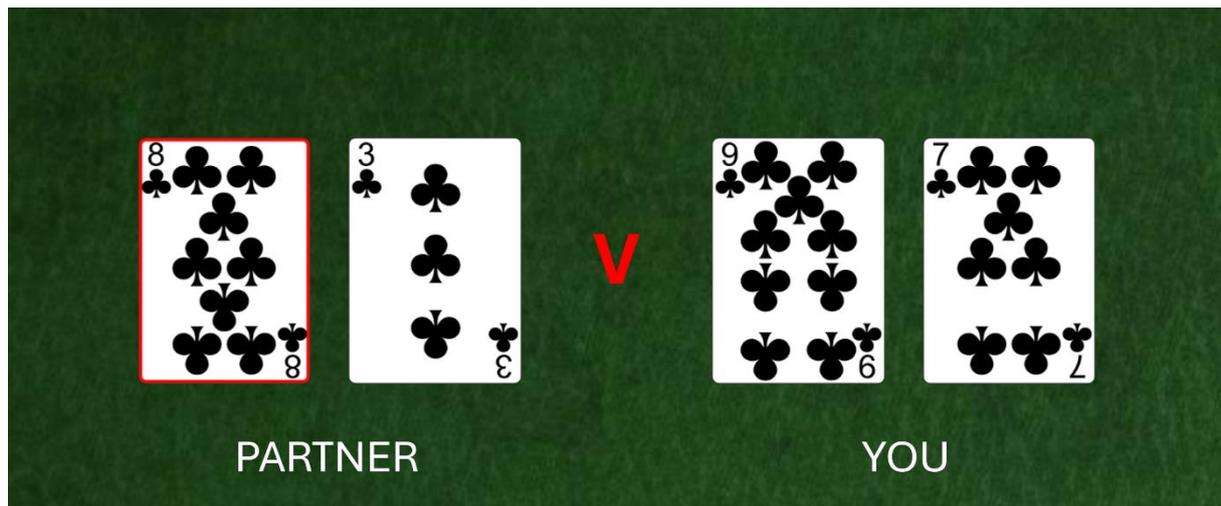
Figure 4-5. Illustration of the experimental task and setup. Participants viewd four horizontally aligned playing cards differing in their numerical value. The two left cards were assigned to a virtual partner, the two right cards were assigned to the participant. At the beginning of a trial, participants first observed their virtual partner selecting one of their cards (indicated by the red frame) and where then required to select one of their own cards shortly after. The card to be selected by participants was instructed by an instructive cue presented in the centre of the screen. In the individual goal condition, the cue instructed participants to select either the higher (Λ) or the lower (V) of their cards (e.g., selecting the clubs seven as the smaller card relative to the clubs nine in the given example). In the joint goal condition the cue instructed participants to select the card that completed an ascending (Λ) or a descending (V) number sequence together with the card selected by their virtual partner (e.g., selecting the clubs seven as a predecessor of the clubs eight in the given example). Participant selected one of their cards by pressing a left (C) or a right (M) key on their keyboard corresponding to the location of the selected card. Crucially, in both the individual and the joint goal condition, selecting the correct card required participants implicitly to adopt either the same or the opposite action goal as their virtual partner (e.g., selecting one's lower card after observing the partner selecting their higher card in the given example). At the same time, selecting the correct card also meant to adopt either the same or the opposite movement goal as their virtual partner (e.g., selecting one's right card, after observing the partner selecting their left card in the given example).

To assess participants' imitative action tendencies, we manipulated whether selecting the instructed card on a given round required participants to adopt either the same or the opposite action goal as their co-actor (selecting the higher or the lower of their cards) and whether that meant to adopt either the same or the opposite movement goal as their co-actor (selecting the left or the right of their cards). Crucially though, to select their own contribution to the task, participants were not required to encode their co-actor's action goal nor their co-actor's movement goal, rendering them both irrelevant for task performance. Instead, participants' task contributions were instructed by an imperative cue. This allowed us to test whether participants

would nevertheless show an *involuntary* tendency to imitate their co-actor's action goal and/or their co-actor's movement goal and how both tendencies might interact with each other. Following associative accounts of imitation, participants should be only affected by imitative congruency between their own and their co-actor's low-level movement goals. In contrast, following goal-directed accounts of imitation, participants should be affected by imitative congruency between their own and their co-actor's higher-level action goals instead, which should modulate participants' involuntary tendencies to imitate their co-actor's lower-level movement goals.

To test whether joint action goals modulate imitative action tendencies, we manipulated the task instructions with regard to the interpretation of the imperative cue. One group of participants was instructed to pursue an individual action goal. Here participants responded to the imperative cue by selecting either the higher (Λ) or the lower (V) of their cards. In contrast, another group of participants was instructed to pursue a joint action goal together with their co-actor. Here, participants responded to the imperative cue by selecting one of their cards that completed either an ascending (Λ) or a descending (V) card sequence together with the card selected by their co-actor before. In contrast to participants receiving individual action goal instructions, participants receiving joint goal instructions were not required to encode their own task contribution in relation to an individual action goal of selecting the higher or the lower of their cards. Instead, they selected their cards in relation to the card selected by their partner to produce a joint outcome together that reflected specific relations between the outcomes of their individual task contributions. We hypothesized that participants receiving instructions to pursue a joint action goal with their co-actor, should be less pronounced to imitate their co-actor's individual-level action goal compared to participants receiving instructions to pursue an individual action goal alone.

Finally, across experiments, we varied *when* the instructive cue was presented. In Experiment 1, the instructive cue was presented *after* participants observed their co-actor

109

selecting one of their cards. In Experiment 2 and Experiment 3, the instructive cue was presented *before* the card selection of the co-actor. This allowed us to test whether possible imitation effects would be modulated by knowledge of individual or joint action goals at the time of observing their co-actor's actions.

4.2.2. Methods

Participants

In total, three samples consisting of 60 (Experiment 1), 60 (Experiment 2) and 214 (Experiment 3) participants took part in the experiments. Our target sample size for all experiments were determined based on a priori power analysis.

Experiment 1 and 2 were powered to detect an ordinal interaction effect of size $f = 0.2$ in a 2 x 2 mixed ANOVA with one between and one within-subjects factor with 80% power at an alpha level of .05. Conducting an a priori power analysis in G*Power with these input parameters, we calculated a required sample size of $N = 52$ participants for Experiment 1 and Experiment 2 respectively[23]. Expecting possible exclusion of participants due to our preregistered criteria, we decided to collect data from 60 participants for Experiment 1 and Experiment 2 respectively. Experiment 3 was powered to replicate an unexpected three-way interaction found in Experiment 2 with adequate follow-up tests (see preregistration protocol for details).

Applying our preregistered exclusion criteria (see below), $N = 59$ participants were retained in Experiment 1 (mean age = 36.1, SD = 11.2; 40 % female, 60 % male), $N = 60$ in Experiment 2 (mean age = 38.2, SD = 10.7; 39 % female, 61 % male) and $N = 209$ in Experiment

---

[23] A subsequent sensitivity analysis showed that the estimated sample size for each between-subject group in Experiment 1 and 2 (i.e., 26 participants per group in each experiment) would also be sufficient to detect main effects and interaction contrasts of size $f > 0.57$ in a 2 x 2 repeated measures ANOVA with two within-subjects factors. This showed that Experiment 1 and 2 should have been sufficiently powered to replicate earlier findings by Ondobaka et al. (2013), reporting effect sizes of $f > 0.6$ for these effects

3 (mean age = 36.9, SD = 11.5; 41 % female, 59 % male). All participants had normal or corrected-to-normal vision and hearing and reported to speak English as their first language. All participants gave written informed consent and received monetary compensation for their participation (2,25 GBP for a total duration of approx. 15 minutes). Ethical approval for the study was granted by an independent research ethics board.

Stimuli and apparatus

Participants completed the respective experiment online in a browser on their own personal computer. Experiment files and data were stored on a JATOS (Just Another Tool For Online Studies; Lange et al., 2015) server. Stimuli presentation and response recording was controlled by a custom made script written in the jsPsych JavaScript framework (de Leeuw et al., 2023).

In all experiments, participants were presented with four horizontally aligned playing cards, displayed against a green background (see Figure 4-5). The cards were drawn from a subset of a standard card deck comprising the numerical card values of the clubs' suit in the range from two to nine. Two of the cards were assigned to the virtual the co-actor and were presented on the left half of the screen. They always showed one out of six possible card combinations from the card set, containing card values in the range from three to eight with a numerical distance between cards larger than two (i.e., [3|6], [3|7], [3|8], [4|7], [4|8], [5|8]). The other two cards were assigned to the participants and were presented on the right half of the screen. Importantly, they always showed the next higher and the next lower card value relative to the card selected by the co-actor on a given trial (e.g., if the co-actor selects clubs' eight, the two cards of the participants were clubs' seven and clubs' nine). Which of their cards participants had to select on a given round was instructed by one of two possible cues (V or Λ) printed in red, centrally presented between the two cards of the co-actor and the two cards assigned to participants. Participants selected one of their assigned cards by pressing the "C"

key for their left and the "M" key for their right card on their keyboard respectively. Figure 4-5 depicts an example arrangement of the four playing cards on an individual trial.

Procedure and design
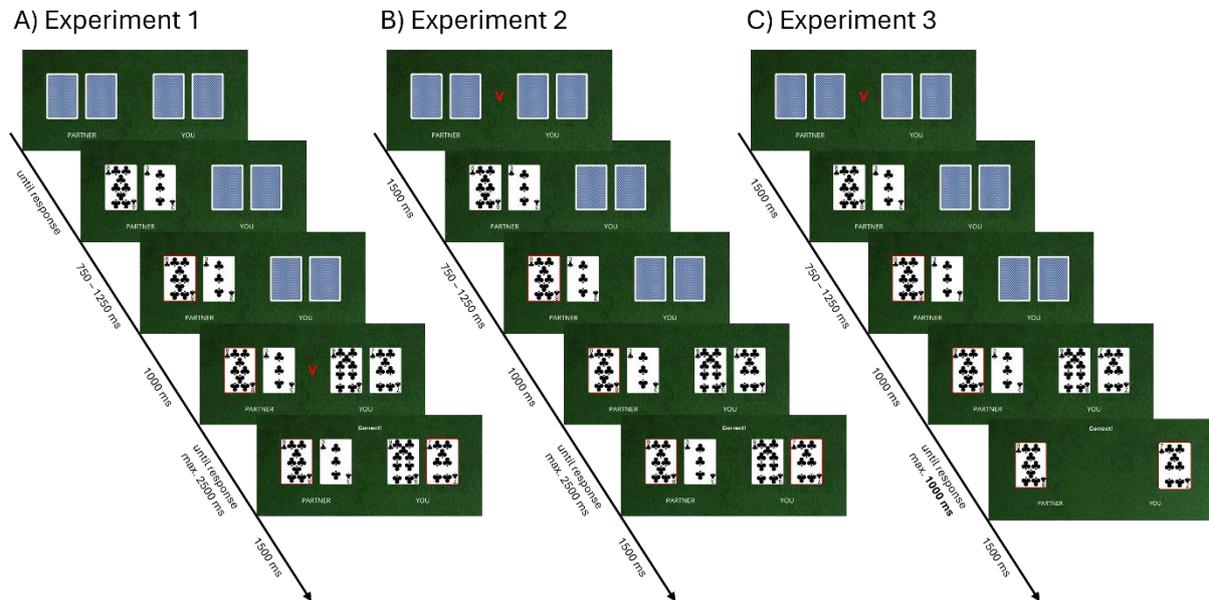
Figure 4-6. Trial timeline in A) Experiment 1, B) Experiment 2, and C) Experiment 3.

The trial time line for Experiment 1, 2, and 3 is depicted in Figure 4-6. In all experiments, each trial started with the presentation of all four playing cards faced down and participants started each trial manually by pressing their space bar. In Experiment 1, pressing the space bar first revealed the two cards of the co-actor, presented for a variable delay between 750 and 1250 ms, which was followed by one of the two cards being highlighted with a red frame, reflecting the co-actor's card selection. After a delay of 1000 ms, the two cards of participants were turned over, together with the presentation of the red cue, instructing participants which of their two cards to select on the given trial. Participants had up to 2500 ms to select one of their cards. If participants responded within this window, their selected card was highlighted with a red frame as well, followed after a delay of 500 ms by written feedback ("correct", "error", or "too slow" if participants failed to respond within the response window) displayed for 1500 ms until the end of the trial. The trial time line for Experiment 2 was identical to Experiment 1, despite the

imperative cue being displayed only at the start of each trial (after participants pressed the space bar) for 1500 ms, with all four cards remaining faced down. The trial time line of Experiment 3 followed that of Experiment 2 with a shortened response window of 1000 ms to enforce faster responding. Additionally, during feedback presentation in the joint goal condition of Experiment 3, only the selected card of the co-actor and the one selected by participants remained on screen to highlight the joint outcome of their separate task contributions.

In all experiments participants were randomly allocated to one of two task conditions that differed with respect to whether the imperative cue instructed participants to select one of their cards to achieve an individual action goal (individual goal condition) or to achieve a joint action goal together with their co-actor (joint goal condition). In the individual goal condition, participants were instructed to select the higher (Λ) or the lower (V) of their cards. In the joint goal condition, participants were instructed to select the card that completed an ascending (Λ) or a descending (V) sequence together with the card selected by their co-actor. Note that as the two cards assigned to the participant always represented the two neighboring cards of the card selected by the co-actor, completing an ascending or descending card sequence in the joint goal condition was equivalent to selecting the higher or lower card in the individual goal condition respectively. Thus, the two task conditions were procedurally and perceptually completely identical, differing only in framing participants card selection as being directed towards an outcome to be achieved alone (higher/lower card) or towards an outcome to be achieved together with their co-actor (ascending/descending card sequence).

After receiving the condition specific task instructions, participants performed a total number of 96 trials, split into six sub-blocks with 16 trials each[24]. Within each sub-block, imperative cues were counterbalanced and half of the trials reflected action goal congruent (e.g.,

---

[24] Proceeding the experiment after the first sub-block required that participants responded correctly in at least 10/16 trials. Participants who failed to reach this performance threshold had to quit the experiment after the first sub-block and got replaced right away.

selecting the higher of one's cards after observing the co-actor selecting the higher of their cards) and the other half action goal incongruent trials (e.g., selecting the higher of one's cards after observing the co-actor selecting the lower of their cards), of which each half reflected movement goal congruent (e.g., selecting the left of one's cards after observing the co-actor selecting the left of their cards) and movement goal incongruent trials (e.g., selecting the left of one's cards after observing the co-actor selecting the right of their cards) respectively. Trial order within each sub-block was randomized. Thus all experiments had a 2 Task Condition (Individual Goal vs. Joint Goal) x 2 Action Goal Congruency (Congruent vs. Incongruent) x 2 Movement Goal Congruency (Congruent vs. Incongruent), mixed factorial design with Task Condition manipulated between and Action Goal Congruency and Movement Goal Congruency manipulated within participants. Each experiment lasted about 15 minutes.

Data analysis

In each experiment, the first 16 trials served as training trials and were excluded from analysis. For the remaining trials, trials with response omissions were removed from analysis (1.44 % of all trials in Experiment 1; 0.35 % of all trials in Experiment 2; 4.54 % of all trials in Experiment 3). Participants with an overall error rate (ER, relative frequency of wrong card selections) larger than 20 % were excluded from the sample (effecting one participant in Experiment 1, none in Experiment 2, and five participants in Experiment 3). For the remaining sample, trials with a response time (RT) that deviated more than three standard deviations from participants' individual mean RT were also excluded (0.97 % of remaining trials in Experiment 1; 1.92 % of remaining trials in Experiment 2; 1.22 % of remaining trials in Experiment 3). For the remaining trials, participants' ERs and mean RTs on correct trials were analyzed by separate 2 (Task Condition: Individual Goal vs. Joint Goal) × 2 (Action Goal Congruency: Congruent vs. Incongruent) × 2 (Movement Goal Congruency: Congruent vs. Incongruent) mixed ANOVA with Task Condition as a between-subjects factor and Action and Movement Goal

Congruency as within-subjects factors. For all inference statistical analysis, the alpha level was set to .05. Significant interactions were followed up by analysis of simple main effects that were Bonferroni corrected for multiple comparisons.

Transparency and openness

Data for Experiment 1 and Experiment 2 were collected in February 2025, data for Experiment 3 in Mai 2025. Data frames, analysis scripts, stimuli and experiment scripts for all experiments are accessible in an Open Science Framework project repository (https://osf.io/wc96q/?view_only=dd5091015d744e81a0295afdb0050d49). The study design, hypotheses, sampling plan and data exclusion criteria of all experiments were preregistered in the Open Science Framework. The preregistration is accessible under https://osf.io/swhpu/?view_only=2e61ecd807a642cc9540bf254f847148.

4.2.3. Results and discussion for Experiment 1

Results

The main findings of Experiment 1 are depicted in Figure 4-7.
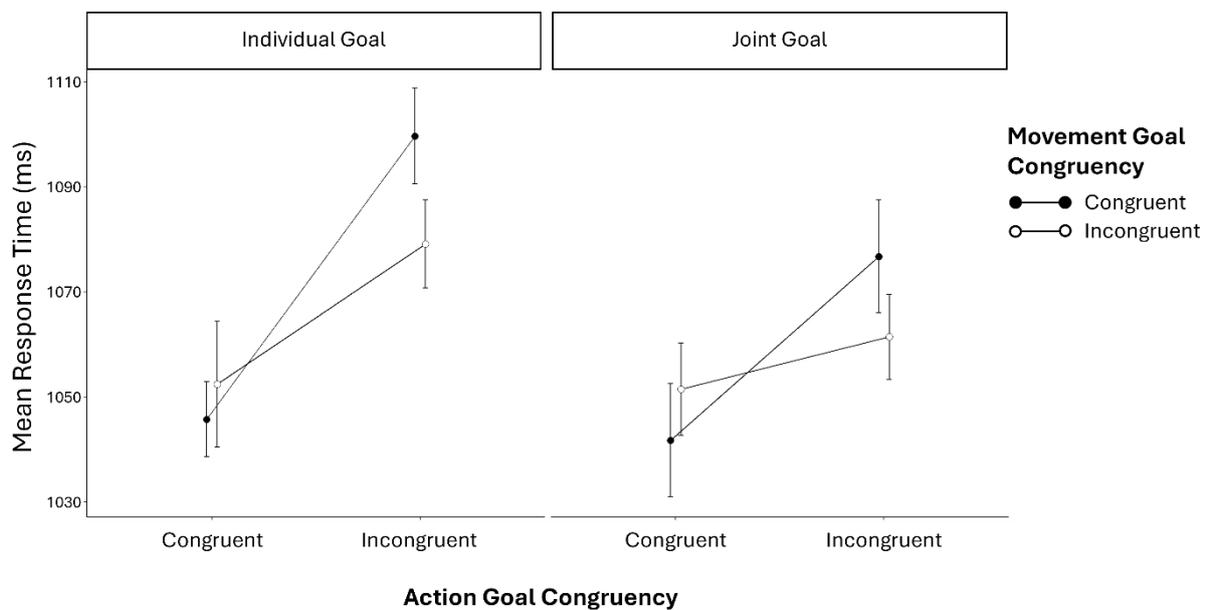

Figure 4-7. Mean RT in Experiment 1 as a function of Task Condition, Action Goal Congruency and Movement Goal Congruency. Error bars represented standard errors of the mean corrected for within-subjects designs.

The RT analysis revealed a significant main effect of Action Goal Congruency, $F(1,57)$ = 15.5, $p < .001$, $\eta_p^2 = .213$, showing that participants selected their cards faster on action goal congruent trials (M = 1048 ms, 95 % CI = [984, 1112]) compared to action goal incongruent trials (M = 1079 ms, 95 % CI = [1013, 1145]). Yet, we found no evidence that the main effect of Action Goal Congruency was further moderated by Task Condition, as indicated by a non-significant interaction between Action Goal Congruency and Task Condition, $F(1,57) = 1.24$, $p = .269$, $\eta_p^2 = .021$. The main effects of Task Condition and Movement Goal Congruency were both non-significant (both $F < 1$). Yet, the analysis revealed a significant interaction between Action Goal Congruency and Movement Goal Congruency, $F(1,57) = 5.53$, $p = .022$, $\eta_p^2 = .089$. Analyzing simple main effects showed that the interactions was driven by descriptively faster RTs on movement goal congruent trials (M = 1044, 95 % CI = [980, 1108]) compared to movement goal incongruent trials (M = 1052, 95 % CI = [987, 1117]) when participants action goal was congruent to that of their partner, $t(57)$ = -0.91, $p_{corrected}$ = .736, while participants responded descriptively faster on movement goal incongruent trials (M = 1070, 95 % CI = [1003, 1137]) compared to movement goal congruent trials (M = 1088, 95 % CI = [1022, 1154]) when their action goal was incongruent to that of their partner, $t(57)$ = 2.28, $p_{corrected}$ = .053.

The ER analysis revealed no significant main effects (all $F < 1$). Yet, similar to the RT analysis we found a significant interaction between Action Goal Congruency and Movement Goal Congruey, $F(1,57) = 4.58$, $p = .037$, $\eta_p^2 = .074$. No other interaction effect reached significance (all $F[1,57] < 3.31$; see Appendix D for the full result table). Follow-up analysis of simple main effects showed that the interaction between Action Goal Congruency and Movement Goal Congruency was driven by descriptively lower ER on movement goal congruent trials (M = 2.53 %, 95 % CI = [1.36, 3.69]) compared to movement goal incongruent trials (M = 3.22 %, 95 % CI = [2.00, 4.44]) when participants action goal was congruent to that of their partner, $t(57)$ = -1.00, $p_{corrected}$ = .647, while participants made descriptively fewer response errors on movement goal incongruent trials (M = 2.58 %, 95 % CI = [1.27, 3.88])

116

compared to movement goal congruent trials ($M = 3.74$ %, 95 % CI = [2.26, 5.23]) when their action goal was incongruent to that of their partner, $t(57) = 1.63$, $p_{corrected} = .219$.

Discussion

Experiment 1 revealed three main findings. First, we found an imitative congruency effect between co-actors' individual action goals: Participants selected their cards faster when imitating rather counter-imitating their partner's action goal of selecting the higher or the lower of their cards. Notably, this effect was present although encoding their task partner's individual action goal was irrelevant for selecting participant's own contribution to the task. This finding suggests that participants encoded the action goal of their task partner automatically and showed a behavioral tendency to adopt the same action goal for their own contribution to the task.

Second, we found that imitative congruency between co-actors' individual action goals modulated effects of imitative congruency relations between co-actors' movement goals: On action goal congruent trials, participants selected their cards descriptively faster and less error prone when *imitating* the movement goal of their task partner, while on action goal incongruent trials participants selected their cards descriptively faster and less error prone when *counter-imitating* the movement goal of their task partner. This interaction pattern replicates previous findings (Marschner et al., 2025; Ondobaka et al., 2012, 2013) and support the idea of an hierarchical control mechanisms for imitative response tendencies in which representations of own and others' actions at higher conceptual levels (i.e., action goals) shape the influence of lower-level sensorimotor representations of own and others' actions (i.e., movement goals) (Ondobaka & Bekkering, 2012, 2013).

Third, however, Experiment 1 revealed no evidence that participants' imitative response tendencies were further modulated by task instructions to adopt an *individual* action goal (to select the higher or the lower of their cards) or a *joint* action goal (to select a card that complements the card selection of their co-actor). Instead, imitative congruency between co-

117

actors' individual action goals seemed to affect task performance in the individual and the joint task goal condition in comparable ways, although participants in the joint goal condition were not instructed to encode their own task contribution in terms of an individual action goal of selecting the higher or the lower of their cards. This finding suggests that instructions to pursue joint rather than individual action goals were unable to eliminate participants' automatic tendency to encode their task partner's individual action goal and to deploy it for their own subsequent task contribution.

To follow up on these initial findings, in Experiment 2 we investigated whether the timing of forming an individual or joint action goal respectively impacts imitative response tendencies in the current task setup. Therefore, we presented the cue instructing participants which of their cards to select (i.e., the higher or the lower of their cards in the individual goal condition or the card that complements an ascending or descending number sequence in the joint goal condition respectively) at the start of each trial, i.e., before participants observed the card selection of their virtual partner. We reasoned that joint action goals may only become able to modulate imitative response tendencies if participants can observe the action of their task partner while already knowing the joint goal they are trying to achieve together.

4.2.4. Results and discussion for Experiment 2

Results

The main findings of Experiment 2 are depicted in Figure 4-8.
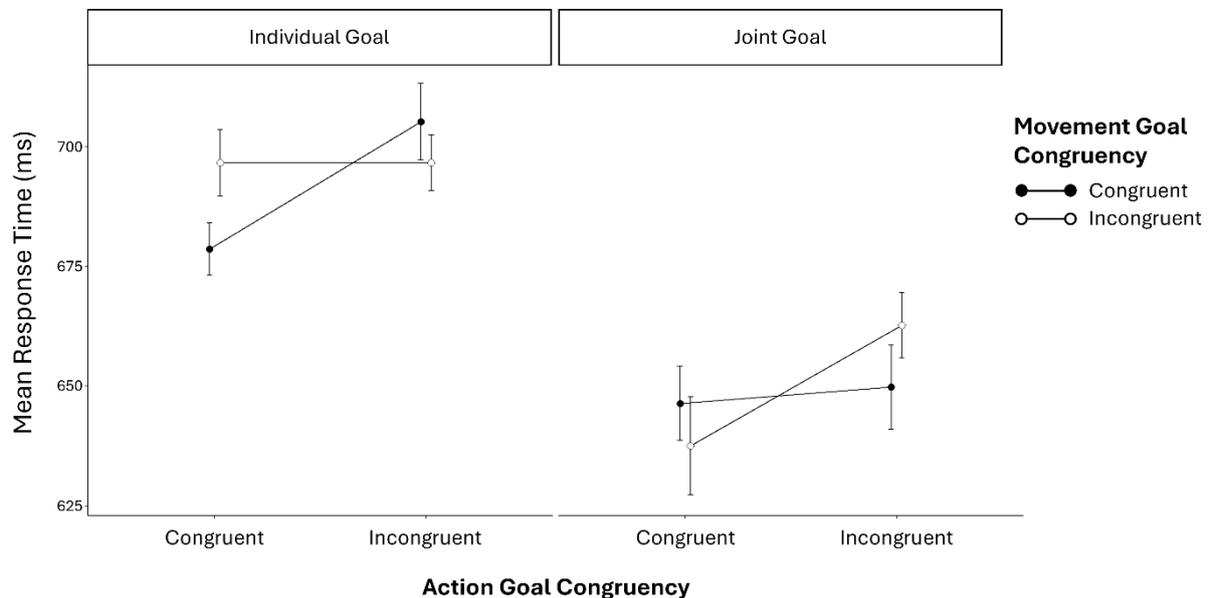
Figure 4-8. Mean RT in Experiment 2 as a function of Task Condition, Action Goal Congruency and Movement Goal Congruency. Error bars represented standard errors of the mean corrected for within-subjects designs.

The ER analysis revealed no significant effects (all $F[1,58] < 1.49$; see Appendix D for full result table). Though, similar to Experiment 1, the RT analysis revealed a significant main effect of Action Goal Congruency, $F(1,58) = 6.90$, $p = .012$, $\eta_p^2 = .106$, showing that participants selected their cards faster on action goal congruent trials ($M = 665$ ms, 95 % CI = [616, 714]) compared to action goal incongruent trials ($M = 679$ ms, 95 % CI = [630, 728]). Yet, as before we found no evidence that the main effect of Action Goal Congruency was further moderated by Task Condition, as indicated by a non-significant interaction between Action Goal Congruency and Task Condition ($F < 1$). The main effect of Task Condition was also non-significant ($F < 1$), although examining Figure 4-8 shows that participants responded descriptively faster in the joint goal condition ($M = 649$ ms, 95 % CI = [580, 718]) compared to the individual goal condition ($M = 694$ ms, 95 % CI = [625, 763]). The main effect of Movement Goal Congruency was also non-significant, as where all two-way interactions (all F

< 1). Yet unexpectedly, we found a significant three-way interaction between Task Condition, Action Goal Congruency and Movement Goal Congruency, $F(1,58) = 6.01$, $p = .017$, $\eta_p^2 = .094$. To follow up on the three-way interaction, we analyzed RTs from participants in the individual and the joint goal condition separately by subjecting RTs in each task condition to separate 2 (Action Goal Congruency: Congruent vs. Incongruent) x 2 (Movement Goal Congruency: Congruent vs. Incongruent) repeated measures ANOVAs. However, both analysis showed no significant effects. Main effects of Action Goal Congruency were non-significant in the individual goal condition, $F(1,29) = 3.75$, $p_{\text{corrected}} = .126$, $\eta_p^2 = .114$, as well as in the joint goal condition, $F(1,29) = 3.24$, $p_{\text{corrected}} = .165$, $\eta_p^2 = .101$. Main effects of Movement Goal Congruency were also non-significant in both analysis (both $F < 1$). Similarly, the two-way interactions between Action Goal Congruency and Movement Goal Congruency were non-significant in the individual goal condition, $F(1,29) = 3.37$, $p_{\text{corrected}} = .154$, $\eta_p^2 = .104$, as well as in the joint goal condition, $F(1,29) = 2.65$, $p_{\text{corrected}} = .229$, $\eta_p^2 = .084$.

Discussion

Similar to Experiment 1, Experiment 2 revealed an imitative congruency effect between co-actors' individual action goals: Participants selected their cards faster when imitating rather than counter-imitating the action goal of their task partner. Furthermore, we still found no evidence that this effect was modulated by task instructions to pursue joint rather than individual action goals, although this time participants observed their task partner's action while knowing which joint goal to pursue together. Interestingly though, the result pattern of Experiment 2 suggested that participants performed the task differently when being instructed to pursue an individual compared to a joint action goal: We observed a significant three-way interaction between Task Condition, Action goal Congruency and Movement Goal Congruency, indicating that the interaction pattern between Action Goal Congruency and Movement Goal Congruency differed between the individual and the joint goal condition.

However, our follow-up analysis remained inconclusive to draw reliable conclusions about the source of these differences.

Therefore, we decided to run a replication of Experiment 2 that was sufficiently powered to detect significant two-way interactions between Action Goal Congruency and Movement Goal Congruency in the individual and joint goal condition respectively (see preregistration protocol for details). Besides increasing the sample size compared to Experiment 2, we also adapted two procedural details in Experiment 3 that we thought would elevate chances of finding significant contrasts in task performance between the individual and the joint goal condition. First, we shortened the response window to 1000 ms to foster faster responding across participants. We did so with the suspicion that relatively lenient constraints on response speed in the previous two experiments may have overshadowed descriptive performance differences between the individual and the joint goal condition. Second, during feedback presentation in the joint goal condition, we now presented only the two cards selected by the virtual partner and participants respectively to highlight the joint outcome of their individual task contributions.

4.2.5. Results and discussion for Experiment 3

Results

The main findings of Experiment 3 are depicted in Figure 4-9.
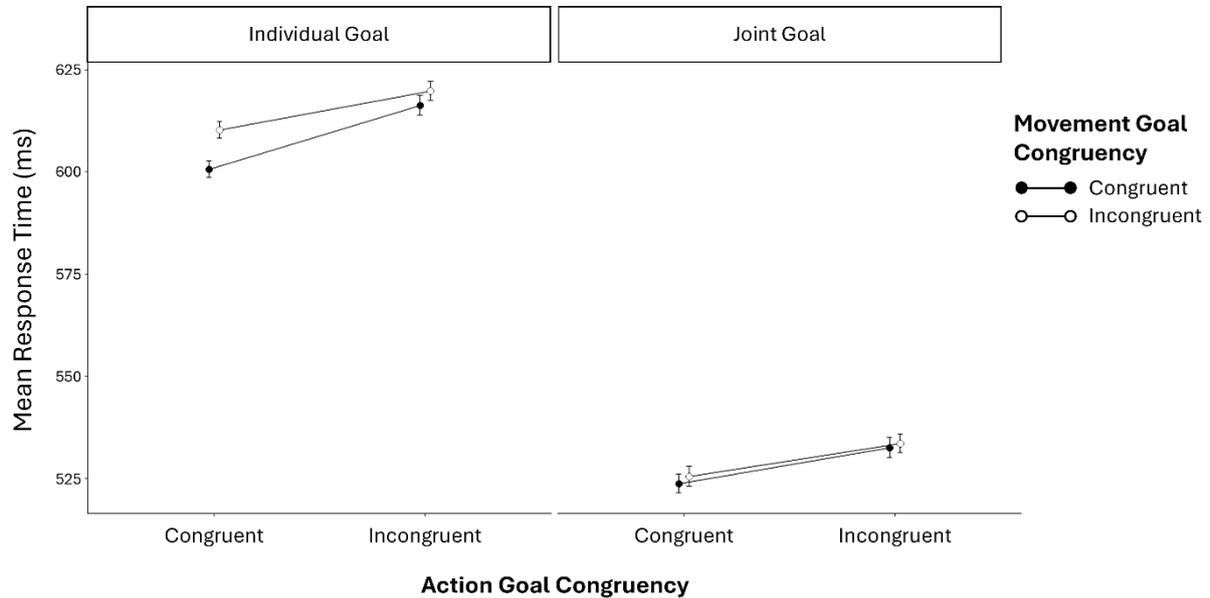
Figure 4-9. Mean RT in Experiment 2 as a function of Task Condition, Action Goal Congruency and Movement Goal Congruency. Error bars represented standard errors of the mean corrected for within-subjects designs.

The ER analysis of Experiment 3 revealed no significant effects (all $F[1, 207] < 2.24$,; see Appendix D for full result table). As in both previous experiments, the RT analysis again revealed a significant main effect of Action Goal Congruency, $F(1,207) = 40.9$, $p < .001$, $\eta_p^2 = .165$, showing that participants responded faster on action goal congruent trials ($M = 565$ ms, 95 % CI = [555, 576]) compared to action goal incongruent trials ($M = 576$ ms, 95 % CI = [565, 586]). This time though, the RT analysis also revealed a significant main effect of Task Condition, $F(1,207) = 61.8$, $p < .001$, $\eta_p^2 = .230$, indicating that participants selected their cards generally faster in the joint goal condition ($M = 529$ ms, 95 % CI = [514, 544]) compared to the individual goal condition ($M = 612$ ms, 95 % CI = [597, 626), and a significant main effect of Movement Goal Congruency, $F(1,207) = 7.46$, $p = .007$, $\eta_p^2 = .035$, driven by overall faster RTs on movement goal congruent trials ($M = 568$ ms, 95 % CI = [558, 579]) compared to movement goal incongruent trials ($M = 572$ ms, 95 % CI = [562, 583]). Yet, we still found no

evidence that the main effect of Action Goal Congruency was modulated by Task Condition, with the interaction between both factors being non-significant, $F(1,207) = 1.57$, $p = .212$, $\eta_p^2 = .008$. Likewise, we found no evidence that the main effect of Movement Goal Congruency was modulated by Task Condition, with the interaction term of both factors being close to but not reaching the statistical significance threshold, $F(1,207) = 3.14$, $p = .078$, $\eta_p^2 = .015$. Descriptively, the difference between movement goal congruent and movement goal incongruent trials was larger in the individual goal condition ($M_{Difference} = 6.59$ ms, 95 % CI = [2.52, 10.7]) compared to the joint goal condition ($M_{Difference} = 1.41$ ms, 95 % CI = [-2.68, 5.50]). The interaction between Action Goal Congruency and Movement Goal Congruency was non-significant, $F(1,207) = 1.00$, $p = .317$, $\eta_p^2 = .005$, as well as the three-way interaction between all design factors ($F < 1$).

Discussion

Overall, Experiment 3 replicated the main observations of the previous two experiments. As in the previous experiments, participants selected their cards faster when imitating rather than counter-imitating their task partner's action goal, providing further evidence for an imitative congruency effect between co-actors' individual action goals.

Furthermore, we still found no evidence that this imitative tendency is easily overcome when own and others' actions are performed to achieve a joint action goal together. Instead, congruency between co-actors' individual-level action goals appeared to affect task performance in the individual and joint goal condition in similar ways.

A novum of Experiment 3 is that we observed an action goal congruency effect across task instructions despite a general performance advantage in the joint compared to the individual task goal condition. A likely explanation of the performance advantage in the joint compared to the individual goal condition is that participants in the joint goal condition could deduce their individual contribution to the joint outcome from observing their partner's card

123

selection. As such they simply had to search for the inferred card value that complemented the instructed sequence once their cards turned over, while participants in the individual goal condition still had to process the relative magnitude relations between their revealed cards to make their selection.

The performance advantage in the joint goal condition clearly shows that participants performed the task differently in the two task conditions and benefited from selecting one of their cards to complement the card selection of their partner rather than selecting the higher or the lower of their two individually assigned cards. Nevertheless, the finding of an imitative congruency effect between co-actors' individual action goals across task conditions indicates that this advantage did not help participants to overcome a remaining tendency to imitate the individual action goal of their task partner. The results of Experiment 3 thus suggest that although acting towards a joint rather than an individual action goal can boost overall task performance, representations of joint action goals appear limited in their ability to modulate imitative response tendencies that stem from representations of own and others' actions in terms of their individual-level action goals.

Lastly, in contrast to the previous two experiments, Experiment 3 also revealed an imitative congruency effect at the level of co-actors' movement goals. Participants selected their cards slightly faster when imitating rather than counter-imitating their task partner's movement goal of selecting the left or the right of their cards. Furthermore, Experiment 3 provided no evidence that this effects was modulated by co-actors' higher-level action goals nor by task instructions to work towards individual or joint action goals respectively. This finding shows the presence of an involuntary tendency to imitate low-level movement features of observed actions in the current study. Yet, it should be noted that the movement goal congruency effect found in the present experiment was rather small which may have masked descriptive differences in the size of this effect between the different levels of our other experimental manipulations (i.e., Action Goal Congruency and Task Condition).

4.2.6. General Discussion

In three Experiments, we investigated the role of joint and individual action goals in shaping involuntary tendencies to imitate aspects of others' actions. We manipulated imitative congruency relations between higher-level action goals and lower-level movement goals of the consecutive actions of two co-actors (of which one was a virtual agent) that either worked towards individual or towards joint action goals. Crucially, irrespective of acting towards an individual or a joint action goal, the individual action and movement goals of participants' task partners were irrelevant for selecting participants' own contribution to the task. This allowed us to assess if participants would nevertheless show involuntary tendencies to imitate their task partner's action and/or movement goals, and how both tendencies might interact with each other. Furthermore, our study tested whether involuntary tendencies to imitate individual-level action or movement goals of an interaction partner become modulated when own and others individual actions are directed to the achievement of a joint action goal.

Across all experiments, we consistently found an imitative congruency effect between co-actors' individual action goals: Participants selected their cards faster when their individual action goal (selecting the higher or the lower of their cards) matched that of their partner. This effect was robust across various experimental manipulations and was observed regardless of whether participants performed their actions to achieve an individual action goal alone or a joint action goal together with their task partner, and regardless of whether participants observed their task partner's actions before or after receiving task-relevant information about their own task contribution. These findings indicate that participants encoded their task partner's individual action goal automatically despite not being required or instructed to do so and provide evidence that automatic encoding of others' individual action goals can interfere with the implementation of subsequent task contributions.

In contrast, evidence for imitative congruency effects at the level of co-actors' movement goals (i.e., selecting cards at the same or opposite relative spatial location) was limited and less consistent in the current study. Evidence for involuntary imitation of movement goals was only found in one out of three experiments. In the other two experiments movement goal congruency effects were modulated by higher-level congruency relations between co-actors' individual action goals. Thus, while effects of movement goal congruency remained small overall, these inconsistencies suggest that involuntary imitation of lower-level movement goals is susceptible to a range of contextual task factors.

Finally, although our findings showed that instructions to pursue joint rather than individual action goals could improve overall task performance, our results provided no evidence that joint action goals enabled participants to overcome an evident tendency to imitate individual-level action goals of their co-actor. Our findings showed that observing a task partner's action while representing a joint action goal can support subsequent action selection, possibly because joint action goals enable people to infer their own task contributions from that of their task partner. However, our findings also show that encoding of and alignment with individual-level action goals of other co-actors appears to be a robust and persistent feature of shared task contexts, even if own and others actions are jointly directed to the achievement of an overarching joint action goal.

Implications for theories of imitation

Our findings inform ongoing debates about the mechanisms and representations that drive people's propensity to imitate others' actions. In line with previous findings (Marschner et al., 2025; Ondobaka et al., 2012, 2013) and goal-directed theories of imitation (Gattis et al., 2002; Wohlschläger et al., 2003) our study provides complementary evidence that cognitive interpretations of observed actions in terms of their higher-level action goals are a central driver of peoples propensity to imitate the actions they observe in others. Our findings show that

people possess an automatic tendency to encode the higher-level action goals of observed actions (c.f., Csibra et al., 1999; Hassin et al., 2005; Woodward, 1998), and tend to imitate these goals to guide performance of their own subsequent actions (c.f., Aarts et al., 2004).

At the same time, the present study revealed only inconsistent evidence for involuntary imitation of lower-level movement goals. This suggests that tendencies to imitate low-level movement features of observed action are much more dependent on specific task contexts than would be predicted by associative accounts of imitation, challenging the notion that imitative behavior is mainly driven by low-level sensorimotor associations (e.g., Heyes, 2011).

Moreover, the findings of Experiment 1 suggest that tendencies to imitate lower-level movement goals become modulated by implicit congruency relations between task partners' higher-level individual action goals. This would extend previous findings showing that involuntary tendencies to imitate lower-level movement goals may not only be shaped by *explicit instructions* to imitate or counter-imitate others' higher-level action goals (Marschner et al., 2025; Ondobaka et al., 2012) but also by *involuntary tendencies* to do so. However, while this finding aligns with proposals that representations of own and others' actions in terms of higher-level conceptual action goals modulate the impact of lower-level sensorimotor representations of own and others' actions (Ondobaka & Bekkering, 2012, 2013), these findings remain only suggestive and need to be backed up by further investigation.

However, limiting the generalizability of our findings, the minor influence of low-level movement goals on imitation behavior in the present study could be due to characteristics of our virtual online setup. We operationalized congruency between co-actors' movement goals by manipulating the relative spatial location of their action targets (i.e., spatial features of movement endpoints), and not by manipulating congruency between own and observed movements per se (i.e., topographical features of bodily effectors and movement trajectories). Although it has been shown that compatibility between spatial features of own and others movement endpoints contributes to a large part to movement imitation effects (Cracco et al.,

127

2018), involuntary tendencies to imitate lower-level movement goals may have been more pronounced in the present study if participants would have observed not only the spatial endpoint but also other features of their partner's movements. Future studies are required to determine whether the dominance of action goal imitation over movement goal imitation suggested by the present findings persists in more naturalistic task settings featuring observation of bodily movement features as well.

Implications for theories of joint action

The present findings also inform current theories of joint action planning. Our findings indicate that co-actors show a persistent tendency to imitate each other's individual action goals even if own and others' individual task contributions are performed to achieve an overarching joint action goal. This finding suggests that representations of co-actors individual-level action goals still possess a dominate role in joint action planning and seem to remain a source of interpersonal influence and interference. One the one hand, this finding highlights the hierarchical nature of joint action planning, comprising both group-level and individual-level representations (Loehr, 2022; Pacherie, 2012; Pesquita et al., 2018; Zapparoli et al., 2022). On the other hand, our findings point to limitations whereby group-level representations (i.e. joint action goals) organize and constrain the interplay between individual-level representations of own and others' actions. While previous findings have shown that joint action goals can counteract people's tendency to imitate the *movement goals* of their task partners (Clarke et al., 2019; Rocca et al., 2023; Sacheli et al., 2018), our findings suggest that joint action goals might be limited in overriding imitative response tendencies that emerge between co-actors' *individual-level action goals*.

This apparent limitation may suggest that acting towards a joint action goal does not automatically imply that co-actors represent their own and their task partners' contribution to a joint task in a highly integrated manner that imposes relational constraints on individual action

128

selection and control processes (c.f., Sinigaglia & Butterfill, 2022). Rather, our findings suggest that in many cases co-actors may default to represent own and others actions merely as parallel task contributions each contributing their individual part to an overarching joint action goal.

However, starting from our findings, future studies will need to establish whether there are conditions where joint action goals could also become able to suppress involuntary tendencies of co-actors to imitate each other's individual action goals. A limitation of the current study could have been, that the joint goal instructions did not specify relations between co-actors individual-level action goals (e.g., same/different relations between individual action goals) but only between specific outcomes of their individual action contributions (i.e., height relations between specific card values). It is thus possible that joint action goals could modulate involuntary tendencies to imitate a co-actor's individual-level action goal when a specific relation between co-actors' individual-level action goals becomes instrumental for the achievement of their joint action goal. Applied to the current study, one would need to find conditions where the achievement of a joint action goal would be contingent on partner A selecting the higher or lower of their card while partner B selects the higher or the lower of their cards respectively.

Conclusion

In summary, the present study provides consistent evidence that people tend to imitate automatically encoded action goals of their task partners despite them being irrelevant for their own task performance. In contrast, the present study revealed only inconsistent evidence for involuntary imitation of low-level movement features of a task partner's actions. These findings support goal-directed theories of imitation by showing that imitative response tendencies are not solely driven by low-level sensorimotor associations but can derive from cognitive interpretations of observed actions in terms of their higher-level action goals as well. Against our expectations, however, the present study revealed no evidence that imitative response

tendencies are further modulated by task instructions to pursue joint rather than individual action goals. This suggests that representations of joint action goals, despite being able to boost overall task performance, appear limited in modulating imitative response tendencies in joint action contexts.

5. General Discussion

The aim of the present dissertation was to offer an account of the content and structure of joint task representations. Building up on previous research findings, I started by proposing three possible structures of joint task representations and spelled out how these structures could support interpersonal coordinating in joint action. In the subsequent chapters, I presented four empirical studies that were all tailored to test central assumptions of one of these proposals that I have referred to as the *relational structure* of joint task representations. The central idea behind this proposal is that joint task representations encode relations between individual action contributions of self and others that capture group-level features of co-actors' joint task performance (c.f., Kourtis et al., 2019; Pezzulo et al., 2025). In the last chapter of this dissertation, I will summarize the key findings of the presented studies and discuss their theoretical implications for the content and structure of joint task representations.

5.1. Key findings

5.1.1. Relations between co-actors' individual task contributions shape anticipated imitation between groups

Study I probed the structure of joint task representations by investigating anticipated imitation effects in a virtual reality experiment simulating social interactions between groups. To this end, Study I employed an anticipated imitation task in which participants performed a stimulus-response task either with or without a virtual co-actor while either one or both member(s) of another dyad of virtual agents imitated participants' task responses.

Building up on the assumption that people use anticipated actions they evoke in others to initiate and control their own contributions to a social interaction (Kunde et al., 2018), this task setup allowed us to test if joint task representations encode group-level relations between co-

actors' individual task contributions: If this is the case, anticipating that group-level relations between own and others' task contributions will be imitated by another group of co-actors should facilitate joint task performance.

The results of Study I supported this prediction with some critical limitations. Our results showed that participants made fewer response errors when the number of imitative responses of another dyad matched the number of task responses performed by participants' own dyad. I.e., performing the task *together* with one's co-actor was facilitated when *both* vs. one member(s) of the opposing dyad performed an imitative response, whereas performing the task *without* one's co-actor was facilitated when *one* vs. both member(s) of the opposing dyad performed an imitative response.

Although the size of this effect was small and a similar descriptive pattern in participants' response times did not reach statistical significance, the presence of this effect indicates that participants' task performance was modulated by congruency relations between groups: Task performance was facilitated when group-level features of the anticipated response of another group (here the number of simultaneously performed actions) matched group-level features of a planned response performed by one's own group. As such, this finding provides a proof of concept that anticipated imitation effects can be shaped by group-level congruency relations between interacting dyads. This supports the assumption that joint action partners represent their own contributions to a joint task not in isolation but in relation to the task contribution of their co-actors, i.e. that they form joint task representations that encode group-level (e.g., *WE* lift two index fingers) rather than individual-level performance (e.g., *I* lift my index finger).

However, a critical limitation of Study I was that the described performance pattern, warranting these theoretical conclusions, showed to be moderated by participants' overall response speed. In an explorative analysis targeting a combined measure of participants' task performance, group-level congruency relations between the two dyads were only found to effect participants who responded relatively slow but not participants who responded relatively fast

132

on the task. This finding indicates that only a subset of participants was affected by anticipated responses of the other agents in the task, pointing towards limitations in how readily people integrate behavioural responses evoked in others into their own task representations. This suggests that the formation of joint task representations that encode relations between own and others individual contributions to a joint action outcome is not an automatic process but rather dependent on a set of contextual and/or motivational factors that need to be worked out in future research.

5.1.2. Action-outcome learning is sensitive to relations between co-actors' individual task contributions

Study II probed the structure of joint task representations by investigating action-outcome learning in synchronous joint action. Starting from the assumption that joint task representations encode relationships between co-actors' actions and their ensuing joint action outcomes in some way, Study II tested how co-acting individuals link jointly produced action outcomes to features of their own and their partners' preceding response contributions.

Study II tested two hypothesis yielding different predictions of how co-actors form action-outcome representations in joint action contexts. If joint task representations take on a minimal structure, co-actors should link the production of joint action outcomes to individual-level features of their own task contribution viewed in isolation. In contrast, if joint task representations take on a relational structure, co-actors should link the production of joint action outcomes to group-level features of their joint task contributions viewed together.

To test these hypotheses against each other, Study II employed a two-stage action outcome learning task derived from research on ideomotor learning (c.f., Elsner & Hommel, 2001) in which the instructed stimulus-response mappings in a choice-reaction time task performed by two co-actors were manipulated to be either compatible or incompatible with the

individual and/or the group-level action-outcome mapping that co-actors experienced in a preceding learning phase.

In support of the hypothesis that co-actors link jointly produced action outcomes to group-level features of their joint response contributions rather than to individual-level features of their own response contributions viewed in isolation, we found that participants' task performance was affected by group-level but not by individual-level compatibility between learning and test phase mappings. However, in contrast to predictions derived from associative accounts of action-outcome learning (c.f., Moeller & Pfister, 2022 for review), this effect was driven by *better* task performance when the stimulus-response mapping was *incompatible* rather than compatible with co-actors' group-level action-outcome mapping of the preceding learning phase.

To explain the unexpected direction of the group-level compatibility effect, we drew on recent proposals attributing action-outcome learning effects in the two-stage action-outcome learning task to strategic rule use rather than to associative response priming (c.f., Custers, 2023; Sun et al., 2022; see Custers, 2024; Hommel & Eder, 2024; Kunde & Janczyk, 2024 for discussion). Following these proposals, we argued that co-actors interpreted the ambiguous stimulus-response rules in the test phase of our experiment in a way that could be most easily reconciled with their previous group-level learning experience. We proposed that this let co-actors to encode the stimulus-response task in the test phase either as a joint task performed together with their partner (if the instructed stimulus-response mapping was compatible with co-actors group-level learning experience) or rather as an individual task performed independently of their co-actor (if the instructed stimulus-response mapping was incompatible with co-actors group-level learning experience). Following this interpretation, the reversal of the group-level compatibility effect could be explained as the result of higher cognitive demands associated with the implementation of joint vs. individual task representation.

The findings of Study II and our proposed interpretation bear several implications for theoretical accounts of joint task representations. First, our findings indicate that experiencing contingent relationships between own actions, others' actions, and ensuing sensory outcomes can lead to the implicit formation of task representations that link joint action outcomes to group-level relations between co-actors' individual task contributions. This suggests that joint task representations embedding a relational structure are readily acquired if co-actors are required to synchronize individual task contributing to produce distinct joint action outcomes.

Second, however, the findings of Study II also indicate that representing group-level relations between own and others' task contributions is not an automatic response to any joint task context. Rather, the proposed interpretation of our results suggests that co-actors may deliberately choose to represent a task performed alongside another co-actor in relation to individual-level performance alone if that helps to overcome conflict with previous learning experience.

Finally, the proposed interpretation of our results further implies that joint task representations embedding a relational structure incur higher processing costs compared to leaner task representations that encode individual task contributions alone. This suggests that representing group-level relations between own and others' contributions to a joint task can become detrimental if the task could also be performed without other co-actors in mind.

5.1.3. Representing relations between own and others' action outcomes does not help to overcome imitative response tendencies

Study III and Study IV probed the content and structure of joint task representations by investigating modulations of imitative response tendencies in joint task settings. Both studies were motivated by recent research findings indicating that representations of shared goals can dissolve automatic tendencies to imitate observed actions of another task partner (Clarke et al., 2019; Rocca et al., 2023; Sacheli et al., 2018). Study III and IV were guided by the idea that

these modulation effects provide a signature of joint task representations that encode relations between co-actors' individual task contributions. Based on this idea, Study III and IV were designed to probe the relational structure of joint task representations by assessing whether shared goals that specify relations between action outcomes of own and others' individual task contributions modulate imitative response tendencies. As another extension of earlier research, Study III and IV assessed modulations of imitative response tendencies at different hierarchical levels of action representations, namely at the level of co-actors' lower-level movement goals and at the level of co-actors' higher-level individual action goals.

In a series of online experiments, we assessed imitative response tendencies in different variants of a sequential target selection task in which participants selected action targets in turns with a virtual co-actor to achieve either individual or shared goals. Across experiments we found consistent evidence for an imitative congruency effect between co-actors' individual-level action goals. Participants executed their own task contributions faster when imitating rather than counter-imitating the individual action goal of their virtual partner. This effect was observed regardless of whether the individual action goal of participants' task partner was relevant (c.f. Study III) or irrelevant (c.f. Study IV) for selecting their own task contributions. Crucially though, this effect remained robust against task instructions to complement the task contribution of a task partner in the service of a shared goal that specified relations between the outcomes of co-actors' individual task contributions. Although acting towards a shared goal showed to be beneficial for overall task performance if it allowed participants to deduce individual task contributions from the task contributions of their co-actor, our findings thus point to limitations whereby shared goals modulate imitative response tendencies.

Thus, the findings of Study III and IV did not confirm the hypothesis that representing relations between own and others' action outcomes help co-actors to overcome imitative response tendencies. Accordingly, Study III and IV did not provide evidence for the relational structure of joint task representations. Rather, their findings suggest that parallel rather than

relational representations of own and others' individual task contributions constrain task performance in many interactive task settings.

5.2. Synthesis: Determinants of the content and structure of joint task representations

Taken together, the studies presented in this dissertation offer a multi-facetted picture of the content and structure of joint task representations. While the results of Study I and Study II supported the proposal that joint task representations can take on a relational structure, both studies also came with their own limitations and complications implying that co-actors do not represent group-level relations between own others' individual task contributions by default. The findings of Study III and Study IV suggest a similar conclusion, namely that the relational structure of joint task representations may reflect an exception rather than the rule in task settings involving coordinated contributions of multiple co-actors.

An implication of these mixed findings is that joint task representations do not default to a particular structure, but are rather flexibly adapted to the specific characteristics and demands of the task at hand. Thus, instead of asking if joint task representations take on either a minimal, parallel or relational structure, it might be more fruitful to investigate the factors that determine which kind of structure co-actors draw on when engaging in a particular task. In the last part of this dissertation, I will therefore discuss how the findings of the presented studies shed light on possible factors that determine the content and structure of joint task representations. These considerations could offer promising starting points for future research on the structure and function of joint task representations.

5.2.1. Coordination demands

A natural starting point to think about possible determinants of the content and structure of joint task representations is to look at the function they ought to fulfil. A guiding assumption

of the work presented in this dissertation was that joint task representations provide cognitive structures that prepare and support coordination between co-actors towards shared goals (c.f., Knoblich et al., 2011; Pacherie, 2012; Sebanz & Knoblich, 2021). With this functional characterisation in mind, an obvious conjecture is that the content and structure of joint task representations is largely determined by the coordination demands imposed by the interactive task they are meant to support.

Looking back at the different studies, the tasks they deployed to investigate the content and structure of joint task representations imposed quite different coordination demands upon participants. Starting with Study I, coordination demands imposed on participants were close to zero, as participants could perform the task by simply adhering to an instructed stimulus-response mapping that was completely independent from the task contributions of any other co-actor. At most, the task implied that a set of other agents coordinated their actions in response to the actions of the participant, making participants the leader of a social interaction completely determined by their own doings. Although implementing further task demands that required participants to monitor the performed actions of other agents in scene (catch trials in which the imitating avatars lifted their respective finger twice), the task used in Study I thus imposed only minimal requirements on participants to take the actions of other people into account to plan and perform their own task contributions. This could provide an explanation for the fact that a bulk of participants in Study I showed no indication of being affected by the behaviour of other agents in the task.

A related point can be made by looking at the coordination demands imposed by the tasks deployed in Study III and IV. Here, although participants had to select their own task contributions in relation to those of their task partner to achieve a shared goal, these coordination demands were still unidirectional as participants' task partners always acted first and were thus primarily in charge of determining the further course of the interaction. Thus, as in Study I, the task imposed a leader-follower hierarchy that prevented a need for mutual

138

adaptation and common coordination requirements between co-actors. The findings of Study III and IV may thus imply that merely acting as a follower who is only in charge of complementing the task contributions of a task partner makes it less likely that people form joint task representations that encode group-level relations between own and others' task contributions.

Examining the coordination demands imposed upon participants in Study II adds to this picture. Across studies, the highest coordination demands were present in the learning phase of Study II. Here, co-actors had to synchronize their individual response contributions with each other to produce a joint action outcome that could only be achieved if co-actors met specific spatiotemporal relations between their individual response contributions. Our findings suggest that this task context was most likely to result in the formation of task representations that encoded group-level relations between co-actors' individual response contributions. In contrast, similar coordination demands were absent in the test phase of Study II, were no synchrony requirements between co-actors' individual response contributions remained. The absence of coordination demands in the test phase may thus have made it unnecessary that co-actors went on to represent group-level relations between their individual task contributions. This may add to the proposed explanation of the reversed group-level compatibility effect which entails that participants who faced conflict between learning and test phase mappings reverted to encode the instructed stimulus-response mapping merely in relation to their individual performance alone.

Taken together, this discussion suggests that the content and structure of joint task representations are shaped by the coordination demands co-actors face in an interactive task setting. As interdependence between individual task contributions increases, and as demands for mutual adaptation and monitoring of others' actions rise, co-actors seem to become more likely to represent group-level relations between their individual task contributions. This interpretation aligns with other research showing that mutual adaptation, common coordination

139

requirements, and egalitarian role distributions are critical task characteristics that foster a sense of acting together as a group rather than as separate individuals (Bolt et al., 2016; Le Bars et al., 2020, 2022; Pacherie, 2014).

5.2.2. Further task characteristics

Beyond the coordination demands imposed by a specific task, the findings of the present dissertation indicate that additional task characteristics are also likely to shape the content and structure of joint task representations.

One critical aspect relates to the legitimate question of whether participants conceptualized the deployed tasks as interactive tasks performed together with other co-actors in the first place. This question arises specifically in the context of Study I, which simulated a social interaction between groups in a virtual environment depicting animated computer avatars, and in Study III and IV, in which participants interacted with a disembodied virtual co-actor (i.e., a computer program). Indeed, previous research indicates that interactive task behaviour is shaped by human-likeness in behaviour and appearance of artificial co-actors (Ciardo et al., 2022; Musco et al., 2025; Sahaï et al., 2019; Weller et al., 2019), suggesting that these features have a critical influence on how people represent joint tasks. It is thus possible that a necessary condition for the construction of joint task representations that integrate task contributions of other co-actors (either in a relational or in a parallel manner) is a certain degree of human-likeness in their appearance and behaviour that have been lacking in some of the presented experiments.

Second, another task characteristic that might determine how joint task representations become structured is whether co-actors' individual contributions to a joint task are performed in synchrony or sequentially (i.e. the temporal structure of the joint task). Our findings suggest that task contexts in which co-actors synchronised their individual task contributions to produce a joint action outcome (Study I and II) were more likely to engage joint task representations

embedding a relational structure than task context in which co-actors performed their individual contributions to a joint task temporally separated from each other (Study III and IV). A possible reason for this could be that synchronous task contributions of self and others are more readily represented in relation to each other (i.e. bound together as a single event representation comprising "our" actions) due to their close temporal contiguity (c.f., Moeller & Frings, 2019). Moreover, synchronous joint action obscure individual action-outcome contingencies, which makes it harder to track the results of co-actors' individual contributions to a joint action outcome in isolation. This could lead co-actors to switch from co-representing own and others' task contributions in parallel to the encoding of group-level features of their synchronized responses instead. In contrast, in sequential joint actions, individual contributions to a joint action outcome are clearly distinguishable. This could increase the likelihood that co-actors represent the individual task contributions of self and others as parallel subplans towards a shared goal.

Lastly, a similar point could also be made about the structure and appearance of the shared goal or joint action outcomes to be achieved in a joint task. If a joint action outcome is composed of clearly distinguishable contributions of each individual co-actor (as it was the case in Study III and IV) this may promote the engagement of joint task representations embedding a parallel structure. On the other hand, if a joint action outcome has a particular appearance or "Gestalt" that cannot simply be reduced to the mere summation of individual task contributions of self and others (e.g., the harmony of a jointly produced chord, as in the learning phase of Study II) this could foster the construction of joint task representations that take on a relational structure. Supporting this idea, a recent study by Zhou et al. (2023) showed that co-actors experience a stronger sense of joint agency (i.e. a feeling of acting as a group rather than as independent individuals) if they synchronize their actions to produce harmonic musical duets rather than a series of constant tones together.

5.2.3. Costs and benefits

Finally, another aspect raised by the findings of the presented studies is that the engagement of joint task representations that take on a relational structure seems to come at some costs. This is suggested by several findings throughout the dissertation. In Study I, those participants who showed indications for relying on a relational structure of joint task representations performed the task slower than participants who showed no indications for integrating other agents' actions into their own representation of the task. In Study II, the reversal of the group-level compatibility effect could be explained by assuming that encoding own and others' task contributions in relation to each other is cognitively more demanding than merely encoding one's own task contribution in isolation. Finally, Study III indicated that participants could have found it less demanding to represent the card selection task in relation to individual-level goals instead of relying on a joint task representation that encoded group-level relations between own and others individual task contributions.

These findings point towards another important factor that is likely to determine the content and structure of joint task representation. Confronted with a novel task, co-actors will need to find a balance between the representational costs associated with a particularly structured task representation and the benefits they offer for successful task performance (c.f., Correa et al., 2023; Ho et al., 2022; Shenhav et al., 2013, 2017). This consideration would predict that co-actors become less likely to engage joint task representations that represent group-level relations between own and others' individual task contributions if relying on a leaner representational structure of the task guarantees equally or even more efficient task performance.

A crucial implication of these considerations is that the content and structure of joint task representations is determined by decision making or construal processes balancing costs and benefits of different representational structures to support optimal task performance. This

consideration in mind, a promising direction for future research would be to link research on the content and structure of joint task representations to normative principles of optimization and utility in human thought and behaviour (Gershman et al., 2015; Griffiths et al., 2015; Lewis et al., 2014) and to optimality principles guiding action identification and task construal (Vallacher & Wegner, 1987, 2012).

5.3. Conclusion

The aim of the present dissertation was to offer a comprehensive account of the content and structure of joint task representations. Building up on previous research, I laid out three possible structures of joint task representations and conducted four empirical studies that war all tailored to test central assumptions of the proposal that joint task representations encode group-level relations between co-actors' individual contributions to a joint task. The findings presented in this dissertation provide partial support for this proposal, offering novel evidence that joint task representations can adopt a relational structure that encodes joint actions at the group level. At the same time, however, our results reveal important limitations regarding how frequently such relational structures are applied in interactive task settings. These limitations suggest, that joint task representations, rather than embedding a fixed structure centred on group-level relations between co-actors' individual task contributions, become flexibly adapted to cope with the specific demands and characteristics of the interactive task they aim to support.

Appendix

A. Supplementary material for Study I

Table A-1

Full ANOVA results for the explorative IES analysis

Effect	df	F	p	η_p^2
Task Condition	1, 233	0.38	.536	.002
Imitative Response	1, 233	0.20	.657	< .001
RT Group	1, 233	323.5	< .001	.581
Task Condition × Imitative Response	1, 233	3.81	.052	.016
Task Condition × RT Group	1, 233	2.08	.151	.009
Imitative Response × RT Group	1, 233	1.15	.284	.005
Task Condition × Imitative Response × RT Group	1, 233	6.61	.011	.028

Note. Inverse Efficiency Scores (IES) were analysed by means of a 2 (Task Condition: Individual vs. Joint Action) x 2 (Imitative Response: Individual vs. Joint Imitation) x 2 (RT Group: Fast vs. Slow Responder) mixed ANOVA with Task Condition and Imitative Response as within- and RT Group as between-subjects factors.

Table A-2

Full ANOVA results for the follow-up IES analysis of slow responder

Effect	df	F	p	η_p^2
Task Condition	1, 116	0.30	.588	.003
Imitative Response	1, 116	0.87	.352	.007
Task Condition × Imitative Response	1, 116	7.21	.008	.059

Note: Inverse Efficiency Scores (IES) were analysed by means of a 2 (Task Condition: Individual vs. Joint Action) x 2 (Imitative Response: Individual vs. Joint Imitation) x 2 (RT Group: Fast vs. Slow Responder) repeated measures ANOVA with Task Condition and Imitative Response as within-subjects factors.

Table A-3

Full ANOVA results for the follow-up IES analysis of fast responder

Effect	df	F	p	η_p^2
Task Condition	1, 117	2.48	.118	.021
Imitative Response	1, 117	0.29	.593	.002
Task Condition × Imitative Response	1, 117	0.33	.569	.003

Note: Inverse Efficiency Scores (IES) were analysed by means of a 2 (Task Condition: Individual vs. Joint Action) x 2 (Imitative Response: Individual vs. Joint Imitation) x 2 (RT Group: Fast vs. Slow Responder) repeated measures ANOVA with Task Condition and Imitative Response as within-subjects factors.

B. Supplementary material for Study II

B1. Learning phase results

Table B1-1

Learning phase performance as a function of subsequent test phase condition

	Solo Learning Phase				Joint Learning Phase (1st Block)				Joint Learning Phase (2nd Block)			
	Individually Comp.		Individually Incomp.		Individually Comp.		Individually Incomp.		Individually Comp		Individually Incomp.	
	Jointly Comp.	Jointly Incomp.	Jointly Comp.	Jointly Incomp.	Jointly Comp.	Jointly Incomp.	Jointly Comp.	Jointly Incomp.	Jointly Comp.	Jointly Incomp.	Jointly Comp.	Jointly Incomp.
Outcome Ratio	47% ±8%	50% ±2%	50% ±2%	51% ±4%	51% ±4%	49% ±3%	50% ±1%	51% ±2%	50% ±2%	52% ±3%	50% ±2%	50% ±1%
Error Rate	0.3% ±1%	0.3% ±0.8%	0.6%, ±1.2%	0%	5.8% ±3.1%	7.7%, ±2.5%	**13.8%, ±9.3%**	5.8%, ±3.5	6.5% ±3.4%	5.4% ±4.8%	4.7% ±4.7%	9.6% ±5.6%

Note. Group means and standard deviations of outcome ratios and error rates in the solo and joint part of the learning phase. Outcome ratios represent the relative frequency of high tones and high chords produced in the solo and joint part of the learning phase respectively. None of the outcome ratios differed significantly from chance (all $t < 1.96$). Analysis of error rates in the first block of the joint learning phase as a function of pairs' subsequent test phase condition revealed a significant interaction between Individual Compatibility and Joint Compatibility, $F(1,32) = 7.71$, $p = 0.009$, $\eta_p^2 = .194$, indicating that pairs' performance in the first block of the joint learning phase was not comparable between the four test phase groups. Subsequent pairwise comparisons between the four test phase groups showed that the interaction was driven by higher error rates in the group who later performed the test phase with an individually incompatible/jointly compatible stimulus response mapping (marked in bold) compared to the other three groups.

B2. Test phase results with Joint Compatibility as within-subjects factor

Table B2-1

Descriptive statistics for the full experimental design

	Jointly Compatible 1st				Jointly Incompatible 1st			
	Individually Compatible		Individually Incompatible		Individually Compatible		Individually Incompatible	
	Jointly Compatible	Jointly Incompatible	Jointly Compatible	Jointly Incompatible	Jointly Compatible	Jointly Incompatible	Jointly Compatible	Jointly Incompatible
IES	632ms ±21.7ms	583ms ±21.7ms	720ms ±46.3ms	646ms ±46.3ms	582ms ±28.7ms	579ms ±28.7ms	546ms ±29.2ms	584ms ±29.2ms
RT	567ms ±15.9ms	541ms ±15.9ms	595ms ±20.7ms	551ms ±20.7ms	519ms ±16.7ms	509ms ±16.7ms	515ms ±13.8ms	521.6ms ±13.8ms
ER	9.61% ±2.61%	6.76% ±2.61%	16.0% ±4.14%	13.5% ±4.14%	10.1% ±2.38%	11.9% ±2.38%	5.38% ±2.74%	9.86% ±2.74%
|Asy|	125ms ±11.6ms	134ms ±11.6ms	165ms ±17.7ms	149ms ±17.7ms	114ms ±13.6ms	115ms ±13.6ms	114ms ±21.8ms	110ms ±21.8ms

Note. Group means and 95% confidence intervals for the test phase performance measures for each cell of the 2 (Individual Compatibility: Ind. Compatible vs. Ind. Incompatible) x 2 (Joint Compatibility: Joint Compatible vs. Joint Incompatible) x 2 (Block Order: Joint Compatible 1st vs. Joint Incompatible 1st) experimental design with Joint Compatibility manipulated within-subjects and Individual Compatibility and Block Order manipulated between-subjects. Confidence intervals are corrected for within subjects designs (Morey, 2008).

Table B2-2

ANOVA results for the full experimental design

Effect	IES			RT			ER			|Asy|		
	$F(1,66)$	p	η_p^2	$F(1,66)$	p	η_p^2	$F(1,66)$	p	η_p^2	$F(1,31)$	p	η_p^2
Ind. Compatibility	2.20	.143	.032	0.76	.388	.011	0.92	.342	.014	2.06	.161	.062
Joint Compatibility	3.93	.051	.056	10.5	**.002**	.137	0.06	.807	< .001	0.27	.605	.009
Block Order	13.4	**< .001**	.166	12.3	**< .001**	.157	1.68	.199	.014	11.8	**.002**	.276
Ind. Compatibility x Joint Compatibility	0.14	.708	.002	0.01	.935	< .001	0.54	.466	.008	2.17	.151	.065
Ind. Compatibility x Block Order	5.12	**.027**	.072	0.26	.615	.004	8.86	**.004**	.118	2.96	.095	.087
Joint Compatibility x Block Order	12.5	**< .001**	.159	8.45	**.005**	.114	7.91	**.006**	.107	0.01	.912	< .001
Ind. Compatibility x Joint Compatibility x Block Order	2.20	.143	.032	2.30	.135	.034	0.29	.589	.004	0.89	.353	.028

Note. IES, RTs, ER and |Asy| were each subject to 2 (Individual Compatibility: Ind. Compatible vs. Ind. Incompatible) x 2 (Joint Compatibility: Joint Compatible vs. Joint Incompatible) x (Block Order: Joint Compatible 1st vs. Joint Incompatible 1st) mixed ANOVA with Joint Compatibility as within-subjects factor and Individual Compatibility and Block Order as between-subjects factors. Significant effects are marked in bold font.

C. Supplementary material for Study III

C1. Additional results for Experiment 1

Table C1-1

Bayesian Model Comparison Results for Response Time Analysis in Experiment 1

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Movement Goal Congruency + Action Goal Congruency + Movement Goal Congruency × Action Goal Congruency	0.053	0.338	9.201	1.000	
Action Goal Congruency	0.053	0.193	4.293	1.757	3.029
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency	0.053	0.169	3.658	2.003	8.692
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Action Goal Congruency × Task Goal	0.053	0.086	1.683	3.956	10.217
Action Goal Congruency + Task Goal	0.053	0.084	1.656	4.016	6.297
Action Goal Congruency + Task Goal + Action Goal Congruency × Task Goal	0.053	0.048	0.905	7.064	6.754
Movement Goal Congruency + Action Goal Congruency	0.053	0.026	0.481	12.991	17.356
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Movement Goal Congruency × Task Goal	0.053	0.024	0.439	14.220	11.878
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Movement Goal Congruency × Task Goal + Action Goal Congruency × Task Goal	0.053	0.012	0.210	29.304	10.631
Movement Goal Congruency + Action Goal Congruency + Task Goal	0.053	0.009	0.170	36.113	7.868

Note. Only the 10 best out of all 19 models are shown. Results were obtained using a default prior specification for model parameters with *r* scale for fixed effects = .5. Model comparison results using two different prior settings can be found in the supplement.

Table C1-2

ANOVA Results for Error Rate Analysis in Experiment 1

Effect	df	F	p	η_p^2
Task Goal	1,188	1.105	.295	.006
Action Goal Congruency	1,188	6.349	.013	.033
Movement Goal Congruency	1,188	3.821	.052	.02
Task Goal × Action Goal Congruency	1,188	0.065	.799	3.4×10^{-4}
Task Goal × Movement Goal Congruency	1,188	0.898	.345	.005
Action Goal Congruency × Movement Goal Congruency	1,188	8.760	.003	.045
Task Goal × Action Goal Congruency × Movement Goal Congruency	1,188	1.002	.318	.005

Note. Based on Type III Sum of Squares
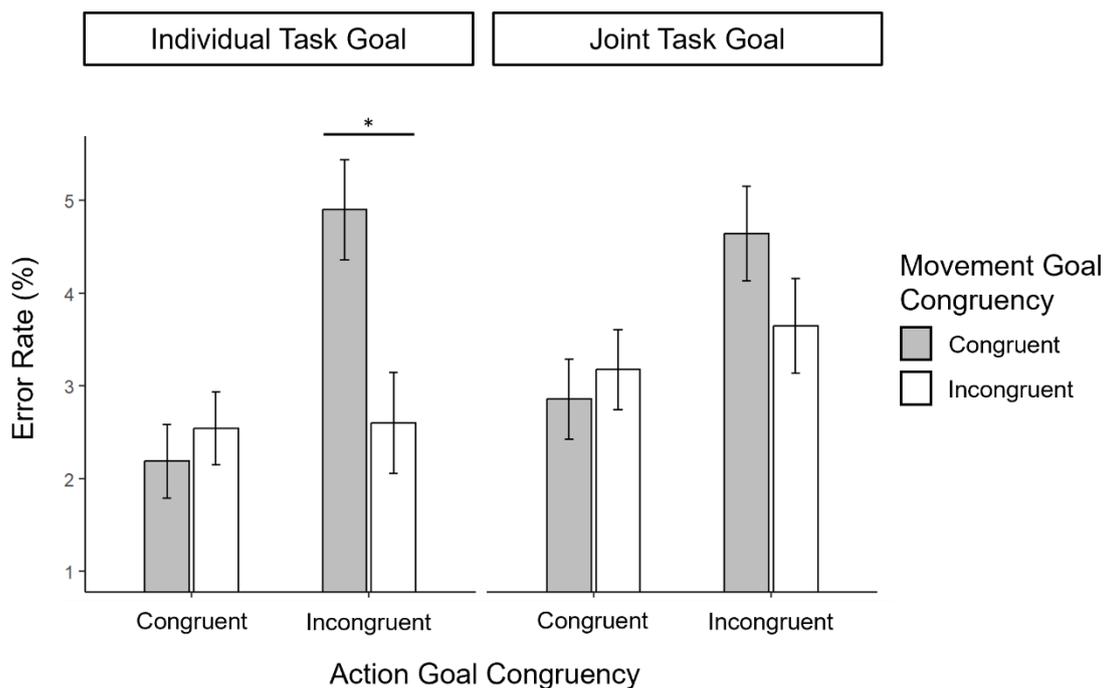

Figure C1-1 Mean error rates in Experiment 1 as a function of Task Goal, Action Goal Congruency and Movement Goal Congruency. Error bars represent standard errors of the mean (SEM) corrected for within subjects designs (Morey, 2008). Asterisks mark simple main effects between factor levels reaching significance after Bonferroni-correcting alpha levels for multiple comparisons. * p < .05.

C2. Additional results for Experiment 2

Table C2-1

Bayesian Model Comparison Results for Response Time Analysis in Experiment 2

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Movement Goal Congruency + Action Goal Congruency + Movement Goal Congruency × Action Goal Congruency	0.053	0.397	11.841	1.000	
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency	0.053	0.272	6.733	1.458	8.974
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Action Goal Congruency × Task Goal	0.053	0.207	4.710	1.913	6.339
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Movement Goal Congruency × Task Goal + Action Goal Congruency × Task Goal	0.053	0.055	1.053	7.181	11.928
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Movement Goal Congruency × Task Goal	0.053	0.052	1.006	7.495	10.774
Movement Goal Congruency + Action Goal Congruency + Task Goal + Movement Goal Congruency × Action Goal Congruency + Movement Goal Congruency × Task Goal + Action Goal Congruency × Task Goal + Movement Goal Congruency × Action Goal Congruency × Task Goal	0.053	0.015	0.281	25.787	33.670
Movement Goal Congruency	0.053	3.5×10^{-13}	6.3×10^{-12}	1.1×10^{12}	4.041
Action Goal Congruency + Task Goal + Action Goal Congruency × Task Goal	0.053	2.4×10^{-13}	4.0×10^{-12}	1.6×10^{12}	6.261
Action Goal Congruency + Task Goal	0.053	1.9×10^{-13}	3.3×10^{-12}	2.1×10^{12}	6.165
Movement Goal Congruency + Action Goal Congruency	0.053	9.4×10^{-14}	1.7×10^{-12}	4.2×10^{12}	4.011

Note. Only the 10 best out of all 19 models are shown. Results were obtained using a default prior specification for model parameters with *r* scale for fixed effects = .5. Model comparison results using two different prior settings can be found in the supplement.

Table C2-2

ANOVA Results for Error Rate Analysis in Experiment 2

Effect	df	F	p	η_P^2
Task Goal	1,184	4.562	.034	.024
Action Goal Congruency	1,184	8.505	.004	.044
Movement Goal Congruency	1,184	0.044	.834	2.4×10^{-4}
Task Goal × Action Goal Congruency	1,184	1.939	.166	.010
Task Goal × Movement Goal Congruency	1,184	0.027	.869	1.4×10^{-4}
Action Goal Congruency × Movement Goal Congruency	1,184	16.632	< .001	.083
Task Goal × Action Goal Congruency × Movement Goal Congruency	1,184	0.206	.651	.001

Note. Based on Type III Sum of Squares

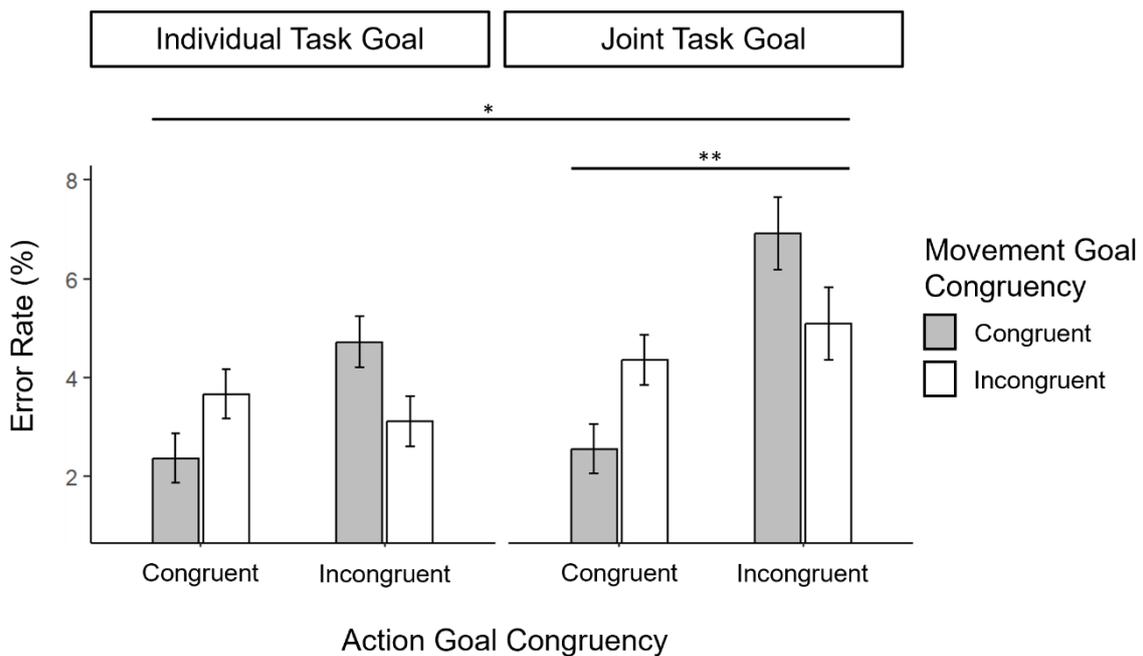**Figure C2-1.** Error rates in Experiment 2 as a function of Task Goal, Action Goal Congruency and Movement Goal Congruency. Error bars represent standard errors of the mean (SEM) corrected for within subjects designs (Morey, 2008). Asterisks mark simple main effects between factor levels reaching significance after Bonferroni-correcting alpha levels for multiple comparisons. ** p < .01, * p < .05.

D. Supplementary material for Study IV

Table D-1

ANOVA Results for Error Rate Analysis in Experiment 1

Effect	df	F	p	η_p^2
Task Condition	1,57	0.17	.678	.003
Action Goal Congruency	1,57	0.20	.655	.004
Movement Goal Congruency	1,57	0.18	.674	.003
Task Condition × Action Goal Congruency	1,57	3.30	.074	.055
Task Condition × Movement Goal Congruency	1,57	0.03	.866	< .001
Action Goal Congruency × Movement Goal Congruency	1,57	4.58	.037	.074
Task Condition × Action Goal Congruency × Movement Goal Congruency	1,57	1.73	.193	.030

Note. Based on Type III Sum of Squares

Table D-2

ANOVA Results for Error Rate Analysis in Experiment 2

Effect	df	F	p	η_p^2
Task Condition	1,58	1.02	.317	.017
Action Goal Congruency	1,58	0.67	.415	.011
Movement Goal Congruency	1,58	0.06	.811	< .001
Task Condition × Action Goal Congruency	1,58	0.23	.635	.004
Task Condition × Movement Goal Congruency	1,58	0.92	.341	.016
Action Goal Congruency × Movement Goal Congruency	1,58	0.07	.790	.001
Task Condition × Action Goal Congruency × Movement Goal Congruency	1,58	1.49	.223	.025

Note. Based on Type III Sum of Squares

Table D-3

ANOVA Results for Error Rate Analysis in Experiment 3

Effect	df	F	p	η_p^2
Task Condition	1,207	1.16	.283	.006
Action Goal Congruency	1,207	2.24	.136	.011
Movement Goal Congruency	1,207	1.20	.275	.006
Task Condition × Action Goal Congruency	1,207	0.28	.600	.001
Task Condition × Movement Goal Congruency	1,207	0.14	.711	< .001
Action Goal Congruency × Movement Goal Congruency	1,207	0.05	.821	< .001
Task Condition × Action Goal Congruency × Movement Goal Congruency	1,207	1.10	.296	.005

Note. Based on Type III Sum of Squares

References

Aarts, H., Gollwitzer, P. M., & Hassin, R. R. (2004). Goal contagion: Perceiving is for pursuing. *Journal of Personality and Social Psychology*, *87*(1), 23–37. https://doi.org/10.1037/0022-3514.87.1.23

Ansorge, U. (2002). Spatial intention–response compatibility. *Acta Psychologica*, *109*(3), 285–299. https://doi.org/10.1016/S0001-6918(01)00062-2

Ansorge, U., & Wühr, P. (2004). A Response-Discrimination Account of the Simon Effect. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 365–377. https://doi.org/10.1037/0096-1523.30.2.365

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: sharing tasks with real and imagined co-actors. *Experimental Brain Research*, *211*(3–4), 371–385. https://doi.org/10.1007/s00221-011-2709-9

Atmaca, S., Sebanz, N., Prinz, W., & Knoblich, G. (2008). Action co-representation: The joint SNARC effect. *Social Neuroscience*, *3*(3–4), 410–420. https://doi.org/10.1080/17470910801900908

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice Hall.

Bardi, L., Bundt, C., Notebaert, W., & Brass, M. (2015). Eliminating mirror responses by instructions. *Cortex*, *70*, 128–136. https://doi.org/10.1016/j.cortex.2015.04.018

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *53*(1), 153–164. https://doi.org/10.1080/713755872

Betti, S., Chinellato, E., Guerra, S., Castiello, U., & Sartori, L. (2019). Social Motor Priming: When offline interference facilitates motor execution. *PeerJ*, *7*. https://doi.org/10.7717/peerj.7796

Böckler, A., Knoblich, G., & Sebanz, N. (2012). Effects of a coactor's focus of attention on task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1404–1415. https://doi.org/10.1037/a0027523

Bolt, N. K., & Loehr, J. D. (2017). The predictability of a partner's actions modulates the sense of joint agency. *Cognition*, *161*, 60–65. https://doi.org/10.1016/j.cognition.2017.01.004

Bolt, N. K., Poncelet, E. M., Schultz, B. G., & Loehr, J. D. (2016). Mutual coordination strengthens the sense of joint agency in cooperative joint action. *Consciousness and Cognition*, *46*, 173–187. https://doi.org/10.1016/j.concog.2016.10.001

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*(2), 124–143. https://doi.org/10.1006/brcg.2000.1225

Brass, M., & Heyes, C. (2005a). *Imitation : is cognitive neuroscience solving the correspondence problem ? 9*(10). https://doi.org/10.1016/j.tics.2005.08.007

Brass, M., & Heyes, C. (2005b). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495. https://doi.org/10.1016/j.tics.2005.08.007

Brass, M., Liefooghe, B., Braem, S., & De Houwer, J. (2017). Following new task instructions: Evidence for a dissociation between knowing and doing. In *Neuroscience and Biobehavioral Reviews* (Vol. 81, pp. 16–28). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2017.02.012

Bratman, M. E. (1992). Shared Cooperative Activity. *The Philosophical Review*, *101*(2), 327. https://doi.org/10.2307/2185537

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: is the Inverse Efficiency Score (IES) a better dependent variable than the mean Reaction Time (RT) and the Percentage of Errors (PE)? *PSYCHOLOGICA BELGICA*, *51*(1), 5–13.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 1–38. https://doi.org/10.5334/joc.72

Butterfill, S. (2012). Joint action and development. *Philosophical Quarterly*, *62*(246), 23–47. https://doi.org/10.1111/j.1467-9213.2011.00005.x

Butterfill, S. (2015). Planning for Collective Agency. In C. Misselhorn (Ed.), *Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation* (pp. 149–168). Springer International Publishing. https://doi.org/10.1007/978-3-319-15515-9_8

Butterfill, S. (2018). Coordinating Joint Action. In M. Jankovic & K. Ludwig (Eds.), *Routledge handbooks in philosophy. The Routledge handbook of collective intentionality* (pp. 68–82). Routledge/Taylor & Francis Group.

Call, J. (2009). Contrasting the Social Cognition of Humans and Nonhuman Apes: The Shared Intentionality Hypothesis. *Topics in Cognitive Science*, *1*(2), 368–379. https://doi.org/10.1111/j.1756-8765.2009.01025.x

Candidi, M., Sacheli, L. M., & Aglioti, S. M. (2015). From muscles synergies and individual goals to interpersonal synergies and shared goals: Mirror neurons and interpersonal action hierarchies. *Physics of Life Reviews*, *12*, 126–128. https://doi.org/10.1016/j.plrev.2015.01.023

Carpenter, M. (2009). Just How Joint Is Joint Action in Infancy? *Topics in Cognitive Science*, *1*(2), 380–392. https://doi.org/10.1111/j.1756-8765.2009.01026.x

Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., & Heyes, C. (2008). Through the looking glass: counter-mirror activation following incompatible sensorimotor learning.

European Journal of Neuroscience, *28*(6), 1208–1215. https://doi.org/10.1111/j.1460-9568.2008.06419.x

Catmur, C., & Heyes, C. (2019). Mirroring 'meaningful' actions: Sensorimotor learning modulates imitation of goal-directed actions. *Quarterly Journal of Experimental Psychology*, *72*(2), 322–334. https://doi.org/10.1080/17470218.2017.1344257

Chiavarino, C., Bugiani, S., Grandi, E., & Colle, L. (2013). Is Automatic Imitation Based on Goal Coding or Movement Coding? *Experimental Psychology*, *60*(3), 213–225. https://doi.org/10.1027/1618-3169/a000190

Ciardo, F., De Tommaso, D., & Wykowska, A. (2022). Joint action with artificial agents: Human-likeness in behaviour and morphology affects sensorimotor signaling and social inclusion. *Computers in Human Behavior*, *132*, 107237. https://doi.org/10.1016/j.chb.2022.107237

Clark, H. H. (1996). *Using Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511620539

Clarke, S., McEllin, L., Francová, A., Székely, M., Butterfill, S. A., & Michael, J. (2019). Joint action goals reduce visuomotor interference effects from a partner's incongruent actions. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-52124-6

Cole, G. G., Atkinson, M. A., D'Souza, A. D. C., Welsh, T. N., & Skarratt, P. A. (2018). Are goal states represented during kinematic imitation? *Journal of Experimental Psychology: Human Perception and Performance*, *44*(2), 226–242. https://doi.org/10.1037/xhp0000429

Cole, G. G., Welsh, T. N., & Skarratt, P. A. (2019). The role of transients in action observation. *Attention, Perception, and Psychophysics*, *81*(7), 2177–2191. https://doi.org/10.3758/s13414-019-01740-5

Cook, R., Press, C., Dickinson, A., & Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(4), 840–852. https://doi.org/10.1037/a0019256

Correa, C. G., Ho, M. K., Callaway, F., Daw, N. D., & Griffiths, T. L. (2023). Humans decompose tasks by trading off utility and computational cost. *PLOS Computational Biology*, *19*(6), e1011087. https://doi.org/10.1371/journal.pcbi.1011087

Cracco, E. (2025). Automatic Imitation of Multiple Agents. In *Automatic Imitation* (pp. 199–218). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-62634-0_10

Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., Coster, L. De, 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., Bernardet, U., Sevenhant, R., Vandenhouwe, N., Copman, F., Durnez, W., Bombeke, K., & Brass, M. (2022). Evidence for a two-step model of social group influence. *IScience*, *25*(9), 104891. https://doi.org/10.1016/j.isci.2022.104891

Cracco, E., Braem, S., & Brass, M. (2021). Observing Conflicting Actions Elicits Conflict Adaptation. *Journal of Experimental Psychology: General*, *151*(2), 493–505. https://doi.org/10.1037/xge0001089

Cracco, E., & Brass, M. (2018a). Motor simulation of multiple observed actions. *Cognition*, *180*, 200–205. https://doi.org/10.1016/j.cognition.2018.07.007

Cracco, E., & Brass, M. (2018b). The role of sensorimotor processes in social group contagion. *Cognitive Psychology*, *103*, 23–41. https://doi.org/10.1016/j.cogpsych.2018.02.001

Cracco, E., De Coster, L., Andres, M., & Brass, M. (2015). Motor simulation beyond the dyad: Automatic imitation of multiple actors. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1488–1501. https://doi.org/10.1037/a0039737

Cracco, E., Keysers, C., Clauwaert, A., & Brass, M. (2019). Representing multiple observed actions in the motor system. *Cerebral Cortex*, *29*(8), 3631–3641. https://doi.org/10.1093/cercor/bhy237

Csibra, G. (2007). Action mirroring and action understanding: an alternative account. In *Sensorimotor Foundations of Higher Cognition XXII* (pp. 435–459). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199231447.003.0020

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition*, *72*(3), 237–267. https://doi.org/10.1016/S0010-0277(99)00039-6

Custers, R. (2023). Thoughts about actions and outcomes (and what they lead to). *Motivation Science*, *9*(4), 261–271. https://doi.org/10.1037/mot0000306

Custers, R. (2024). The Homunculus in the Room: A Reply to Kunde and Janczyk (2024) and Hommel and Eder (2024). *Motivation Science*, *10*(4), 403–406. https://doi.org/10.1037/mot0000377

de Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *Journal of Open Source Software*, *8*(85), 5351. https://doi.org/10.21105/joss.05351

della Gatta, F., Garbarini, F., Rabuffetti, M., Viganò, L., Butterfill, S. A., & Sinigaglia, C. (2017). Drawn together: When motor representations ground joint actions. *Cognition*, *165*, 53–60. https://doi.org/10.1016/j.cognition.2017.04.008

Dewey, J. A., Pacherie, E., & Knoblich, G. (2014). The phenomenology of controlling a moving object with another person. *Cognition*, *132*(3), 383–397. https://doi.org/10.1016/j.cognition.2014.05.002

Di Paolo, E., & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00163

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behavior*, *22*(1), 1–18. https://doi.org/10.3758/BF03199951

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *The Quarterly Journal of Experimental Psychology Section A*, *36*(1), 29–50. https://doi.org/10.1080/14640748408401502

Dolk, T., & Prinz, W. (2016). What It Takes to Share a Task. In *Shared Representations* (pp. 3–21). Cambridge University Press. https://doi.org/10.1017/CBO9781107279353.002

Eder, A. B., & Dignath, D. (2017). Influence of verbal instructions on effect-based action control. *Psychological Research*, *81*(2), 355–365. https://doi.org/10.1007/s00426-016-0745-6

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. In *Journal of Experimental Psychology: Human Perception and Performance* (Vol. 27, Issue 1, pp. 229–240). https://doi.org/10.1037//0096-1523.27.1.229

Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research*, *68*(2), 138–154. https://doi.org/10.1007/s00426-003-0151-8

Essa, F., Sebanz, N., & Diesendruck, G. (2019). The automaticity of children's imitative group bias. *Cognitive Development*, *52*. https://doi.org/10.1016/j.cogdev.2019.100799

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Flach, R., Osman, M., Dickinson, A., & Heyes, C. (2006). The interaction between response effects during the acquisition of response priming. *Acta Psychologica*, *122*(1), 11–26. https://doi.org/10.1016/j.actpsy.2005.09.001

Formica, S., & Brass, M. (2024). Coordinated social interactions are supported by integrated neural representations. *Social Cognitive and Affective Neuroscience*, *19*(1). https://doi.org/10.1093/scan/nsae089

Galang, C. M., Cracco, E., & Brass, M. (2024). Anticipated imitation of multiple agents. *Cognition*, *249*, 105831. https://doi.org/10.1016/j.cognition.2024.105831

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*(4), 160–165. https://doi.org/10.1016/j.tics.2013.02.002

Gattis, M., Bekkering, H., & Wohlschläger, A. (2002). Goal-directed imitation. In *The Imitative Mind* (pp. 183–205). Cambridge University Press. https://doi.org/10.1017/CBO9780511489969.011

Genschow, O., & Cracco, E. (Eds.). (2025). *Automatic Imitation*. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-62634-0

Genschow, O., Hansen, J., Wänke, M., & Trope, Y. (2019). Psychological distance modulates goal-based versus movement-based imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(8), 1031–1048. https://doi.org/10.1037/xhp0000654

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278. https://doi.org/10.1126/science.aac6076

Gilbert, M. (2009). Shared intention and personal intentions. *Philosophical Studies*, *144*(1), 167–187. https://doi.org/10.1007/s11098-009-9372-z

Gordon, J., Knoblich, G., & Pezzulo, G. (2023). Strategic Task Decomposition in Joint Action. *Cognitive Science*, *47*(7). https://doi.org/10.1111/cogs.13316

Grafton, S. T., & Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, *26*(4), 590–616. https://doi.org/10.1016/j.humov.2007.05.009

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review*, *77*(2), 73–99. https://doi.org/10.1037/h0028689

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational Use of Cognitive Resources: Levels of Analysis Between the Computational and the Algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229. https://doi.org/10.1111/tops.12142

Hassin, R. R., Aarts, H., & Ferguson, M. J. (2005). Automatic goal inferences. *Journal of Experimental Social Psychology*, *41*(2), 129–140. https://doi.org/10.1016/j.jesp.2004.06.008

Hazeltine, E., & Schumacher, E. H. (2016). *Chapter Six - Understanding Central Processes: The Case against Simple Stimulus-Response Associations and for Complex Task Representation* (B. H. Ross, Ed.; Vol. 64, pp. 195–245). Academic Press. https://doi.org/https://doi.org/10.1016/bs.plm.2015.09.006

Herwig, A., Prinz, W., & Waszak, F. (2007). Two modes of sensorimotor integration in intention-based and stimulus-based actions. *Quarterly Journal of Experimental Psychology*, *60*(11), 1540–1554. https://doi.org/10.1080/17470210601119134

Herwig, A., & Waszak, F. (2009). Intention and attention in ideomotor learning. *Quarterly Journal of Experimental Psychology*, *62*(2), 219–227. https://doi.org/10.1080/17470210802373290

Heyes, C. (2011). Automatic Imitation. *Psychological Bulletin*, *137*(3), 463–483. https://doi.org/10.1037/a0022288

Heyes, C. (2013). Imitation: Associative and Context Dependent. In *Action Science* (pp. 309–332). The MIT Press. https://doi.org/10.7551/mitpress/9780262018555.003.0019

Heyes, C., Bird, G., Johnson, H., & Haggard, P. (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, *22*(2), 233–240. https://doi.org/10.1016/j.cogbrainres.2004.09.009

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan. *Nature*, *606*(7912), 129–136. https://doi.org/10.1038/s41586-022-04743-9

Hoffmann, J., Lenhard, A., Sebald, A., & Pfister, R. (2009). Movements or targets: What makes an action in action-effect learning? *Quarterly Journal of Experimental Psychology*, *62*(12), 2433–2449. https://doi.org/10.1080/17470210902922079

Hommel, B. (2013). Ideomotor Action Control: On the Perceptual Grounding of Voluntary Actions and Agents. In *Action Science* (pp. 112–136). The MIT Press. https://doi.org/10.7551/mitpress/9780262018555.003.0005

Hommel, B. (2017). *Goal-Directed Actions* (M. R. Waldmann, Ed.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199399550.013.18

Hommel, B., Alonso, D., & Fuentes, L. J. (2003). Acquisition and generalization of action effects. *Visual Cognition*, *10*(8), 965–986. https://doi.org/10.1080/13506280344000176

Hommel, B., & Eder, A. B. (2024). Modern Ideomotor Theory of Goal-Directed Action Is Not What You Think: A Commentary on Custers (2023). *Motivation Science*, *10*(4), 398–402. https://doi.org/10.1037/mot0000359

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849–878. https://doi.org/10.1017/S0140525X01000103

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical Mechanisms of Human Imitation. *Science*, *286*(5449), 2526–2528. https://doi.org/10.1126/science.286.5449.2526

James, W. (1890). *The Principles of Psychology*. Holt.

Janczyk, M., Giesen, C. G., Moeller, B., Dignath, D., & Pfister, R. (2022). Perception and action as viewed from the Theory of Event Coding: a multi-lab replication and effect size estimation of common experimental designs. *Psychological Research*. https://doi.org/10.1007/s00426-022-01705-8

Janczyk, M., Welsh, T. N., & Dolk, T. (2016). A role of goals for social inhibition of return? *Quarterly Journal of Experimental Psychology*, *69*(12), 2402–2418. https://doi.org/10.1080/17470218.2015.1112417

JASP Team. (2024). *JASP (Version 0.19.0)[Computer software]*. https://jasp-stats.org/

Jeannerod, M. (1999). To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *52*(1), 1–29. https://doi.org/10.1080/713755803

Kaup, B., Ulrich, R., Bausenhart, K. M., Bryce, D., Butz, M. V., Dignath, D., Dudschig, C., Franz, V. H., Friedrich, C., Gawrilow, C., Heller, J., Huff, M., Hütter, M., Janczyk, M., Leuthold, H., Mallot, H., Nürk, H.-C., Ramscar, M., Said, N., … Wong, H. Y. (2024). Modal and amodal cognition: an overarching principle in various domains of psychology. *Psychological Research*, *88*(2), 307–337. https://doi.org/10.1007/s00426-023-01878-w

Keller, P. E., & Koch, I. (2006). Exogenous and endogenous response priming with auditory stimuli. *Advances in Cognitive Psychology*, *2*(4), 269–276. https://doi.org/10.2478/v10053-008-0061-9

Keller, P. E., Novembre, G., & Loehr, J. (2016). Musical Ensemble Performance. In *Shared Representations* (pp. 280–310). Cambridge University Press. https://doi.org/10.1017/CBO9781107279353.015

Keller, P. E., Wascher, E., Prinz, W., Waszak, F., Koch, I., & Rosenbaum, D. A. (2006). Differences Between Intention-Based and Stimulus-Based Actions. *Journal of Psychophysiology*, *20*(1), 9–20. https://doi.org/10.1027/0269-8803.20.1.9

Kiesel, A., & Hoffmann, J. (2004). Variable action effects: Response control by context-specific effect anticipations. *Psychological Research*, *68*(2–3), 155–162. https://doi.org/10.1007/s00426-003-0152-7

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing*, *8*(3), 159–166. https://doi.org/10.1007/s10339-007-0170-2

Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An Interference Effect of Observed Biological Movement on Action. *Current Biology*, *13*(6), 522–525. https://doi.org/10.1016/S0960-9822(03)00165-9

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological Research on Joint Action. Theory and Data. In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 54). https://doi.org/10.1016/B978-0-12-385527-5.00003-6

Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 1006–1016. https://doi.org/10.1037/0278-7393.29.5.1006

Knoblich, G., & Sebanz, N. (2006). The Social Nature of Perception and Action. *Current Directions in Psychological Science*, *15*(3), 99–104. https://doi.org/10.1111/j.0963-7214.2006.00415.x

Kourtis, D., Knoblich, G., Woźniak, M., & Sebanz, N. (2014). Attention Allocation and Task Representation during Joint Action Planning. *Journal of Cognitive Neuroscience*, *26*(10), 2275–2286. https://doi.org/10.1162/jocn_a_00634

Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: Social interaction modulates action simulation. *Biology Letters*, *6*(6), 758–761. https://doi.org/10.1098/rsbl.2010.0478

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience*, *8*(1), 31–42. https://doi.org/10.1080/17470919.2012.694823

Kourtis, D., Woźniak, M., Sebanz, N., & Knoblich, G. (2019). Evidence for we-representations during joint action planning. *Neuropsychologia*, *131*(November 2018), 73–83. https://doi.org/10.1016/j.neuropsychologia.2019.05.029

Kunde, W. (2001a). Exploring the hyphen in ideo-motor action. *Behavioral and Brain Sciences*, *24*(5), 891–892. https://doi.org/10.1017/S0140525X01340104

Kunde, W. (2001b). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 387–394. https://doi.org/10.1037/0096-1523.27.2.387

Kunde, W., & Janczyk, M. (2024). Thoughts about "Thoughts about actions and outcomes": Comment on Custers (2023). In *Motivation Science* (Vol. 10, Issue 4, pp. 392–397). American Psychological Association. https://doi.org/10.1037/mot0000341

Kunde, W., Weller, L., & Pfister, R. (2018). Sociomotor action control. *Psychonomic Bulletin and Review*, *25*(3), 917–931. https://doi.org/10.3758/s13423-017-1316-6

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, *27*(3), 145–162. https://doi.org/10.1023/A:1025389814290

Lange, K., Kühn, S., & Filevich, E. (2015). Correction: "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, *10*(7), e0134073. https://doi.org/10.1371/journal.pone.0134073

Langenberg, B., Janczyk, M., Koob, V., Kliegl, R., & Mayer, A. (2023). A tutorial on using the paired t test for power calculations in repeated measures ANOVA with interactions. *Behavior Research Methods*, *55*(5), 2467–2484. https://doi.org/10.3758/s13428-022-01902-8

Le Bars, S., Bourgeois-Gironde, S., Wyart, V., Sari, I., Pacherie, E., & Chambon, V. (2022). Motor Coordination and Strategic Cooperation in Joint Action. *Psychological Science*, *33*(5), 736–751. https://doi.org/10.1177/09567976211053275

Le Bars, S., Devaux, A., Nevidal, T., Chambon, V., & Pacherie, E. (2020). Agents' pivotality and reward fairness modulate sense of agency in cooperative joint action. *Cognition*, *195*(October 2019), 104117. https://doi.org/10.1016/j.cognition.2019.104117

Leighton, J., Bird, G., & Heyes, C. (2010). 'Goals' are not an integral component of imitation. *Cognition*, *114*(3), 423–435. https://doi.org/10.1016/j.cognition.2009.11.001

Lewis, R. L., Howes, A., & Singh, S. (2014). Computational Rationality: Linking Mechanism and Behavior Through Bounded Utility Maximization. *Topics in Cognitive Science*, *6*(2), 279–311. https://doi.org/10.1111/tops.12086

Loehr, J. D. (2018). Shared credit for shared success: Successful joint performance strengthens the sense of joint agency. *Consciousness and Cognition*, *66*(April), 79–90. https://doi.org/10.1016/j.concog.2018.11.001

Loehr, J. D. (2022). The sense of agency in joint action: An integrative review. In *Psychonomic Bulletin and Review* (Vol. 29, Issue 4, pp. 1089–1117). Springer. https://doi.org/10.3758/s13423-021-02051-3

Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring Individual and Joint Action Outcomes in Duet Music Performance. *Journal of Cognitive Neuroscience*, *25*(7), 1049–1061. https://doi.org/10.1162/jocn_a_00388

Loehr, J. D., & Vesper, C. (2016). The sound of you and me: Novices represent shared goals in joint action. *Quarterly Journal of Experimental Psychology*, *69*(3), 535–547. https://doi.org/10.1080/17470218.2015.1061029

Longo, M. R., Kosobud, A., & Bertenthal, B. I. (2008). Automatic imitation of biomechanically possible and impossible actions: Effects of priming movements versus goals. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(2), 489–501. https://doi.org/10.1037/0096-1523.34.2.489

Marschner, M., Dignath, D., & Knoblich, G. (2024). Me or we? Action-outcome learning in synchronous joint action. *Cognition*, *247*. https://doi.org/10.1016/j.cognition.2024.105785

Marschner, M., Knoblich, G., & Dignath, D. (2025). *Do Individual and Joint Action Goals Modulate Imitative Response Tendencies?* https://doi.org/10.31219/osf.io/3ctsx_v1

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social Connection Through Joint Action and Interpersonal Coordination. *Topics in Cognitive Science*, *1*(2), 320–339. https://doi.org/10.1111/j.1756-8765.2009.01022.x

Massen, C., & Prinz, W. (2007). Activation of Action Rules in Action Observation. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(6), 1118–1130. https://doi.org/10.1037/0278-7393.33.6.1118

McEllin, L., Knoblich, G., & Sebanz, N. (2018). Imitation from a joint action perspective. *Mind and Language*, *33*(4), 342–354. https://doi.org/10.1111/mila.12188

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators. *Science (New York, N.Y.)*, *311*(5765), 1297–1300. https://doi.org/10.1126/science.1123007

Meyer, M., Hunnius, S., Van Elk, M., Van Ede, F., & Bekkering, H. (2011). Joint action modulates motor system involvement during action observation in 3-year-olds. *Experimental Brain Research*, *211*(3–4), 581–592. https://doi.org/10.1007/s00221-011-2658-3

Michael, J., & Pacherie, E. (2015). On Commitments and Other Uncertainty Reduction Tools in Joint Action. *Journal of Social Ontology*, *1*(1), 89–120. https://doi.org/10.1515/jso-2014-0021

Milward, S. J., & Sebanz, N. (2016). Mechanisms and development of self–other distinction in dyads and groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1686). https://doi.org/10.1098/rstb.2015.0076

Miss, F. M., Adriaense, J. E. C., & Burkart, J. M. (2022). Towards integrating joint action research: Developmental and evolutionary perspectives on co-representation.

Neuroscience & Biobehavioral Reviews, *143*, 104924. https://doi.org/10.1016/j.neubiorev.2022.104924

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. In *Behavioral and Brain Sciences* (Vol. 32, Issue 2). https://doi.org/10.1017/S0140525X09000855

Moeller, B., & Frings, C. (2019). From simple to complex actions: Response-response bindings as a new approach to action sequences. *Journal of Experimental Psychology: General*, *148*(1), 174–183. https://doi.org/10.1037/xge0000483

Moeller, B., & Pfister, R. (2022). Ideomotor learning: Time to generalize a longstanding principle. In *Neuroscience and Biobehavioral Reviews* (Vol. 140). Elsevier Ltd. https://doi.org/10.1016/j.neubiorev.2022.104782

Monsell, S. (2003). Task switching. In *Trends in Cognitive Sciences* (Vol. 7, Issue 3, pp. 134–140). Elsevier Ltd. https://doi.org/10.1016/S1364-6613(03)00028-7

Monsell, S., Taylor, T. J., & Murphy, K. (2001). Naming the color of a word: Is it responses or task sets that compete? In *Memory & Cognition* (Vol. 29, Issue 1).

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Müller, R. (2016). Does the anticipation of compatible partner reactions facilitate action planning in joint tasks? *Psychological Research*, *80*(4), 464–486. https://doi.org/10.1007/s00426-015-0670-0

Musco, M. A., Paulesu, E., & Sacheli, L. M. (2025). Social and goal-related foundations of interpersonal adaptation during joint action. *Journal of Experimental Psychology: Human Perception and Performance*, *51*(3), 341–356. https://doi.org/10.1037/xhp0001273

Neszmélyi, B., & Pfister, R. (2024). Anticipated Imitation Is Not Affected by the Number of Imitators. *Experimental Psychology*, *71*(6), 352–359. https://doi.org/10.1027/1618-3169/a000637

Newman-Norlund, R. D., Van Schie, H. T., Van Zuijlen, A. M. J., & Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience*, *10*(7), 817–818. https://doi.org/10.1038/nn1911

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2014). Motor simulation and the coordination of self and other in real-time joint action. *Social Cognitive and Affective Neuroscience*, *9*(8), 1062–1068. https://doi.org/10.1093/scan/nst086

Ocampo, B., & Kritikos, A. (2010). Placing actions in context: Motor facilitation following observation of identical and non-identical manual acts. *Experimental Brain Research*, *201*(4), 743–751. https://doi.org/10.1007/s00221-009-2089-6

Ocampo, B., Kritikos, A., & Cunnington, R. (2011). How frontoparietal brain regions mediate imitative and complementary actions: An fMRI study. *PLoS ONE*, *6*(10). https://doi.org/10.1371/journal.pone.0026945

Ondobaka, S., & Bekkering, H. (2012). Hierarchy of idea-guided action and perception-guided movement. *Frontiers in Psychology*, *3*(DEC). https://doi.org/10.3389/fpsyg.2012.00579

Ondobaka, S., & Bekkering, H. (2013). Conceptual and perceptuo-motor action control and action recognition. In *Cortex* (Vol. 49, Issue 10, pp. 2966–2967). Masson SpA. https://doi.org/10.1016/j.cortex.2013.06.005

Ondobaka, S., de Lange, F. P., Newman-Norlund, R. D., Wiemers, M., & Bekkering, H. (2012). Interplay between action and movement intentions during social interaction. *Psychological Science*, *23*(1), 30–35. https://doi.org/10.1177/0956797611424163

Ondobaka, S., Newman-Norlund, R. D., De Lange, F. P., & Bekkering, H. (2013). Action recognition depends on observer's level of action control and social personality traits. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0081392

Oullier, O., de Guzman, G. C., Jantzen, K. J., Lagarde, J., & Scott Kelso, J. A. (2008). Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, *3*(2), 178–192. https://doi.org/10.1080/17470910701563392

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107*(1), 179–217. https://doi.org/10.1016/j.cognition.2007.09.003

Pacherie, E. (2012). The Phenomenology of Joint Action: Self-Agency versus Joint Agency. In A. Seemann (Ed.), *Joint Attention*. The MIT Press. https://doi.org/10.7551/mitpress/8841.003.0017

Pacherie, E. (2013). Intentional joint agency: Shared intention lite. *Synthese*, *190*(10), 1817–1839. https://doi.org/10.1007/s11229-013-0263-7

Pacherie, E. (2014). How does it feel to act together? *Phenomenology and the Cognitive Sciences*, *13*(1), 25–46. https://doi.org/10.1007/s11097-013-9329-8

Pavlov, I. (1927). Conditioned reflexes. In *Conditioned reflexes*. Oxford University Press.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pesquita, A., Whitwell, R. L., & Enns, J. T. (2018). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychonomic Bulletin and Review*, *25*(5), 1751–1769. https://doi.org/10.3758/s13423-017-1393-6

Pezzulo, G., Iodice, P., Donnarumma, F., Dindo, H., & Knoblich, G. (2017). Avoiding Accidents at the Champagne Reception: A Study of Joint Lifting and Balancing. *Psychological Science*, *28*(3), 338–345. https://doi.org/10.1177/0956797616683015

Pezzulo, G., Knoblich, G., Maisto, D., Donnarumma, F., Pacherie, E., & Hasson, U. (2025). *A Predictive Processing Framework for Joint Action and Communication*. https://doi.org/10.31234/osf.io/q4jnr_v1

Pfister, R. (2019). *Effect-Based Action Control With Body-Related Effects : Implications for Empirical Approaches to Ideomotor Action Control*. *126*(1), 153–161.

Pfister, R., Dignath, D., Hommel, B., & Kunde, W. (2013). It Takes Two to Imitate: Anticipation and Imitation in Social Interaction. *Psychological Science*, *24*(10), 2117–2121. https://doi.org/10.1177/0956797613489139

Pfister, R., Dolk, T., Prinz, W., & Kunde, W. (2014). Joint response-effect compatibility. *Psychonomic Bulletin and Review*, *21*(3), 817–822. https://doi.org/10.3758/s13423-013-0528-7

Pfister, R., Kiesel, A., & Hoffmann, J. (2011). Learning at any rate: Action-effect learning for stimulus-based actions. *Psychological Research*, *75*(1), 61–65. https://doi.org/10.1007/s00426-010-0288-1

Pfister, R., Kiesel, A., & Melcher, T. (2010). Adaptive control of ideomotor effect anticipations. *Acta Psychologica*, *135*(3), 316–322. https://doi.org/10.1016/j.actpsy.2010.08.006

Pfister, R., & Kunde, W. (2013). Dissecting the response in response–effect compatibility. *Experimental Brain Research*, *224*(4), 647–655. https://doi.org/10.1007/s00221-012-3343-x

Pfister, R., Weller, L., Dignath, D., & Kunde, W. (2017). What or when? The impact of anticipated social action effects is driven by action-effect compatibility, not delay. *Attention, Perception, and Psychophysics*, *79*(7), 2132–2142. https://doi.org/10.3758/s13414-017-1371-0

Poljac, E., Van Schie, H. T., & Bekkering, H. (2009). Understanding the flexibility of action-perception coupling. *Psychological Research*, *73*(4), 578–586. https://doi.org/10.1007/s00426-009-0238-y

Press, C., Bird, G., Walsh, E., & Heyes, C. (2008). Automatic imitation of intransitive actions. *Brain and Cognition*, *67*(1), 44–50. https://doi.org/10.1016/j.bandc.2007.11.001

Press, C., Gillmeister, H., & Heyes, C. (2007). Sensorimotor experience enhances automatic imitation of robotic action. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1625), 2509–2514. https://doi.org/10.1098/rspb.2007.0774

Prinz, W. (1990). A Common Coding Approach to Perception and Action. In *Relationships Between Perception and Action* (pp. 167–201). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-75348-0_7

Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. https://doi.org/10.1080/713752551

Ramenzoni, V. C., Sebanz, N., & Knoblich, G. (2014). Scaling up perception-action links: Evidence from synchronization with individual and joint action. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1551–1565. https://doi.org/10.1037/a0036925

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

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*(9), 661–670. https://doi.org/10.1038/35090060

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nature Reviews Neuroscience*, *17*(12), 757–765. https://doi.org/10.1038/nrn.2016.135

Rocca, M., Sacheli, L. M., Romeo, L., & Cavallo, A. (2023). Visuo-motor interference is modulated by task interactivity: A kinematic study. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02297-z

Rogers, R. D., & Monsell, S. (1995). Costs of a Predictable Switch Between Simple Cognitive Tasks. In *Journal of Experimental Psychology: General* (Vol. 124, Issue 2).

Sacheli, L. M., Aglioti, S. M., & Candidi, M. (2015). Social cues to joint actions: The role of shared goals. *Frontiers in Psychology*, *6*(JUL), 1–7. https://doi.org/10.3389/fpsyg.2015.01034

Sacheli, L. M., Arcangeli, E., & Paulesu, E. (2018). Evidence for a dyadic motor plan in joint action. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-23275-9

Sacheli, L. M., Musco, M. A., Zazzera, E., Banfi, G., & Paulesu, E. (2022). How shared goals shape action monitoring. *Cerebral Cortex (New York, N.Y. : 1991)*, *32*(21), 4934–4951. https://doi.org/10.1093/cercor/bhac019

Sacheli, L. M., Musco, M. A., Zazzera, E., & Paulesu, E. (2021). Mechanisms for mutual support in motor interactions. *Scientific Reports*, 1–16. https://doi.org/10.1038/s41598-021-82138-y

Sacheli, L. M., Verga, C., Arcangeli, E., Banfi, G., Tettamanti, M., & Paulesu, E. (2019). How task interactivity shapes action observation. *Cerebral Cortex*, *29*(12), 5302–5314. https://doi.org/10.1093/cercor/bhz205

Sahaï, A., Desantis, A., Grynszpan, O., Pacherie, E., & Berberian, B. (2019). Action co-representation and the sense of agency during a joint Simon task: Comparing human and machine co-agents. *Consciousness and Cognition*, *67*, 44–55. https://doi.org/10.1016/j.concog.2018.11.008

Sartori, L., & Betti, S. (2015). Complementary actions. In *Frontiers in Psychology* (Vol. 6, Issue MAY). Frontiers Media S.A. https://doi.org/10.3389/fpsyg.2015.00557

Sartori, L., Betti, S., Chinellato, E., & Castiello, U. (2015). The multiform motor cortical output: Kinematic, predictive and response coding. *Cortex*, *70*, 169–178. https://doi.org/10.1016/j.cortex.2015.01.019

Sartori, L., Bucchioni, G., & Castiello, U. (2013). When emulation becomes reciprocity. *Social Cognitive and Affective Neuroscience*, *8*(6), 662–669. https://doi.org/10.1093/scan/nss044

Sartori, L., Cavallo, A., Bucchioni, G., & Castiello, U. (2012). From simulation to reciprocity: The case of complementary actions. *Social Neuroscience*, *7*(2), 146–158. https://doi.org/10.1080/17470919.2011.586579

Schmidt, R. C., Fitzpatrick, P., Caron, R., & Mergeche, J. (2011). Understanding social motor coordination. *Human Movement Science*, *30*(5), 834–845. https://doi.org/10.1016/j.humov.2010.05.014

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of Interpersonal Coordination. In *Coordination: Neural, Behavioral and Social Dynamics* (pp. 281–308). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-74479-5_14

Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task constraints in joint action. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(8), 1480–1493. https://doi.org/10.1037/xhp0000403

Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2018). Co-actors represent the order of each other's actions. *Cognition*, *181*(November 2017), 65–79. https://doi.org/10.1016/j.cognition.2018.08.008

Schmitz, L., Wahn, B., & Krüger, M. (2023). Attention allocation in complementary joint action: How joint goals affect spatial orienting. *Attention, Perception, and Psychophysics*. https://doi.org/10.3758/s13414-023-02779-1

Seabrooke, T., Hogarth, L., & Mitchell, C. J. (2016). The propositional basis of cue-controlled reward seeking. *Quarterly Journal of Experimental Psychology*, *69*(12), 2452–2470. https://doi.org/10.1080/17470218.2015.1115885

Searle, J. (1990). Collective Intentions and Actions. In P. R. C. Jerry Morgan & M. Pollack (Eds.), *Intentions in Communication*. MIT Press.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Sebanz, N., & Knoblich, G. (2009). Prediction in Joint Action: What, When, and Where. *Topics in Cognitive Science*, *1*(2), 353–367. https://doi.org/10.1111/j.1756-8765.2009.01024.x

Sebanz, N., & Knoblich, G. (2021). Progress in Joint-Action Research. *Current Directions in Psychological Science*. https://doi.org/10.1177/0963721420984425

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition*, *88*(3), B11–B21. https://doi.org/10.1016/S0010-0277(03)00043-X

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1234–1246. https://doi.org/10.1037/0096-1523.31.6.1234

Shanks, D. R. (1989). Selectional processes in causality judgment. *Memory & Cognition*, *17*(1), 27–34. https://doi.org/10.3758/BF03199554

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A Review of Contemporary Ideomotor Theory. *Psychological Bulletin*, *136*(6), 943–974. https://doi.org/10.1037/a0020541

Sinigaglia, C., & Butterfill, S. A. (2022). Motor representation in acting together. *Synthese*, *200*(2). https://doi.org/10.1007/s11229-022-03539-8

Strasser, A. (2015). Can Artificial Systems Be Part of a Collective Action? In *Collective Agency and Cooperation in Natural and Artificial Systems* (pp. 205–218). Springer International Publishing. https://doi.org/10.1007/978-3-319-15515-9_11

Strasser, A. (2018). *Social Cognition and Artificial Agents* (pp. 106–114). https://doi.org/10.1007/978-3-319-96448-5_12

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1746–1759. https://doi.org/10.1037/0096-1523.26.6.1746

Sun, D., Custers, R., Marien, H., & Aarts, H. (2020). Ideomotor Action: Evidence for Automaticity in Learning, but Not Execution. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00185

Sun, D., Custers, R., Marien, H., Liefooghe, B., & Aarts, H. (2022). Examining Mechanistic Explanations for Ideomotor Effects. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(5), 458–466. https://doi.org/10.1037/xhp0000994

Thompson, E. L., Bird, G., & Catmur, C. (2019). Conceptualizing and testing action understanding. *Neuroscience & Biobehavioral Reviews*, *105*, 106–114. https://doi.org/10.1016/j.neubiorev.2019.08.002

Tollefsen, D., & Dale, R. (2012). Naturalizing joint action: A process-based approach. *Philosophical Psychology*, *25*(3), 385–407. https://doi.org/10.1080/09515089.2011.579418

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005a). In Search of the Uniquely Human. *Behavioral and Brain Sciences*, *28*(5), 721–727. https://doi.org/10.1017/S0140525X05540123

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005b). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019). Rationality in Joint Action: Maximizing Coefficiency in Coordination. *Psychological Science*, *30*(6), 930–941. https://doi.org/10.1177/0956797619842550

Török, G., Stanciu, O., Sebanz, N., & Csibra, G. (2020). Joint action planning : co-actors minimize the aggregate individual costs of actions. *42nd Annual Virtual Meeting of the Cognitive Science Society*, 295–301.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory: Vol. Vol. 3* (pp. 200–239). Erlbaum.

Tsai, J. C. C., Kuo, W. J., Jing, J. T., Hung, D. L., & Tzeng, O. J. L. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research*, *175*(2), 353–362. https://doi.org/10.1007/s00221-006-0557-9

Tsai, J. C. C., Sebanz, N., & Knoblich, G. (2011). The GROOP effect: Groups mimic group actions. *Cognition*, *118*(1), 135–140. https://doi.org/10.1016/j.cognition.2010.10.007

Tuomela, R. (2006). Joint Intention, We-Mode and I-Mode. *Midwest Studies in Philosophy*, *30*(1), 35–58.

Uithol, S., van Rooij, I., Bekkering, H., & Haselager, P. (2012). Hierarchies in Action and Motor Control. *Journal of Cognitive Neuroscience*, *24*(5), 1077–1086. https://doi.org/10.1162/jocn_a_00204

Vallacher, R. R., & Wegner, D. M. (1985). *A Theory of Action Identification*. Psychology Press. https://doi.org/10.4324/9781315802213

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, *94*(1), 3–15. https://doi.org/10.1037/0033-295X.94.1.3

Vallacher, R. R., & Wegner, D. M. (2012). Action Identification Theory. In *Handbook of Theories of Social Psychology: Volume 1* (pp. 327–348). SAGE Publications Ltd. https://doi.org/10.4135/9781446249215.n17

van der Wel, R. P. R. D., Becchio, C., Curioni, A., & Wolf, T. (2021). Understanding joint

    action: Current theoretical and empirical approaches. *Acta Psychologica*, *215*(March).

    https://doi.org/10.1016/j.actpsy.2021.103285

van Schie, H. T., van Waterschoot, B. M., & Bekkering, H. (2008). Understanding Action

    Beyond Imitation: Reversed Compatibility Effects of Action Observation in Imitation and

    Joint Action. *Journal of Experimental Psychology: Human Perception and Performance*,

    34(6), 1493–1500. https://doi.org/10.1037/a0011750

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint

    action. *Neural Networks*, *23*(8–9), 998–1003.

    https://doi.org/10.1016/j.neunet.2010.06.002

Vesper, C., Knoblich, G., & Sebanz, N. (2014). Our actions in my mind: Motor imagery of joint

    action. *Neuropsychologia*, *55*(1), 115–121.

    https://doi.org/10.1016/j.neuropsychologia.2013.05.024

Vesper, C., & Sebanz, N. (2016). Acting Together: Representations and Coordination

    Processes. In *Shared Representations* (pp. 216–235).

    https://doi.org/10.1017/cbo9781107279353.012

Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2013). Are you ready to

    jump? Predictive mechanisms in interpersonal coordination. *Journal of Experimental*

    Psychology: Human Perception and Performance, *39*(1), 48–61.

    https://doi.org/10.1037/a0028066

Vogel, D., Rudolf, M., & Scherbaum, S. (2020). The impact of the verbal instruction and task

    characteristics on effect-based action control. *Cognitive Processing*, *21*(2), 271–285.

    https://doi.org/10.1007/s10339-020-00960-0

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.*

    Harvard University Press.

Waszak, F., Wascher, E., Keller, P., Koch, I., Aschersleben, G., Rosenbaum, D. A., & Prinz, W. (2005). Intention-based and stimulus-based mechanisms in action selection. *Experimental Brain Research*, *162*(3), 346–356. https://doi.org/10.1007/s00221-004-2183-8

Watson, P., Steenbergen, H. Van, Wit, S. De, Wiers, R. W., & Hommel, B. (2015). *Limits of ideomotor action – outcome acquisition*. *1626*, 45–53. https://doi.org/10.1016/j.brainres.2015.02.020

Weller, L., Pfister, R., & Kunde, W. (2019). Sociomotor actions: Anticipated partner responses are primarily represented in terms of spatial, not anatomical features. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(8), 1104–1118. https://doi.org/10.1037/xhp0000658

Welsh, T. N., Elliott, D., Anson, J. G., Dhillon, V., Weeks, D. J., Lyons, J. L., & Chua, R. (2005). Does Joe influence Fred's action? *Neuroscience Letters*, *385*(2), 99–104. https://doi.org/10.1016/j.neulet.2005.05.013

Wenke, D., Atmaca, S., Holländer, A., Liepelt, R., Baess, P., & Prinz, W. (2011). What is Shared in Joint Action? Issues of Co-representation, Response Conflict, and Agent Identification. *Review of Philosophy and Psychology*, *2*(2), 147–172. https://doi.org/10.1007/s13164-011-0057-0

Wild, K. S., Poliakoff, E., Jerrison, A., & Gowen, E. (2010). The influence of goals on movement kinematics during imitation. *Experimental Brain Research*, *204*(3), 353–360. https://doi.org/10.1007/s00221-009-2034-8

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460–473. https://doi.org/10.1037/0033-2909.131.3.460

Wit, S. De, & Dickinson, A. (2009). *Associative theories of goal-directed behaviour : a case for animal – human translational models*. 463–476. https://doi.org/10.1007/s00426-009-0230-6

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 501–515. https://doi.org/10.1098/rstb.2002.1257

Wolfensteller, U., & Ruge, H. (2011). On the timescale of stimulus-based action-effect learning. *Quarterly Journal of Experimental Psychology*, *64*(7), 1273–1289. https://doi.org/10.1080/17470218.2010.546417

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 593–602. https://doi.org/10.1098/rstb.2002.1238

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732. https://doi.org/10.1016/S0960-9822(01)00432-8

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*(1), 1–34. https://doi.org/10.1016/S0010-0277(98)00058-4

Zapparoli, L., Paulesu, E., Mariano, M., Ravani, A., & Sacheli, L. M. (2022). The sense of agency in joint actions: A theory-driven meta-analysis. *Cortex*, *148*, 99–120. https://doi.org/10.1016/j.cortex.2022.01.002

Zhou, Z., Christensen, J., Cummings, J. A., & Loehr, J. D. (2023). Not just in sync: Relations between partners' actions influence the sense of joint agency during joint action. *Consciousness and Cognition*, *111*, 103521. https://doi.org/10.1016/j.concog.2023.103521

Zwosta, K., Ruge, H., & Wolfensteller, U. (2013a). No anticipation without intention: Response–effect compatibility in effect-based and stimulus-based actions. *Acta Psychologica*, *144*(3), 628–634. https://doi.org/10.1016/j.actpsy.2013.09.014

Zwosta, K., Ruge, H., & Wolfensteller, U. (2013b). No anticipation without intention: Response-effect compatibility in effect-based and stimulus-based actions. *Acta Psychologica*, *144*(3), 628–634. https://doi.org/10.1016/j.actpsy.2013.09.014